Guardian Angel:
Patient-Centered
Health Information Systems

Peter Szolovits, Jon Doyle, William J. Long
Laboratory for Computer Science
Massachusetts Institute of Technology

Isaac Kohane
The Children's Hospital

Stephen G. Pauker
New England Medical Center

May 1994

Introduction

Technical reports normally concern the *results* of research project, not their plans. Nevertheless, we believe it important sometimes that new ideas, reflected in a proposal to carry out a research project, be exposed to the scrutiny and suggestions of our colleagues. This serves both to let them teach us valuable lessons in criticizing our plans and (we hope) to sway them toward our views on what constitutes fruitful research directions. This report contains most of the text of a proposal we submitted in March, 1994 to ARPA in response to BAA 94-13, for the Health Information Infrastructure Program. This report differs from the proposal only in its organization and in omitting budgetary and administrative details. To a large extent, we have retained the language typical of proposals, and hope for the reader's patience with its tone.

The mid-1990's promise to usher in an era of comprehensive medical record keeping and the beginning of intelligent computer-based agents that can exploit such records to improve the quality or reduce the cost of health care. Most current efforts to achieve these goals center on putting in place ideas that have been proposed over the past several decades, but made more feasible by declining costs of computing and made more urgent by our society's determination to study the outcomes of medical interventions and to restrict funding to only those interventions that seem most clearly worthwhile. Such efforts typically center on building clinical information systems for hospitals, clinics, individual health care providers, insurers, and government reviewers.

Our proposal comes from an attempt to rethink the sources of possible leverage in improving health care, and focusing on the role of the patient in maintaining his or her health and responding to disease. Thus, we plan to build systems to support the health information needs of the consumers of health care rather than its providers. This shift of focus will, we hope, make possible dramatic improvements in health maintenance and health care delivery, by making useful information available in a timely manner to the patient. The shift also allows us to look at many of the technical issues of how to collect, maintain and interpret comprehensive health information from a very different viewpoint, and might therefore speed the advance of medical informatics.

The review of this proposal was favorable, and although we have as of now no assurance of funding, we are very much committed to pursuing the project and exploring its societal and technical consequences.

Peter Szolovits
May 1994

Part I

Lifetime Patient-Centered Health Information Systems

Current health information systems are built for the convenience of health care providers and consequently yield fragmented patient records in which medically relevant lifelong information is sometimes incomplete, incorrect, or inaccessible. We are constructing information systems centered on the individual patient instead of the provider, in which a set of "guardian angel" (GA) software agents integrates all health-related concerns, including medically-relevant legal and financial information, about an individual (its "subject"). This personal system will help track, manage, and interpret the subject's health history, and offer advice to both patient and provider. Minimally, the system will maintain comprehensive, cumulative, correct, and coherent medical records, accessible in a timely manner as the subject moves through life, work assignments, and health care providers. Each GA is an active process that performs several important functions: it collects patient data; it checks, interprets, and explains to the subject medically-relevant facts and plans; it adapts its advice based on the subject's prior experiences and stated preferences; it performs "sanity checks" on both medical efficacy and cost-effectiveness of diagnostic conclusions and therapeutic plans; it monitors progress; it interfaces to software agents of providers, insurers, etc.; and it helps educate, encourage, and inform the patient. All this serves to improve the quality of medical decision-making, increase patient compliance, and minimize iatrogenic disease and medical errors.

The "Guardian Angel" Concept

We propose a major shift of primary focus away from information systems based on the hospital, clinic and medical practice, to one based on the individual. Such a system, which we call "Guardian Angel" (GA), integrates over a lifetime all health-related information about an individual (its "subject"), thus providing, at minimum, a comprehensive medical record that is often virtually impossible to reconstruct in a timely manner as the subject moves through life and work assignments. But GA is to be not merely a passive repository of information, but an active process that:

(a) engages in data collection, sometimes by interacting with the subject and sometimes by automatic tracking and recording of instruments,

(b) monitors the progress of medical conditions and the effects of therapy with respect to expectations, and checks for side effects,

(c) interprets facts and medically-related plans and helps explain them to the individual,

(d) allows the patient to customize therapy plans within bounds established by care providers, giving the patient "ownership" of his or her therapy,

(e) performs "sanity checks" on the appropriateness of diagnostic conclusions and therapeutic plans,

(f) contains some understanding of the subjects' preferences, represents these in a broad range of negotiations with other systems, including setting therapeutic guidelines and scheduling appointments, and also uses these in structuring interactions between GA and the patient,

(g) interfaces to information systems used by care providers, insurers, researchers, etc., to provide access to personal medical history information as authorized by the individual, and

(h) provides patient education functions, including access to general and specialized medical encyclopedias, and explanations of diagnostic findings and therapy plans specific to the individual,

(i) implements patient reminding and alerting functions, including reminders of scheduled therapy, medications and appointments, integrating these with personal scheduling tools,

(j) provides patient support functions such as contacts with support groups and other patients, queries to pharmaceutical companies, government agencies, etc.

We believe, with many others, that active monitoring of accurate and comprehensive information about an individual's lifetime medical history can greatly improve the quality of medical decision making for that person, reduce errors in health care, and allow people to make better personal decisions and commitments about their care.[1] Although experimental evidence to support this widely-held view is scant, recent reports in the areas of diabetes [DCCT93, Lask93] and hypertension management both strongly indicate the value of active, continual management.

In addition, we believe that there are dramatic improvements to be gained in both the effectiveness and the efficiency of health care if we can empower the user to take a much more active role in monitoring his or her own health status and care, and to take greater responsibility for making informed and guided decisions concerning that care.

Principal benefits for doctors and other health-care providers include access to accurate, comprehensive data, the opportunity to be alerted to changes in the patient's health that are either dangerous in themselves or deviate from an expected course of therapy, and the ability to communicate reliably with the patient.

Although in the long run we envision GA as a routine health assistant for everyone, whether they are ill or well, in the short term it will be easier to apply the architecture as adapted for high-intensity medical interventions, for specific populations of patients who are undergoing active and complex therapy. In such circumstances, some of the monitoring of that therapy can be offloaded to the patient, with the help of GA, making the patient, the "human in the loop". For example, patients with chronic conditions such as diabetes can help to monitor and adjust their own care, and the ambulatory patient undergoing acute care such as chemotherapy may be able to judge certain aspects of his or her own care and tune them, with the aid of GA, for best results. GA could also help make possible the scheduling of visits to the care provider based on the patient's actual response to care rather than on a general schedule. Most uses of GA could benefit greatly from the development and deployment of small, non-invasive sensors that can autonomously and continually monitor characteristics of the patient that are of vital concern.

Architecture for "Guardian Angels"

The objective of current medical information systems is to record and use only limited classes of information tailored to the operations of specific organizations. We propose to develop a highly flexible information flow architecture for Guardian Angels, one that permits variation of the timing, hardware and functionality of computations as the patient's situation permits. GAs will run on a wide and evolving range of hardware, initially personal digital assistants (PDA's), personal computers (PC's) or workstations, and eventually credit-card or dogtag-sized portable computers. They will employ distributed networked facilities to acquire and deliver information, relevant software modules, and access to specialized services, and to provide secure and reliable backup. We will also develop an extensible ontology for medical knowledge adequate to capturing patient and provider data, using the UMLS metathesaurus[UMLS91] as a point of departure, and an interchange language based on this ontology and on ARPA efforts such as KIF [Gene92] and the KQML [Baal94] interchange mechanisms.

---

[1]For example, investigators have been able to demonstrate early detection of disorders such as malnutrition, tumor-associated acquired growth hormone deficiency, or inflammatory disease increases the likelihood of earlier, less costly intervention and with lesser long term morbidity [Bith92].

Another group has shown that tracking the longitudinal health status of children at risk has been recognized as an important tool in improving the health status of these underprivileged children [Will92].

Guardian Angel Interactions and Interoperations

Current provider-oriented medical information systems do little to facilitate the intelligent use of health care resources by patients and by other organizations. We are designing the GA architecture to facilitate (within strict bounds of confidentiality) a wide variety of interactions with other software agents, including interactions aimed at scheduling medical procedures, polling by public health authorities to detect or delineate epidemics or contaminations, facilitating drug "recalls", providing access to information infrastructure resources, and enabling patient support groups.

Experimental Guardian Angel Applications

We plan to apply the GA concept to creation of a small number of prototype systems for use by narrowly-defined patient populations who normally require relatively high-intensity out-patient care. With our partners, we are selecting among the following medical conditions: insulin-dependent diabetes, hypertension, angina, chronic anticoagulation, chronic renal disease, and pulmonary disease (COPD). These applications will also provide us a testbed in which to investigate both technical aspects of the system and the efficacy of its application to these particular populations of patients, and will provide a natural locus of collaboration among the participants.

Example Scenario

The following extended scenario illustrates our vision of the kinds of life-long patient-centered care that the GA concept can foster. The scenario (due to Dr. Kohane) assumes a specific type of hardware and communication architecture that is just one of many possibilities we plan to explore, but it does suggest the functionality we seek:

Abby Kaye is a 14 year old girl who has had insulin-dependent diabetes mellitus for the last five years. Two years ago she began to measure her blood glucose and administer insulin without direct parental supervision. She has recently been enrolled as a participant in the Guardian Angel program.

She wakes this morning, measures her blood sugar which is 180 mg/dl. This is automatically downloaded into the PDA which is connected to her glucometer. GA_PDA stores an internal note to itself to remind Abby tonight that her blood sugar has been greater than 150 mg/dl before breakfast for 5 out of the last 7 days. Perhaps she should increase her nighttime dose of long acting (NPH) insulin? However, Abby is about to give her morning dose and GA_PDA does not distract her with this information at this time.

Abby is uncertain what insulin dose to give this morning as she has a double session dance class at 10:00 and she remembers all too well that she has had mild hypoglycemic symptoms towards the end of even single session dance classes. She draws an exercise symbol spanning 10 to 11:30 on her daily schedule on the GA_PDA interface and then selects the Advise Dose icon. The GA_PDA informs her that she can either keep the dose unchanged if she thinks she can manage a double carbohydrate snack before the dance class or she can reduce her morning dose of insulin by two units of short acting (regular) insulin.

Two weeks later, Abby's morning blood sugars have continued to climb and they are in the mid 200's. Abby has not modified the GA_PDA default authorization to communicate with her parents' desk-top computer. Therefore, when according to schedule, the weekly blood sugars are uploaded by the GA_PDA component to the GA_home_computer component this worrisome trend is displayed to the parents. The GA_home_computer makes several suggestions as to how to modify the nightly NPH. If the parents do not feel comfortable following these suggestions, they are given several options for communications with the health-care providers including electronic mail with attached measurements, directions for paging or an emergency phone number. If they are comfortable with the suggestions, then the recent data along with the alert are uploaded during the routine daily modem connection of GA_home_computer with the GA_hospital_process running on the hospital's information system.

Five months later: Abby has just come back from school upset because one of her friends teased her about becoming chubby. She steps on the bathroom scale and notes her weight; she has gained 10 pounds since her last doctor visit 4 months ago. She has already tried skipping her snacks but that leaves her feeling awful from hypoglycemia (and even hungrier) subsequently. Perhaps, the GA_PDA can help. She taps on the diet icon for advice. The

GA_PDA is aware of several unexplained instances of hypoglycemia in the past month. The pattern of insulin dosing and recorded periods of exercise makes it most likely that these hypoglycemic episodes are due to skipped snacks. The GA_PDA makes a note to itself to use the opportunity of a query about diet to ask her at the end of the interaction whether she has been missing snacks and to warn her about the dangers. The GA_PDA is also aware of the heights and weights obtained at the doctor's office in the last two visits (and downloaded from the IHIS by the GA_hospital_process to the GA_home_computer and then to the GA_PDA) and recognizes that Abby has indeed an increased weight for height over the past visits. GA_PDA asks Abby if she would like to review her meals and snacks for the past days as well as go over her favorite foods. She selects from GA_PDA's large store of food types in quickly generating this description. GA_PDA then identifies the highest caloric items in her diet and proposes several lower calorie substitutions. This information is also passed along to Abby's parents via the GA_home_computer along with an optional shopping list for the grocery store.

The next day, GA_PDA asks Abby if she would be interested in joining a group of children with diabetes to discuss the problems and difficulties of managing diabetes. The discussion can take place via the GA_PDA or GA_home_computer which provides Abby with an interface to her peers over the Internet. Eventually Abby meets some of her Internet group friends at a summer diabetes camp. The discussion group become a way for them to keep in touch during the rest of the year when they live in different cities. Abby's parents also use the GA_home_computer to stay in touch with other parents of children with diabetes and to remain abreast of any developments in research which might make diabetes easier to manage.

Two months later: Abby's father realizes that next week Abby will be performing with her entire dance class for a school performance. However, this performance occurs the same day as a scheduled visit with Abby's endocrinologist. Abby's father goes to the weekly schedule view of the GA_home_computer and cancels the endocrinologist visit. The GA_home_computer asks him if he wants to reschedule (he does) and then negotiates with the hospital scheduling system (via the GA_hospital_process) two possible appointments for Abby in the coming month. After one is selected, the GA_hospital_process sends appropriate notifications to the hospital-based care providers.

Another month has passed. After speaking with her endocrinologist during the last clinic visit, Abby was pleased to hear that her average blood sugar, indirectly measured from an assay of glycosylated hemoglobin, had fallen since the previous visit. She resolves to build upon this progress by "tightening" the control of her blood sugar. However, the GA notes that there has been a downward trend in her pre-lunch and pre-dinner blood sugars and if it continues she will become dangerously hypoglycemic. It suggest that she decrease her morning insulin dose. On their last interaction on diet, the GA_PDA became aware that Abby was already ingesting too many calories, so it does not suggest any increase in snacks. Despite this suggestion, Abby does not decrease her dose of insulin.

The next day, before the GA_PDA has a chance to upload this information to GA_home_computer, Abby goes to school. She checks her blood glucose early, at 11:00 a.m. because she is feeling abdominal cramps and weak. The glucometer communicates the value of 22 mg/dl back to the GA_PDA. The GA_PDA immediately makes a loud sound and displays on its screen a request for Abby to drink orange juice or eat a fruit or candy. When Abby does not acknowledge the alert after two minutes, GA_PDA sends a wireless message to GA_home_computer. If there is no response, it also sends a message to GA_hospital_process. Both GA instances have as their first priority notifying the parents and school care takers. The diabetes nurse practitioner is paged with a message if none of the GA instances have received any acknowledgment from humans.

In fact, Abby had not passed out but had found a candy bar. She was too miserable to work with the GA_PDA and therefore was only able to acknowledge after 3 minutes. The GA_PDA then sends a reassuring notation to the other GA instances. The diabetes nurse practitioner has a telephone conversation with Abby's parents about the need to moderate her target blood sugar levels.

Six months later the desktop computer running GA_home_computer has a serious disk crash. All data is lost without local backups. After re-initializing the disk, the GA software is re-installed and the GA_home_computer automatically rebuilds its data structures after a brief communication with GA_hospital_process (for the exhaustive long-term record) and with GA_PDA for the most recent data.

Two months later: GA_home_computer identifies a cyclic rise in blood sugars that has been occurring with periodicity of 30 to 45 days. The rises last 4 to 7 days before resolving. Insulin adjustments are typically too late to catch these trends from raising the average blood sugar. The GA_home_computer asks whether Abby has started to have menstrual periods. If she has, it recommends increased vigilance for a rise in blood sugar towards the expected end of the cycle so that insulin dose can be raised more responsively. It offers Abby the option of having GA_PDA provide early alerts for this rise.

After the hospital obtains parental consent, Abby's full clinical data set is sent by GA_hospital_process to the data-base where a large prospective multi-center trial is being conducted for an antihypertensive drug to reduce the incidence of renal disease associated with diabetes. After reviewing the data to see if she meets criteria, the GA_hospital_process receives an invitation to join the study which it passes on to GA_home_computer. After reviewing a few lay articles (retrieved by GA_home_computer) and a discussion with Abby's endocrinologist, Abby's parents decline to enroll her because the drug appears to have been insufficiently tested with children. Meanwhile, those adult patients who are already enrolled and have a GA are given an added software component to their GA instances which allows them to report adverse events while on this test drug. These adverse events are then directly reported to the drug manufacturer, the FDA and the patient's clinician and clinical researchers.

Part II

Objectives

Our work includes four major components: (1) design of the Guardian Angel architecture, (2) selection, implementation, and integration of the components of GA that are to be used in our experiments and that will form reference implementations for future use and dissemination, (3) final selection of the medical domains in which we will conduct our experiments, selection of specific sites, patient groups, and goals for the tests, and knowledge engineering of the medical content of the GA tools specialized to the selected domains, and (4) conduct of limited experimental tests of the system to allow us to assess the validity of our designs. The first three of these activities will take place concurrently, although we describe them separately below.

Architecture

We are defining GA initial reference architecture by developing an ontology of the entities, interactions, and information involved in GA activities, as elaborated below.

We will complete and refine the domain model to describe, using a taxonomic language, all the potentially significant entities, their relationships, and their possible interactions, as in the DSSA methodology.

We will identify the types of software, hardware, and communications systems used in the entities with which the GA is to interact.

We will complete and refine the functional model to describe, using a taxonomic language, the inputs, outputs, physical or mental states, and operations or functions of the GA and other entities in the domain. Similarly, we will identify and describe all the specializations of these operations for different specific classes of entities in the domain model.

We will formulate a list of all the types of questions one might pose to the general GA system, and all the types of questions it might pose to other entities, according to the type of the entities. We will formulate similar lists for the additional questions specific to the medical foci applications (diabetes, hypertension, anticoagulation, etc.).

We will prepare or obtain a taxonomic description of medical concepts and relations relevant to the foci applications, making use where possible of existing resources such as the UMLS[UMLS91], and GALEN[GALE93] systems.

Reference Implementation

We are identifying the representation methodology and taxonomic language for use in the GA prototype, and implement methods for translating between the chosen representation system and KIF or other representation standards.

We will identify commercial platforms and applications for the various routine parts of the prototype, i.e., communication, database, scheduling software, etc., making use where possible of systems already exploited by the other entities in the GA's environment.

We will identify initial hardware platforms for both a mobile GA (the GA_PDA of the extended scenario) and a fixed GA (the GA_home_computer of the scenario).

We will identify initial database systems for use in making the system persistent and reliable, such as the THOR system[Lisk90] for stable repositories under development by Barbara Liskov at MIT.

We will identify the initial clinical information systems with which the prototype GA will interact.

We will develop an initial backup strategy and implementation to assure continued integrity of GA's data.

We will build prototypes of the various components of GA, including prototypes of the user interfaces presented to both the patient and health care providers.

We will integrate the components of GA to assure that it functions as a coherent whole.

We expect that many parts of component implementation, adaptation and integration will be recurrent steps, as we learn from experience with earlier iterations. Therefore, we plan to build fully-integrated, complete systems as early as practicable, to give us adequate time to perform several iterations.

Experimental Domains

We will make a final selection of the specific two medical conditions for which we will test the GA approach. Along with these choices, we will also need to identify the particular institutions and individual care providers who will cooperate with us in the planned experiments and the particular group of patients who we intend to use as experimental subjects. These choices will also constrain us to use and support the database and data collection tools at the target institutions.

We will identify protocol-based studies of care for selected medical foci [Fiel92]. With these in hand, we will encode these protocols in formal procedures for use by the GA, and use these procedures to guide the development of the knowledge base for selected medical foci.

We will delimit, for each of the two medical domains, the exact bounds of what we plan to implement and test. This will involve determining all of the information that will be acquired and stored about each patient and the set of patient characteristics that the GA will try to influence.

We will select, implement, and integrate tools that will generate prototype modules. Where possible, we will attempt to make use of existing systems for automatic specialization of procedures to available knowledge, such as the system developed for automated construction of specialized scheduling algorithms developed by Doug Smith at Kestrel Institute for ARPA [Smit92].

Using these tools, we will implement prototype GA patient management systems for two selected medical foci. These prototype systems will provide the basic persistent personal medical history, therapy management, explanation, scheduling, and alerting functions.

We will identify and get rights to use domain-specific material for patient education and design and develop connections between the general material and patient-specific information.

We will integrate the implementation of these prototypes with one or more institutional information systems to demonstrate the feasibility of patient-centered medical records.

Experimental Testing

We intend to determine experimentally whether the ideas we propose can be successfully reduced to practice and whether, when used, they lead to the improvements in health care that we anticipate. Because of the severe ethical and legal responsibility that we undertake in the conduct of experiments on human subjects, we will work carefully with our collaborators and with the Institutional Review Boards of our own and cooperating institutions to define experimental protocols that will be based on properly informed consent of all participant subjects and that will assure safety during the experiments. Experimentation on human subjects with a newly-developed set of ideas and techniques is always a slow process if done carefully; therefore, we expect to achieve only limited experimental tests of Guardian Angel during the three-year exploration and development project proposed here. If the studies we propose here are successful, then we will plan more elaborate trials subsequently.

We will first try out GA in our selected medical domains by simulating patient use, with experimenters acting as surrogate patients. We will use realistic scenarios collected from actual patient experience.

We will then try out GA prototypes with actual volunteer patients, but only under supervised settings. This will be during visits to the doctor's office and in the presence of experimenters, and will be intended to test the usability of the system, not its effectiveness.

If these early studies are encouraging and we have assured the safety of subjects, we will conduct some limited trials (see below).

We plan to evaluate the experiments, and to use their results to:

(1) Continue evolution of the knowledge base for the selected applications.

(2) Revise the architecture and implementation based on trial results

If the limited trials continue to be positive, then we will plan a more substantial evaluation based on the anticipated capabilities of the system and its possible extensions.

We will also plan and conduct technology transfer activities to assure dissemination of the architecture and implementation.

Technical Rationale

Current health information systems are built for the convenience of health care providers and focus on orderly record keeping for an individual health care setting. This orientation leads to two classes of severe problems, which we plan to address through the proposed research: alienation of the patient from participation in decision-making concerning his or her own health care, and inability of health records to integrate information over the lifetime of the patient. Furthermore, many interesting opportunities are missed that would apply new technologies to improve decision making and to reduce the incidence of errors in carrying out the decisions made.

Roots of the Problems

a. Health information *for* patients.

Although medical care is obviously given to the patient, exists for the benefit of the patient, and must (legally and ethically) ultimately be under the control of the patient, most patients do not have a thorough understanding of their health status, do not appreciate their alternative diagnostic or therapeutic options, and do not fully understand most of the decisions taken in their behalf by doctors, nurses and other health care practitioners. Even medical doctors, when they have themselves become patients, have complained of the anonymous treatment they receive. Patients without medical training express their disillusionment by exhibiting "poor compliance," and by avoiding clinics until they have obvious major symptoms such as pain or visible alterations of anatomy or physiology. The common practice among geriatric patients of sharing each other's prescriptions, the popularity of "alternative medicine" and its unproven interventions, and anger at health care providers and researchers may all be seen (at least in part) as a lack of understanding and a lack of trust by patients in their providers.

Traditionally, responsibility fell on the doctor's shoulders to explain to the patient the meaning of symptoms and tests, the possible further diagnostic and therapeutic interventions being considered, the known and the unknowable in reaching decisions, and the likely outcomes of current plans. As medical practice has become more capable and more complex, it has become more and more difficult to explain all this to the average patient. In addition, the rapidly increasing pace of medical care means that there is no leisure time for the doctor to provide lengthy explanations and to make sure that the patient actually understands. Instead, patients are provided brochures and videotapes (for the more common ailments) that describe in some generic way the disease they have been diagnosed as having or the treatment that has been selected for them. In the absence of such specially-prepared patient education material, the patient can at best turn only to package inserts in their medication or to general texts on health care, ranging from the Merck Manual to the latest New Age writings. But these static information sources (even if accurate) do not help to understand just why this diagnosis explains the particular constellation of symptoms worrying this particular patient, or what effect the therapy will have on the specific plans or concerns of this individual. No wonder that patients turn to friends and neighbors for interpretation and advice.

We believe that health information systems should be designed and built specifically to meet the needs of their subjects, the patients. Therefore, support for helping the patient understand and participate in all aspects of medical decision making and its consequences becomes not merely a desirable peripheral issue in system design but a clear, central focus.

b. Life-long information.

A medical colleague once observed that every hospital in the country is absolutely convinced that its own specific procedures for collecting and keeping medical records are unique, and that no "ready-to-wear" system could possibly suit it. The history of dissemination of computerized medical information systems amply reinforces this view. There is still only a handful of comprehensive hospital information systems—ones that are concerned more with clinical records than administrative and financial ones—installed at civilian hospitals. Even though some of these systems have been in service for many years and are well-proven within their home institution [McDo90, Pryo83], attempts to propagate them to other hospitals have been generally unsuccessful. For example, two commercial vendors have been promoting adoption of the HELP system [Warn72] from LDS Hospital in Utah for nearly a decade, but (as of 1993) fewer than a dozen other hospitals have adopted it. Even large administratively-oriented

systems such as those sold by IBM for two decades typically underwent such major customizations during installation that the resulting systems shared little in common. At least hospital-wide information systems have provided common facilities across their single institution. At most institutions, even that level of integration has not been achieved. There, information known to one departmental system is simply unavailable to others, and it is only through the traditional paper record that one can integrate a coherent view of a patient's condition.

When a patient moves or otherwise changes doctors or hospitals, chances are that many of his or her medical records become effectively lost. The new primary care provider can re-learn the highlights from taking a careful history, but much of the detail of previous care will be lost and much that is retained will be at least subtly incomplete or erroneous. In the exceptional case where a past test result can lead to a clear solution of a new problem, old records can be found and copies forwarded, but this is an exceptional, not a routine process. Analysis of current data must therefore surely suffer. It has been known for over a decade, for example, that variations of laboratory values within a single patient are more narrowly distributed than are population distributions for the same test [Stat76]. Therefore, even small deviations from past values can signal something of possible clinical significance, even though the absolute test values are not yet abnormal enough to be notable on their own.[1]

The deficiencies of current record-keeping systems and proposals to develop new national standards for the maintenance and interchange of patient data among providers, insurers, etc., are well described in the Institute of Medicine's recent study, *The Computer-Based Patient Record* [IOM92]. Even this modern viewpoint, however, assumes that the primary locus of medical data about a patient is with the hospital, clinic or doctor's office. It then becomes a massive problem to integrate all these data when they are needed. The standardization efforts now under way should at least make that data integration possible, but as long as the locus of health-care information remains with multiple providers the problems of gaining access to it all in a timely manner and properly integrating it remain large.

We propose instead to re-focus health-care information systems on the individual. Logically, life-long medical information will be collected and maintained within an individual's personal system, thus assuring the completeness and continual availability of his or her records whenever they are needed. The patient-specific components of health records in possession of health-care providers are then viewed as secondary, for the convenience of the provider. The primary information can always be accessed directly from the individual's system. In addition, active processes within the system can augment the record-keeping function to provide a pro-active component that represents the patient's interests.

Domain Modeling

Most of the capabilities we seek to provide through GA are novel. Although others have certainly thought of the possibility of personal medical advisor programs, means of communicating between an at-home device and a hospital support system, and many other parts of the GA vision, very few such components have actually been built, and they have never been integrated into the sort of system we wish to build. Experience suggests, therefore, that flexibility will have to be a key virtue of our work, because our understanding of precise requirements will evolve greatly, technology to support our goals will improve dramatically, and we will certainly make our share of poor choices in design and implementation.

We share the methodology of the Domain-Specific Systems Architecture (DSSA) community [Huff94] in recognizing the importance of modeling explicitly the domain's potentially significant entities and their possible interactions. The resulting ontology will provide a terminology within which the GA reference architecture(s) can be defined, and will give a set of shared concepts for researchers to use in studying, describing, and negotiating specific requirements.

Building a suitable domain model is, of course, one of the early steps of the research. Nevertheless, we can sketch at least a rough outline of some of the concepts that are certain to play an important role in the ontology:

---

[1]Our recent work on identifying growth abnormalities in children demonstrates the same point. [Haim93]

People
– Patient
– Primary Provider
– Other Providers
– Consultants
– Parents or Guardians
– Family (spouse, children, …)
– Commander/supervisor and subordinates

Institutions
– Clinic
– Hospital
– Patient Support Group
– Information Service Provider
– Insurer
– Government
– Employer/military unit

Settings
– Home
– Outpatient visit
– Hospital stay
– Travel away from home
– Battlefield

Etiologic Entities
– Disease-causing organisms
– Environmental factors
– Accidents
– Deliberate harm
– Idiopathic causes

Processes
– Life
– Natural course of disease
– Diagnostic and laboratory tests
– Normal therapeutic activity
– Routine functioning of an institution
– Monitoring
– Computation

Procedures (descriptions that, when performed, evolve processes)
– Institutional Standard Operating Procedures
– Computer programs

Relations
– Causal
– Taxonomic
– Geographic
– Legal
– Familial
– Organizational
– Financial

Data Sources
– People (observations, reports)
– Instruments
– Laboratories
– Databases
– Libraries
– World-Wide Web

Architecture

The long-term development of GA will require a highly flexible, open architecture that will support expansion of the concept and evolution of its implementation. Yet after GA is in actual use, such changes must not ever require a "reset." We cannot, of course, hope to solve the full problem of maintaining installed bases, but we plan to exploit, as explained below, the insights of the DSSA methodology to develop an open architecture capable of encompassing a wide range of long-term technological developments.

In addition, it is absolutely critical to our ability to build the system that virtually all technology components of GA be provided by generic, preferably standards-based facilities. These include:

data repositories and retrieval, perhaps based on the object-oriented data model

reliable checkpointing and backup

computer-based patient record

security and authentication

network-based communication

alerting and notification, including pagers, electronic FAX, and email

interfaces to data-capture instruments

easy-to-use user interfaces

formal expressions of guidelines, practice standards, etc.

hypertext-based bodies of medical information

There is no surer guarantee of failure for this project than to take on the task of building, from scratch, technology components that can be adequately provided by commercially available means. Thus, for example, one could easily get side-tracked into an investigation of the best methods for assuring the integrity of life-long records within a small portable computer that would implement GA. One can imagine catastrophes that would result in destruction of the device itself, loss or theft, temporary unavailability, accidental or deliberate alteration, etc. Although it might be an interesting exercise to work out specific architectural ideas that will overcome these problems, it is even more sensible to architect a solution that will rely on anticipated commercial services that will provide comparable facilities for a broader computing market. Then, specific technology choices can be framed in terms of engineering trade-offs.

To continue our example of long-term integrity, one could choose an infrequent (say monthly) backup tape for people with no serious current conditions. At the other extreme, a direct cellular call to a system providing guarantees of data integrity may be worthwhile when significant events occur in a patient being monitored for possibly life-threatening events. Fruitful synergies may also arise in such an architecture. For example, a hospital may choose to go into the commercial business of providing data integrity for its patient's GA systems; in that case, communication costs can be reduced and reliability increased if the same telephone call can be used both to report time-critical data (part of the data repository and the alerting and notification functions) and to assure that it is backed up in case the GA device might malfunction (part of the backup function).

Some of these facilities are already being provided commercially. Others are the focus of active current and proposed research on the National Information Infrastructure. A few will require at least prototype development by this project, if not otherwise provided. There is enough work to do on integrating these facilities and using them to try out the GA concept that we must be very careful not to take on more than what is absolutely essential here.

The figure below shows a schematic sketch of a portion of the overall GA domain and the corresponding pathways for communication.

At the center is the individual, holding what we have called the GA_PDA in the scenario, above. The patient is potentially in direct communication with a set of instruments that help gather information about him or her. For the diabetes example, these could include a glucometer and a scale with a computer interface forming part of a medical BodyNet [Schi94]; but other variations might include communication with exercise machines, or with temperature, respiration, and cardiac sensors for infants at risk of crib death (SIDS). We assume that the home, in the form of a home computer or access to some networked provider, provides more permanent record keeping and access to other services and institutions. Other connections link the patient with his or her doctor's office, to provide a pathway for information that requires urgent attention, to civic emergency services, and to wide-area network service such as the Internet for information and support services.
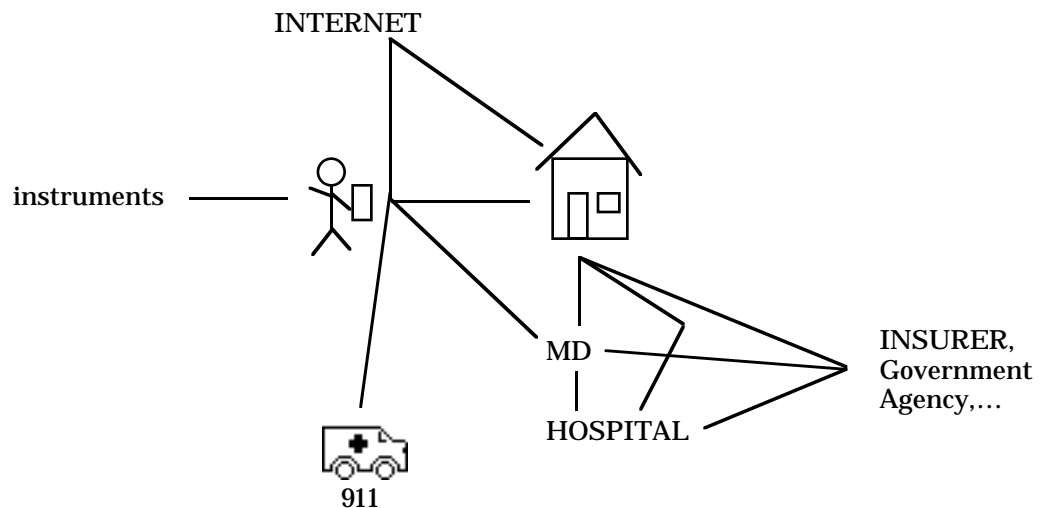
This diagram is highly schematic as it omits the probable and possible connections between all the other people, institutions, and data sources entering into the domain. In principle, any of these entities and their associated GA processes (if any) may communicate with any other, but we expect the communication to be direct in only a few important cases that may change as the technology evolves.

Special Functionalities

The DSSA concept [Huff94] is most easily applicable to domains with existing limited systems, and provides a methodology for modularizing and generalizing these systems to facilitate future reuse of components in constructing new or variant applications in the same domain. In the GA domain, however, there is no existing system from which to generalize, so we must begin by designing a specific system. Nevertheless, the concerns of the DSSA methodology provide an important guide in this task, as the GA concept involves an array of specializations and variations of immediate interest.

It seems unlikely that any single GA device or overall functionality will be appropriate for every person, at least for quite a long time. The different physical and medical needs of different people call for very different specific functionalities. The scenario of juvenile diabetes presented above calls for functionalities and a sensor (glucometer) not necessary in the GA_PDA of the average person. While technological progress may eventually permit realization of these capabilities in every GA at negligible cost, in the short run we can expect the GA provided to a diabetic to provide functionalities extending or refining those present in the generic GA. Similar variations of special skills can be expected for GAs appropriate for other special subjects, such as hypertensives, stroke victims, and others. More generally, we foresee variations of GA functionality across ages as well. For example, attaching the GA_PDA in the scenario above to an infant would not be desirable or feasible, but even in the short run we expect a very limited GA to be desirable for infants. This might consist of a purely passive personal record of allergies, medical conditions, etc., for use by emergency personnel; indeed, one can imagine replacing the standard delivery-room identifying bracelet with a more permanent bracelet containing the initial personal medical information. As technology permits, this minimal functionality might be expanded to include temperature, respiration, and cardiac sensors to alert parents to infant distress and the risk of impending crib death. At the other end of life, the generic GA_PDA would also not be appropriate to people suffering from Alzheimer's disease or for comatose patients, and specialized designs might provide the appropriate functions for them. For example, the GA of an Alzheimer's patient might monitor the intake of medications and provide address and contact information to bystanders if the patient becomes lost.

In view of this wide range of possible functionalities, we plan to focus our design on the GA processes appropriate to the generic adult and to the special medical applications selected (diabetes, hypertension, anticoagulation, etc.). Designing for this small variation and keeping the wider possible variation in mind will help identify the basic functional modules and their possible refinements and alternatives. In some cases, the refined versions will be generated automatically from the generic versions and the specifics of the requirements for the variation; in other cases, the alternative versions will have to be constructed manually as with the generic versions.

Generic Functional Modules

Examining the list of GA functionalities presented above, we observe the following generic functions of the GA.

- Recording and representing information
- Monitoring processes
- Analyzing events and results
- Interpreting and explaining plans and references
- Adapting and customizing plans and preferences
- Checking plans
- Communicating, translating, providing, and retrieving information

Coordinating functionality distributed among GA processes at different sites

Educating the subject

Reminding and alerting the subject and others

Supporting and socializing the subject

Scheduling

Each of these functions will be performed by one or more architectural modules. Some functions correspond to a single isolable activity, and will be performed by a single module in a given GA, with the only variation being variation across GAs due to the differing needs of the subjects. For example, the basic module performing sanity checks on plans might well serve all GAs without exception; but in the short run, one can expect different specializations according to the special medical knowledge required for different subjects. That is, the generic sanity checker analyzing the plans for average subjects need not know the details relevant to checking plans for diabetic subjects. Such variations because of differing specialized knowledge should be especially amenable to automatic construction and refinement. Other functions, however, will be performed by a set of related modules specialized to the types of information or the entities involved in the function. For example, modules for recording data from bodily sensors will be different as a matter of engineering from modules for assessing and recording the preferences of the subject, for representing these preferences to other people and institutions, and possibly for representing these preferences to the scheduling module.

Hardware Realizations

Seeking maximum flexibility in adapting the GA architecture as hardware technology develops entails placing no restrictions on just how the memory, computation, or communication functions of the GA domain are realized in hardware systems. The diabetes scenario presented above envisions a personal GA residing on something like a current PDA, the home GA residing on a personal computer, and makes no particular assumptions on the residence of the hospital GA process. But fast downloading abilities, the shrinking physical size of computational power, and improvement in the resilience of memory might obviate the need for a home GA altogether except as a communication terminal, and improvements in cellular telephony and wireless connections to the Internet might obviate the need even for that role. Our initial architectural designs will accordingly identify hardware realizations that make sense currently and in the near future, but the domain model, functional architecture, and modularization will not be dependent on the particular hardware choices in force at a particular time.

Information Infrastructures and I-95

We anticipate that the current push toward universal network access for all Americans will provide vast new opportunities in health care as well as in other areas of commerce and industry. We expect that early exploiters of this universal access will build vertical applications that serve specific, targeted users with narrow services. Soon, however, many users and providers will recognize the great advantages of open systems that promote interoperability and the incremental addition of services to existing systems. For users, this will lead to convenience of use and consistency of operation. For providers, it means that one can concentrate on doing the best job at providing a specific new service without having to build a complete application.

The I-95 architecture [Tenn93], in whose creation we participated, envisions a layered model of services. At bottom are traditional network services and concerns such as reliable communications, addressability, and scalability, plus new services such as security, authentication and bounded-delay data streams for isochronous data. At the top are application level services and connections to a variety of user interfaces. In-between, we foresee a layered set of capabilities that provide uniform support for a large variety of transactions. These include electronic money (or, in general, ways of being billed for the use of network services), methods to support negotiation and binding commitment in transactions, and means to help navigate essentially unbounded volumes of data and text.

We plan to capitalize on future developments of this set of architectural ideas to support GA. Ideally, GA will contribute its own set of layered services and will in turn rest on services provided in greater generality for many other kinds of transactions. For example, we expect

not to have to develop our own standards for encryption and authentication, but only to design the application of those methods [Rive78] to the specific needs of patients and providers in health care.

Prototype Applications

Implementing the wide range of Guardian Angel applications that we anticipate eventually building is a task too large for a single research project. Indeed, we intend for the results of this project to provide the appropriate architecture and examples that will enable others to build the many specific applications. To help us develop and refine the GA concept, we need to choose one or two specific medical applications that can represent the larger medical problem and for which we can build demonstrable and testable prototypes. For these prototypes, we plan to choose medical domains in which the intensity of required care is higher than normal for the general population. This is justified by the medical importance of the applications themselves and by the fact that we are likely to learn more from experiments in which "much happens" than otherwise.

Our initial choice of prototype application domains is insulin-dependent juvenile diabetes and management of chronic hypertension.

Insulin-Dependent Juvenile Diabetes

There are approximately twelve million patients in this country with diabetes mellitus, of which one million patients had juvenile onset. The cost of the complications due to the microvascular and macrovascular disease of patients with diabetes mellitus is well over four billion dollars [Huse89]. Therefore, any improvements in efficacy of treatment will have significant national impact. After the Diabetes Control and Complications Trial (DCCT) [DCCT93], a recent multi-center trial, it has become clearer that minimizing the average blood sugar of patients will lead to major decreases in the incidence (e.g. 50% decrease in retinal disease) and severity of diabetic complications. Also, the Department of Health & Human Services has identified in Healthy People 2000 [PHS91] several diabetes-related national objectives including reducing the incidence of blindness and end-stage renal disease. However, as noted in an editorial commentary on the implications of the DCCT in the New England Journal of Medicine [Lask93]: "We now face the considerable task of providing patients and practitioners with the support they need to make this therapy generally available. The importance of the patient-centered team in achieving the results of the DCCT suggests that policies affecting the way care is organized, assessed, and paid for are likely to be as important as those affecting benefits, insurance coverage and access to care. "

Even current, non-intensive management of pediatric patients with diabetes mellitus is intensive and expensive in its use of health care resources. In addition to regular clinic visits, each diabetologist (nurse or physician) will field dozens of phone calls per week from patients or their families to help adjust insulin dosage (or oral hypoglycemic agents), diet or exercise. Despite this investment of clinicians' time, the interval between calls for a particular patient will range from weeks to months, during which time the current therapy may have become progressively less well tailored to the patient's needs. Sub-optimal therapies will tend to increase the likelihood of acute complications (e.g. ketoacidosis and dehydration or hypoglycemia) and the risk of developing progressive chronic complications (because of suboptimal glycemic control). To meet the requirement for more stringent glycemic control (which has been shown lead to substantial decreases in vascular complications in the DCCT) will require greater patient education, increased clinician/patient communication and heightened responsiveness to changes in physiology (e.g. during a gastroenteritis) or lifestyle.

Several attempts have been made in the past ten years to provide patients with diabetes with automated assistance in the management of their blood glucose control [Stra85, Bell88, Buys89, Chya90, Fuge92]. These systems have not had a significant acceptance beyond their initial testing sights or lasting impact upon the care of most individuals with diabetes mellitus for a variety of reasons, including: inadequate user interfaces, hospital oriented care, burdensome, unreliable or low bandwidth communications protocols. Also, several of these systems suffered from a too narrow focus on the specific clinical task [Fuge92] rather than the broader mandate of the Guardian Angel project. Nonetheless, we intend to include these components of prior systems that have proven to be successful such as the insulin adjustment heuristics.

The objectives of the implementation of a Guardian Angel system for Diabetes Mellitus include:

(1) Decreased patient reliance on expensive tertiary care specialists for routine advice on home management of diabetes mellitus. At the same time, timely communication of relevant clinical data to the diabetologist (nurse or physician) will be increased.

(2) Improved early intervention to correct harmful blood glucose levels, thereby reducing the likelihood of hospital admissions or specialist consultations for acute complications and potential chronic complications.

(3) Increased patient autonomy.

(4) Improved clinical data collection to help conduct outcomes research and suggest improved therapeutic interventions.

The extended scenario, above, also introduces additional considerations that are important in this disease domain.

Chronic Hypertension

Hypertension affects 20% of the population of the United States. The prevalence of hypertension increases with age, with a higher prevalence in men until about age 50 and then a higher prevalence in women [NIH93]. The natural history of the disease is that it gradually becomes more severe. Over many years it gradually damages the blood vessels, the heart, the kidneys and other organs. The heart hypertrophies and eventually fails. The increased pressure accelerates atherosclerosis and produces also hemorrhages in end organs and aneurysms in large and small blood vessels, affects the retina, the kidneys, the brain, and the supply of blood to the legs. The aorta is subject to dissection (tearing), producing rupture and the cutoff of blood vessels. There is a strong, predictive correlation between the level of blood pressure and the incidence of coronary heart disease, stroke, heart failure and renal disease.

The presence of hypertension is usually determined by a regular physical exam or other screening, since symptoms are unusual in the early stages of the disease. The majority of hypertension is essential , meaning that there is no known cause that can be treated. However, it is important to rule out possible causes, since secondary hypertension is treated differently from primary hypertension and because treatment can be definitive (curative) for some of these causes. Since essential hypertension is progressive and incurable, the disease usually needs to be treated for the rest of the patient's life. For the borderline hypertensive, the therapy consists of diet, exercise, and monitoring. For mild hypertension, drugs such as diuretics, beta-blockers, calcium blockers, and angiotensin converting enzyme inhibitors are all appropriate, in different subsets of patients. For more severe hypertension there is a wide range of choices of single or multiple drug combinations. The large armamentarium is needed because patient response varies, both in degree of response to therapies and in prevalence of and tolerance to side effects. It is common to try several different therapies in a patient before finding one that provides an adequate degree of control for the hypertension without more side effects than the patient is willing to tolerate.

One of the real challenges is that most hypertension therapies have some side effects and the patient often feels better untreated than treated. Many of the common side effects of anti-hypertensives are fairly nonspecific, such as headaches, fatigue, dizziness, depression, and sexual dysfunction, that might not be attributed to the therapy by the patient without appropriate questioning. Since there are usually some drawbacks to every therapy, the decision to reject a therapy can depend on the attitude of the patient. As the disease progresses, the therapy may have to change to maintain an adequate degree of control. The therapy may also need to be changed because of end-organ damage, changing sensitivities and allergies.

Another challenge of hypertension management is tracking the blood pressure. Blood pressure can vary considerably and is often higher in the doctor's office than in a different setting. It also follows a pattern over the course of a day and increases with stress and other factors. Thus, regular monitoring of the blood pressure provides a much better picture of the state of the patient's disease than the pressures read in the physician's office [Pick87]. If the physician-read blood pressures are the sole source of data, it takes multiple visits to establish the diagnosis of hypertension and frequent visits to track the progress of therapy. Appropriate trend

analysis of home blood pressure measurements can provide the physician a better measure of the hypertension and the effect of the therapy as well as giving the patient more timely feedback and encouragement than waiting for the visits to the physician.

The blood pressure is not the only focus of management in the hypertensive. Life style modification is a very important part of the overall therapy, including changing diet (salt, calories, and cholesterol reduction), exercise, often weight loss, stress management, and smoking cessation. Since the therapies and other management strategies often have more impact on the patient's day-to-day life than the disease, keeping the patient convinced of the importance of control, encouraged, and informed is a major challenge in assisting the patient. Part of the challenge is getting the patient to take ownership of the management process, something that takes instruction and the involvement of the patient in determining a patient-specific strategy for managing the disease. This can be very time consuming and the assistance of the GA system could benefit the patient and take some of the educational burden from the physician. As the GA monitors for early complications (often by asking the patient about signs and symptoms), it will reinforce the importance of blood pressure control. It also will show the patient the success of therapy and life style changes by showing their impact on the patient's risk of complications. For example, several of the Framingham risk equations for predicting the development of coronary artery disease depend on blood pressure and many of the associated life style changes. The GA will graphically demonstrate how his/her risks change over time. This feedback of changes in prognosis for a therapy to prevent complications in minimally symptomatic patients should be very effective as it will show them the expected effect of therapy.

Managing Hypertension Using a GA

At different stages in the management process, the GA could serve different functions for the patient and for the providers. When hypertension is first suspected or noted on a routine examination, the physician will load and activate the hypertension GA in the patient's computer. Initially, the GA could assist in the screening and diagnosis process for the physician. The GA would collect any additional needed demographic information and information about the risk factors for myocardial infarction and stroke. This function may be of particular importance because modern trends in health care continually limit the time available for patient-physician contact. How complete a history can be taken in an eleven minute visit? The patient could also get information about hypertension and the various risk factors. If a health care provider (nurse or physician's assistant) is entering blood pressure measurements into the office, clinic or hospital system, that information would be transferred to the patient's GA computer. If a hypertension workup is appropriate, the GA would gather the rest of the history needed by the physician. At this stage of the interaction, the GA could give the patient a preview of the implications of risk factors that are apparent from the history and initial blood pressure readings along with other information that would help the patient be more informed when interacting with the physician.

When the patient's GA transfers the data to the GA physician's process, the data needed to begin the diagnosis will be available to the physician. If the patient already has some of the necessary screening tests, those data will be immediately available and some unnecessary duplication will be avoided. The GA will also be able to provide data about trends in tests over time, for example how the serum creatinine may have changed (reflecting renal disease, both a cause and a complication of hypertension). The diagnostic module on the GA physician's process can then provide patient-specific source of information about the diagnosis and management of hypertension. It could provide suggestions for therapy or a critique of proposed therapy in the context of other diseases (co-morbidities) or special circumstances that may be affecting the patient [Mill84]. For example, in a patient with a history of depression, a beta-blocker would be inappropriate. The information could be provided in the form of a guideline or in response to questions, as appropriate to meet the needs of the physician. The GA physician's process would also provide access to other information sources, such as references to literature on secondary causes of hypertension, the current costs for drugs, availability of classes or therapy groups for the patient on diet, quitting smoking, exercise, or living with hypertension. It would also provide information about the patient's response to drugs given before (about which the current provider may not be aware), as well as allergies or how other drugs might interact. Data would also be provided to the patient. For example, most common over the counter remedies for upper respiratory infections warn the patient to avoid them without

consulting their physicians. In the case, the GA could provide a selection of minimally problematic drugs and then monitor the patient's blood pressure response to those drugs and note the results for future reference.

Once the patient has been diagnosed with hypertension and the physician has determined an appropriate course of action, the patient's GA will assist by helping the patient take ownership of the plan. To successfully combat hypertension requires a long-term commitment and consistency, so the patient must understand the situation, understand the implications in managing the disease, and be committed to the management plan. The GA will provides an effective tool to take the patient through these steps. It will answer questions about the patient's risk of heart attack or stroke or the other consequences of hypertension and frame this information in a variety of ways that might help the patient understand. It will help the patient understand the additional risk incurred from smoking, diet, obesity, stress, and so forth. The GA will also help the patient personalize the management plan so that the goals are ones the patient is committed to achieve. This management plan will incorporate the GA as a monitor and guide by including a commitment to consult with the GA on a regular schedule.

Once the patient and physician have established a management plan, the GA will help the patient follow the plan and monitor the progress. For most patients, who are able to check their blood pressure with the home pressure cuffs now available, the GA will track the pressure against the expectations for therapy effect and use the data to provide correction and reinforcement for the patient. The GA will also monitor other aspects of the plan including drug compliance, weight, and lifestyle changes. The degree of monitoring of side-effects and lifestyle changes such as diet and exercise will depend on the needs and desires of the patient. It could vary from simply making information available to having the patient fill out a progress report each session. Importantly, the GA will be able to ask about most of the potential side effects of the drugs. From our interviews with hypertension patients it is clear that some like to know what the possible side effects are while others would prefer having them monitored in case they happen. Indeed, there are preliminary data that suggesting to a patient that certain side effects might occur, e.g., insomnia or erectile dysfunction on beta-blockers, will markedly increase the frequency of those side effects. The GA will accommodate individual preferences for information compared to monitoring. The GA will help the patient evolve the plan to meet changing needs, such as revising an exercise program to accommodate a change in schedule. By occasionally linking the GA to the physician's GA module via modem, the GA will schedule follow-up with the physician. Since the GA has the monitoring data, it can determine the urgency of the visit. Indeed, if expectations have been met and there are no indications of side effects, it may be appropriate just to add the data to the physician's data base and delay the visit, unless of course it is time for some regularly scheduled laboratory test (e.g., a potassium or an echocardiogram to monitor changes in left ventricular mass) or unless some other problem requiring the physicians attention has developed.

The GA will act as a resource to answer the patient's questions as they arise. The GA would assist in this process by keeping track of what information the patient has seen and by suggesting related information, information consistent with the patients interests and knowledge. Many kinds of information about the disease, the therapies, lifestyle changes, and so forth will be available with a simple hypertext interface. The GA will be able to access functionality on other computers to supplement that available locally. For example, if the patient has agreed to a certain kind of diet, transparent access through modem to another server may provide a list of restaurants and dishes that would meet those requirements when the patient wants to go out for a meal. Such network access will also allow the patient to participate in support groups or mailing lists to exchange views, keep informed, and be encouraged by people in similar situations. Often the best way to stay on track in managing a lifestyle change is to interact with someone else fighting a similar battle and be accountable to each other. Access to network mailing lists or information servers will also make available timely responses to articles that appear in the press or television about some aspect of hypertension.

Chronic Anticoagulation

Anticoagulants are drugs which limit the ability of blood to clot and are used commonly in the treatment of cardiovascular disorders and therapies that subject the patient to a risk of thromboembolism. Classic indications have been prior embolic stroke, the insertion of prosthetic heart valves, the presence of acute and recurrent venous thrombosis (phlebitis) and

pulmonary emboli, the presence of cardiomyopathy (disease of the heart muscle) which limits the strength of cardiac contraction, and the presence of certain cardiac arrhythmias, e.g., atrial fibrillation. Recent studies have underscored the need for anticoagulation and careful monitoring of the patient's response to these drugs (as typically measured by the prothrombin time, a measure of how long it takes the blood to clot under controlled circumstances in the laboratory). Patients responses to these drugs (most commonly warfarin) are quite variable from patient to patient and from moment to moment in a given patient. Recent improvements and corrections in laboratory science have minimized some of the non-biological variation, but the situation is sufficient unstable that laboratory studies of a patient's response a measured at least monthly and often far more often. If anticoagulation is inadequate, the patient is at risk for thrombo-emboli; if it is excessive, the patient is at risk for hemorrhage. The toxic-therapeutic ratio is quite low and variable.

Almost irrespective of the purpose of the anticoagulant therapy, the management is the same. A therapeutic goal (a target range of the laboratory value, the INR (international normalized ratio) is established as well as action levels outside that range, with actions such as repeating the study, changing the dose, given an antidote or another anticoagulant (e.g., heparin as a supplement), and hospitalizing the patient. Because of the frequency of problems and changes in therapy, it is critical that the patient be a partner in the therapeutic process, understand the implications of laboratory values, and understand the signs and symptoms of the complications of excessive or inadequate therapy. A very large set of other drugs also interact with anticoagulants, causing the risk of bleeding or of thromboembolism to increase, sometimes without changing the prothrombin time. The patient must be educated about these effects, because an intercurrent physician treating an acute problem may inadvertently prescribe a drug that interferes with safe and adequate anticoagulation. The patient must also be taught about a wide variety of over the counter drugs that can cause bleeding or inadequate anticoagulation.

Given the frequency of problems with anticoagulant therapy and the importance of this lifesaving therapy that is often continued lifelong, we believe that a GA for anticoagulation will be an important benefit. It should make this therapy safer and more effective, while keeping the patient involved and responsible. There are even portable instruments (one made by Dupont, the primary manufacturer of warfarin (Coumadin)) with the patient can purchase to measure the prothrombin time at home at about one dollar per test. The output is digital and should be able to be directly interfaced to the patient's GA. The frequency of testing and changes in therapy and the need to monitor for side effects and drug interactions make this domain a prime candidate for GA software agents. The frequency of changes in therapy (especially when therapy is first started or when it must be temporarily discontinued, e.g., for surgery or even visiting a dentist) means that we should be able to do substantial testing of this application in a relatively short time. This domain is also interesting because the GA will not be focused on a single disease but rather on the effects an the management of a common therapeutic agent. It may be prototypical of a large set of software agents that might be prescribed for a patient (loaded into his/her computer) when a drug is prescribed, to teach the patient about the drug and its use and necessary precautions.

We shall develop and anticoagulant GA which will monitor the effects of the drug, trying to keep the patient's response (INR) within the prescribed target range. It will educate the patient, help the patient understand how drug dose should change with laboratory response, watch for complication and try to prevent or at least compensate for drug interactions. Because the protime might be more frequently and closely monitored at home than can now be done by visiting a laboratory and getting a blood sample obtained by venipuncture, control should be tighter and safer. When scheduled intercurrent events are planned (e.g., a visit to the dentist or a sigmoidoscopy), the GA will help the patient and the physician minimize the window during which the patient is relatively unprotected from thrombo-embolic events. The GA will keep the physician informed about the patient's response, progress, and complications, and its recommendations to the patient will (can) be monitored or adjusted by the prescribing physicians. Some busy physicians even now delegate anticoagulation management to assistants or nurses, so this pattern of care is already partially commonplace.

Other Domains

The domains described above are only three of a number of medical problems that could be fruitfully attacked by prototype GA implementations. We have mentioned angina, chronic renal disease, and pulmonary disease (COPD) as other possibilities. If time permits, we may undertake additional prototype implementations in these domains, to further study the adaptability of the GA concept, the GA reference architecture we will build, and the reusability of components built for the initial the domains. In selecting the three domains we have tried to select problems and expose different opportunities for GA functions, e.g., managing drugs, monitoring physiology, educating patients, watching for early complications, looking for interactions with other diseases or drugs.

Angina, for example, is a domain in which a GA has the potential to reduce the incidence of myocardial infarction and death. Since the majority of patients experiencing a myocardial infarction describe a change in their anginal pattern within a couple months prior, the detection of such changes in a more reliable way than relying on patient observation and concern, offers the opportunity to intervene in many cases that would otherwise proceed to infarction.

Implementation

Designing our initial implementation strategy is one of the tasks in our research. Nevertheless, we can outline some of the approaches we plan to pursue in implementing the capabilities we propose to build. We focus first on initial hardware platforms, and then on the software environment.

We assume that, in the short term, lightweight devices with relatively limited computational power will be used routinely by patients in our experimental studies. This will mean either a "personal digital assistant" class machine such as the Apple Newton or a very small, lightweight portable computer such as an HP "palmtop" or a larger machine such as a Duo. During the term of this project, such machines are not likely to have either the computational power or the storage capacity to hold all the material that GA will require. Therefore we posit the presence of a home computer (a PC-class machine or workstation) with a large disk drive, CD-ROM, keyboard, and large color video display. Communication paths between the PDA, home computer, and other aspects of the network will have to be always available, reliable, high-bandwidth, and convenient to use. For close proximity, infrared devices appear to be appropriate. For more distant connections from PDA to other machines, the current technology of choice is probably cellular telephone, though this may change as satellite-based direct communication channels come on line. We assume that, at least by the time we are ready to conduct field trials, the home computer will be attached to the Internet via high-speed telephone switches or set-top cable converter boxes, and therefore the communication facilities of the Internet will be available in the home. Therefore, we plan to use Internet services as the basis for communication with the doctor's office, hospitals, insurers, third-party payers, government organizations, other patients and support groups, etc. If, contrary to our expectations, high-speed networked connectivity to the home is not available, we would have to turn to existing alternatives such as high-speed modems or leased lines to conduct our studies.

Our collaborative team brings a very valuable suite of software tools to our project. Gensym's G2 and GDA (G2 Diagnostic Assistant) software provides an excellent environment for prototyping GA components and for developing monitoring and tracking modules. It has been used in process-control applications to construct reliable agents with long-term plans and agendas, and should allow us to create multiple independent software agents that collectively constitute the reasoning part of GA for a specific individual. G2 and Gensym also have considerable use and experience in interfacing to a variety of different information systems, and provide tools that make creation of data interfaces relatively simple. Because we will need to interface with a variety of record-keeping systems in doctors' offices, hospitals and other institutions, this will be a critical contribution. In addition, the temporal aspects of this software will provide us models for forms of knowledge representation we expect to be important, but which are not yet fully supported in KIF [Gene92] and KQML [Baal94].

IBM's Applications Solutions Institute has developed, in their collaboration with Kaiser Colorado, a set of tools for compiling workflow specifications into program code that implements those workflows. These capabilities will be very important to the components of the GA architecture that run at institutions away from the patient, where the needs and concerns of

many patients are intertwined and need to be simultaneously managed. IBM is currently using these tools to develop software for more conventional HMO operations, and we plan to investigate how they are applicable to the management of the far more patient-controlled data streams that will result from use of GA. IBM has also developed a set of tools to manage taxonomies of medical terminology, and is currently using these to define controlled vocabularies and medical relations for use in their HMO record keeping systems. We believe that a coherent set of terminology, though not absolutely essential, makes record-keeping and interpretation far easier and more effective. We plan to use the UMLS as our base vocabulary, where its coverage is sufficient, and to use the IBM tools and vocabularies as much as possible to supplement UMLS with terminology needed to cover descriptions of patient data.

The MIT Laboratory for Computer Science plans to pursue an aggressive research and development strategy to make use of the Word-Wide Web (W3) [Bern92] architecture as a foundation for our planned I-95 extensions. In concert with this, we plan to use W3 extensively as a foundation for accessing general medical information, patient support groups, the technical literature, and patient-specific records. If suitably augmented with privacy features, existing tools that provide HTML access to various forms of databases and widely-available front-end tools such as NCSA Mosaic can then provide high-quality interface components that require little effort on our part to develop, yet provide excellent functionality and uniform interface conventions for our users.

Not every implementation component of GA can be obtained "off the shelf," of course, and we certainly expect to have to build various software modules and code to tie together and integrate others. As we have mentioned, these implementation commitments are subject to re-evaluation before the project commences as technologies change. They are, however, representative of our commitment to leverage as much as possible from our own previous work and the outstanding work of others.

One of our initial concerns will be the form of the data model. In the GA environment the data needs to be useful for the life of the patient and by modules yet to be conceived. Thus, we can not assume that because no current modules use certain data, that it will not be needed later. Furthermore, the data comes from multiple contexts and sources, which may be pertinent to its appropriate use. Unlike a hospital based system, there is no underlying assumption that all data has been collected or filtered by appropriate clinical personnel. Thus, it will be necessary to store contextual information along with the data. For example, in the diabetes example, when a glucose level is recorded, the datum must include that it was done by the patient at home, the type of glucometer, and that it was recorded by direct connection (as opposed to typed in). Other information, such as the food and exercise state of the patient, are also pertinent but would be available as data items themselves. The actual implementation of such a data scheme does not have to be burdensome, since context information can be carried over sets of data. The starting point for designing the data model is the domain model. The relationships among the entities in the domain determine the classes of information that must be available in the data.

There are several different kinds of data that need to be a part of the GA system, with different implications for storage and access. The most obvious type is the disease findings, such as the glucometer findings or exercise logs. These constitute the basic track of the disease. In contrast, the GA also needs to keep track of the educational materials the patient has reviewed in order to offer new materials when the patient has questions and to have a general characterization of the patient's disease understanding. The requirements for data security are very different for these two kinds of information. The loss of the raw disease data would represent a loss of part of the basic patient disease record. The loss of the patient education trace is little more than an inconvenience. Thus, the data model needs to capture the varying needs for reliability, security, and access for different kinds of data.

Comparison with Ongoing Research

To our knowledge, this proposal represents a novel departure from both our own and our colleagues' ongoing work in the development of medical information systems. The history of medical informatics begins in the 1960's with the computerization of the collection, storage and retrieval of clinically-relevant data, initially focusing on data from the clinical laboratories. Over the past three decades, this focus has expanded to making possible the comprehensive collection and maintenance of clinical data from many other sources, including radiology,

pathology, pharmacy, ICU monitoring equipment, etc. [Ball91]. Hospital information systems being introduced today are moving away from the monolithic centralized systems of earlier days and now support networked interaction among heterogeneous components, with broad conventions and policies governing communication (e.g., HL7 [HL7] as a network communication format), queries (e.g., SQL for access to stored data), and interactions with other hospital responsibilities (e.g., a policy requiring that billing be performed based only on data recorded in the clinical information system). Nevertheless, all such systems are still centered on the care-providing institution, not the lifetime experience of the individual patient. Current efforts to define a set of standards for communicating information in the patient record (e.g., CPRI) are intended to facilitate the exchange of relevant patient information among various providers, but integration of this information is still planned to be optional and to remain the responsibility of the current provider. No provisions are made for long-term patient-centered evaluation of these data or for active processes that can represent the patient's viewpoint and preferences in medical decision making or that can prospectively identify errors as defined by the patient.

Current efforts in medical AI center on a number of different types of tasks: (1) Development of model-based reasoning techniques and systems that can accurately reason about complex medical domains. Our group, for example, has been developing a program to reason about diagnostic and therapy management problems in cases of heart failure, helping physicians to manage very ill heart patients. (2) Create methods to learn medical knowledge from the growing volume of data available in information systems. Examples here include statistical learning methods like CART [Brei84] and C4.5 [Quin93] Bayesian network induction methods [Verm90], case-based reasoning [Koto88, Jang93], and various approaches to knowledge-base debugging. (3) Develop tools that make the construction and deployment of relatively conventional (well-understood) systems easy and inexpensive. See, for example, the hierarchy of tools embodied in the Protege project of Musen, *et al.* [Muse89] (4) Provide integrative functions across heterogeneous data sources, with slightly intelligent interfaces in "physician workstation" projects [Koha90] . (5) Standardize methods of expression of clinical data (e.g., CPRI), medical terminology (e.g., UMLS) and isolated chunks of interpretive expertise (e.g., Arden Syntax MLM's [Hrip90]). (6) Codify and help to implement guideline-based care for routine problems, based on flowcharts and possibly some exception mechanisms [Fiel92]. Although each of these approaches is relevant to aspects of our proposal here, none is addressed at our specific task and plan. We hope to make a major contribution to medical informatics with the ideas sketched here, as developed through the conduct of the proposed research.

Performance Evaluation

Evaluating a new paradigm for health care delivery provides many methodological challenges. The evaluation of incremental changes in a system is relatively straightforward, because there is a clear null hypothesis—the new system is no different than the old—and either parallel or cross-over studies can provide evidence about this hypothesis. Problems normally revolve around finding suitable outcome measures (e.g., are improvements in measures of the process of health care sufficient, or must we show ultimate benefit in terms of reduced patient mortality and morbidity), and designing studies with sufficient power to reject the null hypothesis without inordinate expense.

In our case, evaluation is made more complicated by the fact that both our own ideas on how to build GA and the more traditional approaches to providing medical informatics will evolve during the course of the project. Because of this expected evolution, it makes less sense to perform a single final evaluation of the total system than to combine small, incremental evaluation steps with continued research and development. Many of the questions we will be asking are not of the form "does method x perform task t better than method y?", but more "how can we accomplish t at all?" The first steps in evaluation are, therefore, less a matter of empirical analysis of the performance of existing systems than part of the engineering design effort. We expect much of the first two years of our project to be devoted to research and development toward the GA goals, and therefore to involve mostly "how can we…" questions.

The third year of the project will focus on the performance evaluation of the GA. The objective of the evaluation is to prove the viability of the idea of patient-centered health information systems. Furthermore, we will use this evaluation to prepare the GA system for

appropriate clinical trials. In addition to evaluating the performance of GA in specific application areas, we are also interested in judging the success of its architectural design in making possible the creation of new GA applications in new settings. We do not expect to be able to carry out controlled trials of this aspect of the system during a three year study, but we are designing all stages of our evaluation to provide feedback on this question when possible.

There are four levels at which a system such as this needs to be evaluated, from system functionality to patient outcome. The following are the hypotheses that need to be tested at each of these levels:

1. The system successfully implements the performance characteristics we have set out. That is, it is able to collect and maintain the appropriate data, communicate with the other processes, handle the expected tasks for the patient and other users, and handle the various possible failure modes of the distributed system.

2. The system gives good advice.

3. The system is judged to be useful by the users.

4. The system improves patient outcome or compliance.

Because each of these levels of evaluation rests on the previous level and we are testing a completely new paradigm in medical assistance, most of our effort will, of necessity, be focused on the lower levels of evaluation. However, we will do some evaluation at each level. The objective of this evaluation is to determine whether the system merits the expense and time required for a full clinical trial. This is a formative evaluation, so we will make improvements to the program as we uncover limitations. The following paragraphs discuss what we will do to evaluate the program at each of these levels.

To judge whether the system successfully implements the desired performance characteristics, we will design a series of scenarios that test the behavior of the system for all of the desired functionality in the two medical domains. The scenarios will include examples of all of the desired kinds of communication the system will have to handle. In addition, they will include the range of likely failures in these communications. For example, we will test whether the system does appropriate alternative notification in case of emergencies when the primary destination is unavailable. Because the ultimate intention is to have a system that assists patients for their lifetime, we will test that the system state is reliably maintained by testing that the system is able to recover from the failure of any single component. We will include scenarios that involve the addition of new modules to the system to test its extensibility. When it is clear that the desired functionality is well in hand, we will commence with the second level of evaluation. For both of the medical domains, we need to test the domain knowledge of the GA. For this we will use a combination of designed scenarios and, later, cases constructed from data gathered from patient use of the GA. We will use a set of scenarios that covers as wide a spectrum of decision situations as possible to give the domain knowledge a thorough test. We will collect the analysis, actions, and advice of the system at each decision point in the scenarios. These decisions will be given to domain experts to judge. The experts will be asked to judge each decision as good, acceptable, or wrong. The decisions judged wrong will be reviewed to determine the basis for the judgment and the appropriate way to deal with the problem. Once the program has achieved a level of domain performance in which the mistakes are infrequent, the system will be advanced to the third level of evaluation.

The object of the third level of evaluation is to determine whether the system is perceived by the patient and physician to be useful. This is a necessary hurdle for the system to clear because the system depends on acceptance, especially by the patient, for it to be used and to serve its function. To conduct this evaluation we will select a small number of patients (probably 5) who have each target disease and equip each with the system for three months, under supervision of a physician who will have the corresponding part of the system for the physician. For the hypertension domain it may be necessary to give the GA to the patient for a longer period of time, perhaps up to six months, because interactions with the system will be less frequent than with diabetes. We will enlist patients at different stages of their diseases in order to get a better idea of the strengths and weaknesses of the system. The system will be instrumented to collect comments from the patient and physician, as well as the complete record of the use of the system. We will design this test to minimize the potential risk to the patient in collaboration with our Institutional Review Board (IRB). For example, all advice given to the

patient will also be given to the physician. Any advice that involves a change in therapy (such as a response to side-effects) will require the prior approval of the physician. These safety features will have to be balanced against the responsiveness of the system. After the first session of use, at an intermediate point, and at the end of the study period, we will ask the patient and physician to fill out questionnaires about the strengths and weaknesses of the GA. This field trial will also be a vehicle for testing whether the performance characteristics of the system hold up in practice. In particular, it will give us a good test of the reliability of the input data and the willingness of the patients to comply with the GA over an extended period of time.

Using the data from this field test of the system, we will analyze the record, looking for instances in which actions were taken as a result of the information provided by the GA system. These will be classified as to whether they increased or decreased the need for physician intervention and hastened or delayed optimal therapy. The resulting balance sheet will be used as a rough measure of the amount of improvement that could be expected from the GA system in each domain. The expectation is that this will provide enough data to do the power analysis for a full clinical evaluation. Such an evaluation will be planned, but it will not be carried out during the proposed project period.

Conclusion

We expect to publish reports that delineate the domain of active intelligent patient-centered health information systems, architectural documents specifying the reference architecture of Guardian Angel and its interfaces, a succession of prototype software implementations of GA and its applications, and research reports describing the above and evaluations conducted to assess the validity of our assumptions and (ultimately) the efficacy of the underlying concepts. We expect the example applications to measurably improve medical care, and we will also make those applications available to interested adopters. Because we plan to make available both a reference architecture and a set of reference implementations, potential adopters can choose whether to build new implementations based on the architecture or to adopt our reference implementations. To the extent possible, we would like to make the reference implementations available to others at little or no cost, but because they are likely to include components that are commercially developed, this may not be possible. We can guarantee for-fee availability of the reference implementation.

Schedule

We propose a three-year project, incorporating the following tasks:

Year one:

> Define GA initial reference architecture
>
> Select initial implementation strategy: specific medical, institutional and hardware environments.
>
> Identify protocol-based studies of care for selected medical foci.
>
> Begin development of knowledge base for selected medical foci.
>
> Begin selection or implementation of tools that will generate prototype

Year two:

> Complete implementation of authoring support system
>
> Implement prototype GA patient management systems for two selected medical foci.
>
> Integrate implementation with institutional information system(s)
>
> Simulate use with experimenters acting as surrogate patients
>
> Plan evaluation based on anticipated capabilities

Year three:

Continue evolution of the knowledge base for the selected applications

Conduct limited field trial(s) and perform their evaluation

Revise the architecture and implementation based on trial results

Plan and conduct technology transfer activities to assure dissemination of architecture and implementation

Part III

Project Partners

The project will be done as a collaboration among the Clinical Decision Making Group at the MIT Laboratory for Computer Science and researchers at two Boston-area hospitals, two companies, a large HMO, and a part of the Veterans Administration.

IBM Research has agreed to collaborate with us in basic research towards building health care solutions and plans to contribute to this program some appropriate existing and under developed software tools under fee-waived license agreement. IBM's Dr. Ifay Chang, who heads the Applications Solutions Institute (ASI) within the T. J. Watson Research Laboratory, is currently leading a very large-scale effort to develop a modern clinic-based information system for the Kaiser Colorado health plan. As part of that activity, ASI has developed a number of tools that we expect will be useful for prototyping and that will probably serve as reference implementations for portions of our reference architecture. These include tools for application definition and information and workflow modeling, for defining and maintaining a medical lexicon, and for building clinician's workstations from abstract specifications. As a commercial company, it is reasonable to expect that IBM will be interested in turning these tools (or their offspring) into commercial products. For this research program, Dr. Chang has agreed to work with us to develop mechanisms that would assure potential adopters of our system the ability to license, at a reasonable cost, components of IBM technology that we have incorporated in our reference implementations.

Gensym, Inc., is a small software company in Cambridge, tracing its roots to earlier MIT projects about 15 years ago, and is the most successful commercial vendor of tools for intelligent process control. Their graphically-oriented expert system shell, G2, has been used to build hundreds of systems for keeping track of evolving manufacturing processes, and they have built various tools to augment G2 for diagnostic reasoning and for interface to heterogeneous databases (two examples relevant to our proposal). Lowell Hawkinson, CEO of Gensym, has agreed to grant to this project six licenses to Gensym's commercial tools G2, GDA and associated support tools, and will dedicate one staff member (if funded via this proposal) to help develop the architecture and reference implementation for long-term GA agents. Gensym's products are available for commercial purchase by any interested user.

One of the concerns of all the involved parties is the question of liability for the medical knowledge contained in our implementations, and its potential impact on individual patients. We plan to maintain a strict separation between the software tools that we develop or adopt and the medical content that is encoded with those tools. This is not only good engineering practice, but will also alleviate concerns about legal liability on the part of the tool providers.

Our other partners in this project, Tufts/New England Medical Center, The Children's Hospital, The Boston Development Center of the VA, and Kaiser Colorado have no proprietary claims to any existing programs or to the software that we develop. They will have rights to the formalization of medical expertise contained in the system, to the extent of their contributions to it.

Although we have not at this stage made final decisions about the exact nature of the collaborative relationship, everyone involved in the project agrees to the following principles: (1) The GA architecture and ideas should not be encumbered by proprietary claims, except perhaps those assigned to a consortium that is committed to public dissemination. (2) Products contributed by participating commercial vendors shall be freely available for use during the conduct of this project, but without ceding commercial rights. (3) Vendors commit to making available, either as products or under special licensing arrangements, any components of reference implementations produced.

To date, there are no proprietary claims by any external party on the ideas of this project. The MIT researchers named in this proposal have full control over all intellectual property rights. Our intent is to maintain unencumbered all such rights to the architecture and technical specifications of Guardian Angel, so that if the project is successful we will have the greatest likelihood of being able to encourage its widespread use.

By the end of this project, we anticipate having created a domain-specific software architecture for GA agents and their supporting environments. Further, we plan to have defined a reference architecture appropriate to contemporary computational and communication technologies. In addition, we plan to have produced one or more reference implementations

consistent with this architecture, both to support our experimental applications and to give potential adopters an easy place to "get on board". (This is, we believe, one of the main positive lessons from the successful X-consortium run from our laboratory during the past decade.) The cost of building such reference implementations from scratch would be prohibitive, and we have organized a collaborative team intended to make this process much more efficient.

Previous Accomplishments

This proposal brings together a moderately large, heterogeneous group of computer science researchers, physicians, and software developers to explore a new paradigm of using computing to improve health care. The major participants are:

> Computer scientists at MIT: Drs. Peter Szolovits, William J. Long and Jon Doyle
> Collaborating physicians in Boston: Stephen G. Pauker, M.D., Mark Eckman, M.D., and John Wong, M.D., of New England Medical Center, and Isaac Kohane, M.D., Ph.D., of The Children's Hospital
> Michael L. Meistrell, M.D., Chief of Medical Informatics at the VA's Boston Development Center
> Gensym, Inc.: Lowell Hawkinson, CEO, Steven Fraleigh, and a software engineer
> IBM Corporation, specifically the Applications Solutions Institute of the T. J. Watson Research Laboratory, directed by Dr. Ifay Chang
> Kaiser Permanente of Colorado, currently implementing a clinical information system with the IBM group.

We sketch the background of each of these groups below.

MIT Laboratory for Computer Science

Attempts to build intelligence into computer programs operating in health information systems began with early programs in the 1960's based on flowcharting and simple statistical methods. G. Anthony Gorry, the founder of our group at MIT, for example, explored sequential Bayesian methods for optimizing diagnostic reasoning. Based on difficulties of these early methods, researchers, including us, in the 1970's turned to simulating the cognitive processes of expert human clinicians as a basis for building expert advisory systems. Many of these early "AI in Medicine" (AIM) systems used clinical associations and rules, and encountered grave difficulties in clinical problems involving interactions among multiple disorders, partially-treated diseases, and the need to monitor ongoing treatment. As a result, we led a move toward "deeper," model-based medical reasoning systems starting in the late 1970's, using inferences about pathophysiological models at various levels of detail to deal with unanticipated situations. Our research has broadened to include problems of explanation generation, to use more complex probabilistic and heuristic inference models, to learn from experience, to deal with time-dependencies systematically, and to model and use preferences in decision-making. We have also contributed to medical knowledge representation, KR in general, truth-maintenance, qualitative reasoning, and modeling repetitive decision-making. We illustrate some of these developments by a brief historical recap of our more significant projects.

Gorry introduced the sequential Bayesian diagnostic method in 1967 [Gorr68]. It uses information theory to select "most informative" questions, and decision analysis to decide which costly tests and treatments are worth doing. Applications in congenital heart disease, acute renal failure, and others, illustrated the strengths and limitations of the simple independence assumptions embedded in the probabilistic model. Nevertheless, the technique proved quite valuable in such medical specialties as managing the diagnosis and treatment of Hodgkin's disease.

Our first diagnostic program that attempted to model the reasoning processes of clinicians was the "Present Illness Program" (PIP), reported in *American J. Med.* in 1976 [Pauk76]. It adopted a hypothetico-deductive framework for diagnostic reasoning, using strong cues from the patient presentation to trigger hypotheses, both logical criteria and a pseudo-probabilistic scoring scheme to confirm or eliminate hypotheses, and explicit differential links to revise hypotheses when discrepant information arose. Later versions introduced a simple model of time, categorizing both patient data and a hypothesis-oriented time line along the dimension: past, recent-past, now, near-future, future. Gratifyingly, this facility led unexpectedly to a simple but reasonable prognostic ability, as PIP could use current data to predict future instances of disease. Our interest in temporal reasoning has continued through the doctoral

work of Kohane [Koha86, Koha87], exploring temporal constraints in diagnostic reasoning; Russ [Russ90, Russ91], who designed a control structure that supports reasoning about unreliable streams of time-oriented data and applied it to diabetic ketoacidosis; and Haimowitz [Haim93, Haim93a], who studies trend detection in pediatric growth data and in ICU monitoring. Wu's PhD in 1992 [Wu92] demonstrated dramatic speedups of classical associational diagnostic reasoning using two heuristic principles: problem decomposition to identify clusters of findings, and occasionally temporarily ignoring information distinguishing alternative diagnoses, thus containing the exponential explosion of intermediate diagnostic states.

At the same time as PIP's development, we created a program to advise on digoxin therapy, using a hybrid inference method that employed pharmacokinetic modeling to assess drug levels, but a clinically-derived experiential model that suggested how to titrate ideal doses given the patient's clinical state, evidence of possible digitoxicity, and urgency of care. This program also served as the vehicle for several programs on explanation and justification. Perhaps the most interesting was Swartout's PhD thesis [Swar83], in which he used an automatic programmer to generate the digitalis advisor from underlying medical knowledge and treatment principles, which then formed the basis for a sophisticated explanation capability that guaranteed consistency with the program's actual operation. Other explanation work used sophisticated English language generation techniques to turn program fragments and causal models into comprehensible natural language.

Our frustration with programs based on associational knowledge led, around 1980, to the first model-based medical reasoning programs. In Patil's ABEL program [Pati81] we modeled the acid-base and electrolyte domain in terms of multiple levels of pathophysiologic detail. Then, for cases in which associational models led to discrepant predictions, the program could instantiate a deeper model to explore causal mechanisms that can explain the discrepancy. For example, a patient with independent sources of acidosis and alkalosis could exhibit a normal pH, a finding not predicted by either condition but consistent with their combination. Long's Heart Failure program (HF) [Long92, Long92a, Long94], addresses the complex treatment of patients with heart failure, providing both a diagnostic and therapy planning component. Diagnosis is based on an approximate probabilistic method that works over a network of clinically-significant causal concepts, and therapy prediction is based on predicting the influence of possible interventions in a complex feedback system by using signal-flow analysis techniques. HF has proven to be quite effective at diagnosis in certain subdomains, and remains under active development to augment its diagnostic acumen and to further develop and test its therapeutic side.

Because hand-building programs is so difficult, we pursued the use of case-based experience to augment and make more efficient the operations of a model-based reasoner. Koton's PhD project, Casey [Koto89, Koto89a], adapted past case solutions to new cases that were similar as judged by their causal similarity, given by HF's causal models. Jang's dissertation [Jang93] enhances this approach by decomposing cases to be learned into their causally-coherent subparts (again as determined by HF), thus matching parts of a new problem and assembling the parts into a globally consistent account. We have also studied various statistical learning and classification techniques applied to real-world data.

Our colleagues at Tufts/NEMC have pioneered the application of decision analysis for individual patient cases [McNe82, Pauk75, Pauk87, Pauk88], and we have contributed via the development of computer-based tools and models. The Bunyan decision-tree critiquer [Well89], built in the mid-1980's is an expert system that examines decision trees and critiques both gross and subtle errors. For example, it would note that a diagnostic test whose outcome does not influence any later decisions is not being modeled correctly. More subtle critiquing expertise could notice that some phenomenon is being modeled in greater detail on one branch than on another—a common but subtle modeling bias. Leong, in her doctoral thesis [Leon94], is completing a new tool to support reasoning about long-term outcomes of medical interventions by building semi-Markov decision models from a graphical and medically-meaningful treatment description. Szolovits has recently built a prototype genetic counseling program, Geninfer [Szol92], that uses Bayes network solution methods to perform risk calculations to estimate the probability of heritable diseases. This program is being readied for trials under NIH funding. He is also involved in more fundamental explorations of probabilistic

reasoning; with Andersen of Aalborg and Shachter of Stanford [Shac94], we have some new results on the relationship between cutset conditioning and clique-tree formation.

Because much of medical knowledge is imprecise, we continue to explore qualitative descriptions of medical processes and to develop techniques that yield useful inferences from less detailed models. Kuipers began his work on QSIM [Kuip86] in our group in the mid-1980's, developing qualitative models of cardiovascular physiology. Sacks' 1988 PhD [Sach90] explored the use of sophisticated state-space representations to get much stronger characterizations of some qualitatively-described systems and extended these to derive probabilistic predictions with Doyle [Doyl91a]. Wellman's 1988 PhD [Well90] introduced the qualitative probabilistic model, which abstracts from particular numerical dependencies to more generic concepts such as "positively influences" and "interacts synergistically with." Although such models cannot make detailed tradeoff decisions, they provide very robust models that can be used to show dominance of whole classes of possible plans over others, and thus can greatly simplify planning in large plan spaces. Yeh in his 1990 PhD [Yeh90] explored a variety of inference methods to determine bounds on possible behaviors of probabilistically-defined systems.

Doyle and Wellman continue to investigate suitable representations for individual preferences [Doyl91b, Doyl94, Well91, Well92]. Their characterization of "preference all other things being equal" provides a powerful formal tool for abstract reasoning, without requiring a fully-specified multi-attribute utility function. This work continues with theoretical investigations aimed at applying it to capture basic preference attitudes of the subjects of GA.

We also have a long-term participation in knowledge representation efforts. Hawkinson and Szolovits worked in the mid-1970's on the OWL [Szol77] and BrandX [Szol81] representation schemes that provided great flexibility and opportunities to exploit linguistic analogies but suffered from a lack of semantic rigor. When current more restrictive KR systems were built in the 1980's, we tried to use KL/ONE to represent medical knowledge and found that too much expressive ability had been sacrificed for semantic cleanliness and computational efficiency [Haim88, Haim88a]. Doyle and Patil produced a major and influential critique of this trend for the KR community [Doyl91].

Long and Szolovits are both Fellows of the American College of Medical Informatics, and Doyle and Szolovits are both Fellows of the American Association for Artificial Intelligence.

Our group has pursued a wide variety of research projects in diagnostic and therapeutic reasoning, we have made fundamental contributions to AI and to medical informatics, and we believe that we are uniquely well-situated with our collaborators to make fundamental advances toward our vision of the "Guardian Angel" approach to health care.

As an illustration of the background of the group, we attach as an appendix a selected listing of reprints often requested by correspondents.

Tufts/NEMC

The faculty at the Division of Clinical Decision Making at the New England Medical Center have been developing decision support for health policy and for individual patients over the past two decades. They were the first physicians to apply decision theory to the care of individual patients. In a 1975 article, they described acquiring and using patient utilities in deciding about coronary bypass surgery for three patients with angina. They are acknowledged as the premier group in the area. Their software has become the standard for medical decision- and cost-effectiveness analyses. They have collaborated with faculty and supervised students at MIT-LCS for two decades and continue a close working relationship.

Their interest in patient utilities dates back to the paper addressing the need for surgery in coronary artery disease. In 1976, they developed models and techniques that allow patients to more fully participate in decisions about prenatal diagnosis. Both of these models have been extensively applied to individual patients over the past two decades, the latter to over 1500 patients. In now classic articles in the *New England Journal of Medicine*, they describe the acquisition of patient preferences (utilities) in patients with lung cancer and cancer of the larynx, and in other publications describe the impact of preferences on both individual decision making and health policy. They have been developing platforms for the automated acquisition and delivery of patient preferences for the past five years.

Although the group spans all of medicine in their clinical interests, they have had a particularly impressive publication record in cardiovascular disease, including coronary artery disease, valvular heart disease, cardiac arrhythmias and hypertension. They have been involved in multi-centered studies, developing decision support tools funded by the AHCPR,

with collaborators at Duke University, the University of Minnesota, the University of Manitoba, UCSF, and the New York State Department of Health. They were major participants in the recent Short-Term Mortality Conference on coronary surgery, held at Duke University, which emphasized the importance of patient specific factors such as co-morbidities.

These investigators also have substantial interest in adult diabetes and have been the major modeling group for the Diabetes PORT, funded by the AHCPR. They are just completing an analysis of the treatment of foot infections in diabetics, an analysis showing the importance of patient preferences and co-morbidities. Their interest in AIDS and HIV disease extends back over six or seven years, including articles emphasizing the interpretation of testing in low risk populations and the early treatment of HIV positive individuals.

Much of their theoretical work has involved the assessment of co-morbidities and their impact on survival, patient preferences, and decision making. Dr. Stephen Pauker, the director, is Professor of Medicine at Tufts University and a Fellow of the American College of Cardiology, the American College of Physicians, the American Heart Association and the American College of Medical Informatics. Dr. Mark Eckman is Associate Professor of Medicine at Tufts University and is a Fellow of the American College of Physicians and the American College of Medical Informatics. Dr. John Wong is an Assistant Professor of Medicine at Tufts University and is a Fellow of the American College of Physicians.

Children's Hospital

Isaac S. Kohane first began to work on automated diagnosis as a doctoral student in 1985. He joined the Harvard Medical School faculty and the staff of Children's Hospital in July, 1992. He holds a medical degree and practices pediatric endocrinology.

His PhD was awarded for his research in artificial intelligence in medicine and for his work on the Temporal Utility Package [Koha87] and its application to medical diagnosis. Dr. Kohane currently spends 80% of his time on work in medical informatics. In 1991 he completed the implementation of an on-line medical chart (the Clinician's Workstation—CWS) [Koha90] for the Division of Endocrinology. This system has now been in full operation for 3 years and provides on-line access to clinic notes, clinic measurements, demographics, pharmacy data, laboratory results, problem lists and reports from ancillary departmental systems (e.g. radiology). The CWS is currently being brought to the nephrology and nuclear medicine divisions at Children's Hospital. Over the past year, Dr. Kohane designed and led the implementation of a data integration and display system for the Multidisciplinary Intensive Care Unit. Data-sets on diabetes monitoring from this integration effort were distributed over the Internet for use by participants in the Spring Symposium of the American Association for Artificial Intelligence on "Interpreting Clinical Data" in Palo Alto (co-chaired by Dr. Kohane).

Dr. Kohane's current research is in the investigation of computer representations and architectures for clinical event monitoring. He is particularly interested in monitoring pediatric populations for early detection of pathological processes. In collaboration with Ira Haimowitz at MIT, he has developed a system for early detection and diagnosis of pathology from on-line clinical data. He has helped conduct preliminary trials to validate the methodology, especially in the domain of growth and development.

Dr. Kohane's clinical practice is split between general pediatric endocrinology and the care of patients with insulin-dependent diabetes mellitus (IDDM). Of the latter, he follows approximately 50 patients and supervises the care of another 50. He has conducted several clinical trials to study the growth and development of several populations including children with IDDM, with heart disease, with adrenal disorders, and children treated with bone-marrow transplantation for malignancy.

Veterans Health Adm.: Boston Dev. Center

The Boston Development Center is a field support unit serving the needs of senior VHA management located in Washington, DC. The BDC provides medical informatics support to improve the linkage of clinical processes, budget processes, and resource allocation processes. The BDC develops and maintains large relational databases that provide patient-specific data, including workload, and costing, as well as limited diagnostic and medical procedure information for all patients treated by the VHA, which maintains 172 medical centers, and many outpatient clinics, nursing homes, and domiciliaries. The VHA system provides approximately one million inpatient admissions and 45 million outpatient clinic stops annually. Approximately 2.5 million unique individuals are treated each year.

Dr. Michael L. Meistrell is chief of medical informatics at this center. He practiced as a thoracic surgeon, and did fellowships in medical decision making with Dr. Robert Beck at Dartmouth and Dr. Stephen Pauker at Tufts/NEMC. Dr. Meistrell is interested in working with us on the definition of the GA architecture, and will help to recruit physicians in the VA system for trials of our applications.

International Business Machines Corp.

Applications Solutions Institute, T. J. Watson Research Laboratory

The Applications Solutions Institute at IBM's T. J. Watson Research Laboratory is organized to re-establish IBM's strong presence in the health care information business by developing and deploying new technologies based on open systems to solve real customers' needs. IBM has developed a strategy for the development of clinical information systems in which several critical areas have been identified: (a) clinical repository and data modeling, (b) medical lexicon authoring tool and server, (c) application and work flow modeling tool and process manager, (d) clinical workstation and user interface, (e) communication and system interface, and (f) decision support. IBM's project with Kaiser Colorado was initiated in 1992, designed with the above architecture specification and the intention to expand into multimedia and image support, and to multimodal input technologies.

The IBM group has also prototyped an individual patient medical record system. This prototype is based on object-oriented techniques, and provides a small, portable machine with about 500MB of storage on a disk, and data import/export capabilities to a number of platforms [Chan91].

Our group at MIT has been tracking the development of this IBM system and its deployment at Kaiser Colorado. We view many of these facilities as useful components of a reference implementation of GA prototypes.

The group at Kaiser Colorado is also quite interested in the Guardian Angel concept, and has expressed an interest in participating after the IBM-developed clinical information system is installed. Today, part of the system is in alpha-test mode. Kaiser Colorado and IBM are currently developing a medical lexicon authoring tool. The schedule of the IBM-Kaiser Colorado effort matches very well with the proposed Guardian Angel schedule. We plan to pursue the Kaiser Colorado collaborative opportunity (with IBM) before we enter year two of the project, because it provides not only an excellent test-bed in addition to the VA, but also a significant point of leverage toward eventual adoption of our ideas by a major health care provider.

Gensym Corporation

Gensym is a small software corporation that provides intelligent real-time process control software to a world-wide client base. Its principal product, G2, is a system for developing and deploying complex applications that require continuous and intelligent monitoring, diagnosis, and control. It is designed for a wide range of real-time applications in such fields as process control, telecommunications, manufacturing, aerospace, medicine, finance, and robotics.

G2 utilizes a distributed, client/server architecture. This architecture provides the ability to do the following: (a) Scan an application, as a human operator would, and focus on key areas when it detects potential problems or opportunities. (b) Reason about and control events in a continuously changing environment. (c) Respond to events when they occur (e.g. without continually polling sensors). (d) Apply knowledge represented in multiple forms, e.g., rule-based heuristics, procedures, graphical networks of objects, and dynamic or discrete-event models. (e) Apply this knowledge where appropriate, using object-oriented representations that closely model the actual domain. (f) Express and use relationships between objects. (g) Represent the permanent and transient aspects of an application. (h) Acquire information from any number of data sources, both local and remote. (i) Provide information to and respond to requests from users, both locally and at remote client nodes. (j) Communicate with other G2s and applications (for example, with external simulation programs).

Over 1500 G2 licenses are installed worldwide. G2 is increasingly applied in pharmaceutical, biomanufacturing, and health-care applications. For example, G2 is being used in a clinical laboratory in Ontario that will process 15,000 blood samples per day. The system monitors and makes decisions on work flow, system trends and alarms; balances workload among multiple analyzers; releases normal results; and refers abnormal results for further

testing, human interpretation, etc. Gensym has considerable experience interfacing G2 to corporate databases. Gensym is interested in working with our project and shares our vision of how to influence the future of health care.

Results, Products and Transferable Technology

We plan to assure dissemination of our results and technology transfer in the following three ways: First, as is normal with research projects, we will publish papers and reports, give conference presentations, and encourage our students and colleagues to pursue research projects based on what we learn during the conduct of the proposed research. Second, we plan to develop both an architecture of Guardian Angel and reference implementations as applied to the specific medical areas we pursue. These will be made available to other researchers, and can form the basis for future projects and products. Third, through our collaborations with IBM, Gensym, New England Medical Center, the Children's Hospital, Kaiser Colorado, and the Veteran's Administration, we expect to imbue both industry and potential prime supporters of our approach with enthusiasm for GA. This enthusiasm will be based, we expect, on their successful cooperation in defining the GA architecture, helping to bring about its implementation, and observing the success of its applications.

Our project is likely to produce interesting new scientific ideas as we address difficult problems of modeling and architecture, and as we challenge existing technologies and build new ones. Our strong emphasis on tailoring the actions of the GA to the needs and preferences of the patient will result in new insights into issues of patient modeling and the expression, elicitation and use of preference models. As we have often done in the past, we will continue to use the difficult challenges of medical reasoning and knowledge representation to test ideas of knowledge representation, which have often proven inadequate to the tasks posed by medical reasoning. Our emphasis on frequent and seamless communication and exchange of information will provide interesting models for the evolving national information infrastructure. Our use of knowledge representation and interchange languages developed in current knowledge-sharing efforts should also yield opportunities for testing and validation of these efforts. We also expect to apply existing methods and develop new expert systems techniques for incremental interpretation of data and for user modeling. Finally, if our beliefs about the importance of the Guardian Angel concept for medical care prove correct, we can achieve a significant improvement in the delivery of medical care. Even if the concept is worked out with only partial success, we will have identified a very important set of research problems for future attention. These results will be disseminated by traditional scientific publications.

We also expect that the architecture we develop for GA, the reference implementations we build for our chosen application domains, and the tools we assemble and construct will form a valuable set of concepts and engineering artifacts that others may find useful. Because we plan to make available both a reference architecture and a set of reference implementations, potential adopters can choose whether to build new implementations based on the architecture or to adopt our reference implementations. To the extent possible, we would like to make the reference implementations available to others at little or no cost, but because they are likely to include components that are commercially developed, this may not be possible. We can guarantee for-fee availability of the reference implementation. We hope that our architecture will be adopted both by other researchers and by those more interested in building new applications. If interest warrants, we will establish a consortium between interested parties to promote and guide future development.

A less tangible but very important result we hope to achieve is to generate a commitment among our partners and others who provide health-care services and health-care technologies to the principles of patient-centered computation and record keeping. American medicine is very likely to undergo a dramatic change with the introduction of government-mandated universal health-care coverage and with the growing emphasis on cost-effectiveness of care. In such a time of upheaval, we hope to be able to create a paradigm shift in the application of informatics techniques as well. Guardian Angel may be the right vehicle.

Bibliography

[Baal94]. Baalen, J. V., Fikes, R. E. The role of reversible grammars in translating between representation languages. Proceedings of Fourth International Conference on Principles of Knowledge Representation and Reasoning (KR'94), 1994.

[Ball91]. Ball, M. J., Douglas, J. V., O'Desky, R. I., Albright, J. W. Healthcare Information Management Systems, Springer-Verlag, 1991.

[Bell88]. Bellomo, G., Santucci, S., Mannino, D., Alessi, R. A simple computer program for insulin dose adjustment in diabetic patients. *Comput Methods Programs Biomed* 1988, 26, 257-8.

[Bern92]. Berners-Lee, T. J., Cailliau, R., Groff, J.-F., Pollermann, B. "World-Wide Web: The Information Universe", *Electronic Networking: Research, Applications and Policy*, Vol. 2, No. 1, pp. 52-58, Spring 1992.

[Bith92]. Bithoney, W. G., Dubowitz, H., Egan, H. Failure to thrive/growth deficiency. *Pediatrics in Review* 1992, *13*, 453-460.

[Brei84]. Breiman, L., Friedman, J., Olshon, R., Stone, C. Classification and Regression Trees, Wadsworth, Pacific Grove, CA, 1984.

[Buch92]. Buchanan, B. G., *et al.* Involving patients in health care: explanation in the clinical setting. Proceedings of the Annual Symposium on Computer Applications in Medical Care. 1992.

[Buys89]. Buysschaert, M., Jacques, D., Donckier, J., Lambert, A. E. Effect of compter-assisted insulin delivery on glycemic control of type I diabetic patients: a preliminary experience. *Acta Clin Belg* 1989, 44, 169-73.

[Chan91]. Chang, I., Dimitroff, D. An object oriented approach to automated patient medical records. Proceedings for the Fourteenth Annual International Computer Software and Applications Conferencee. 1990.

[Chia90]. Chiarelli, F., Tumini, S., Morgese, G., Albisser, A. M. Controlled study in diabetic children comparing insulin-dosage adjustment by manual and computer algorithms. *Diabetes Care* 1990, 13, 1080-4.

[DCCT93]. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 329(14):977-86, 1993.

[Dick91]. Dick, R.S., Steen, E.B. The Computer-Based Patient Record: An Essential Technology for Health Care, Institute of Medicine, 1991.

[Doyl91] Doyle, J., Patil, R. S. Two theses of knowledge representation: Language restrictions, taxonomic classification, and the utility of representation services. *Artificial Intelligence,* 48(3):261-297, April 1991.

[Doyl91a] Doyle, J., Sacks, R. P. Markov analysis of qualitative dynamics. *Computational Intelligence*, 7(1):1-10, 1991.

[Doyl91b] Doyle, J., Shoham, Y., Wellman, M. P. A logic of relative desire (preliminary report). Ras, Z. W., Zemankova, M., editors, Methodologies for Intelligent Systems, 6, volume 542 of Lecture Notes in *Artificial Intelligence*, pages 16-31, Berlin, October 1991. Springer-Verlag.

[Doyl94] Doyle, J., Wellman, M. P. Representing preferences as ceteris paribus comparatives.. Hanks S., Russell, S., Wellman, M. P., editors, Proceedings of the AAAI Spring Symposium on Decision-Theoretic Planning, 1994.

[Fiel92]. Field, M. J., and Lohr, K. N., *Guidelines for Clinical Practice: From Development to Use*, Institute of Medicine, 1992.

[Fuge 92].   Fuge, C. F., Assal, J. P. Designing computer assisted instruction programs for diabetic patients: how can we make them really useful? Proc. Annual Symposium Computer Applications in Medical Care 1992, 215-9.

[GALE93].   Generalized Architecture for Languages Encyclopaedias and Nomenclatures in Medicine, GALEN Deliverable 6, The Master Notation, Version 1, 1993.

[Gene92].   Genesereth, M. R., Fikes, R. E. Knowledge Interchange Format Version 3.0 Reference Manual. Report Logic-92-1, Computer Science Department, Stanford University, June 1992.

[Gorr68].   Gorry, G. A., Barnett, G. O. Sequential Diagnosis by Computer" *JAMA*, 205: 849-854, 1968.

[Haim88].   Haimowitz, I. J. Using NIKL in a large medical knowledge base. TM 348, MIT-LCS, Cambridge, MA, January 1988.

[Haim93].   Haimowitz, I. J., Kohane, I. S. "Automated Trend Detection with Multiple Temporal Hypotheses," IJCAI-93, Chambery, France, pp 146-151, 1993.

[Haim93a].   Haimowitz, I. J., Kohane, I. S. Encoding patterns of growth to automate detection and diagnosis of abnormal growth trends. Paper presented at the American Society of Pediatrics/Society of Pediatric Research Meeting, Washington, DC., 1993.

[Haim88a].   Haimowitz, I. J., Patil, R. S., Szolovits, P. Representing medical knowledge in a terminological language is difficult. In Symposium on Computer Applications in Medical Care, pages 101-105, 1988.

[Huff94].   Hufnagel, S., Harbison, K., Doller, H., Silva, J., Mettala, E. National Healthcare System Information Architecture: Using DSSA Scenario-Based Engineering Process. Annual Conference of the Healthcare Information and Management System Society of the American Hospital Association, February 1994.

[Jang93].   Jang, Y. Hydi: A hybrid system with feedback for diagnosing multiple disorders. TR 576, MIT-LCS, October 1993.

[HL7].   Health Level Seven Standards: An Application Protocol for Electronic Data Exchange in Health Care Environments, Version 2.0, HL7: Philadelphia, PA, 1989.

[Hora90].   Horan, P.P., *et al.*, Computer-assisted self-control of diabetes by adolescents. *Diabetes Educator*, 1990. **16**(3): p. 205-11.

[Hrip90].   Hripcsak G., Clayton, P. D., Pryor, T. A., Haug, P., Wigertz, O. B., van der Lei, J., "The Arden Syntax for Medical Logic Modules," SCAMC-14, pp 200-204, 1990.

[Huse89]. Huse, D. M., Oster, G., Killen, A. R., Lacey, M. J., Colditz, G. A. The economic costs of non-insulin-dependent diabetes mellitus. *Journal of the American Medical Association* 1989, *262*, 2708-2713.

[IOM92].   The Computer Based Patient Record. Institute of Medicine, National Academy of Sciences. 1992.

[Koha86].   Kohane, I. S. Temporal reasoning in medical expert systems. In R. Salamon, B. Blum, and M. Jorgensen, editors, MEDINFO 86: Proceedings of the Fifth Conference on Medical Informatics, pages 170-174, Washington, October 1986. North-Holland.

[Koha87].   Kohane, I. S. Temporal Reasoning in Medical Expert Systems. MIT-LCS TR-389, 1987.

[Koha90].   Kohane, I. S., McCallie, D. P. A Dynamically Reconfigurable Clinician's Workstation with Transparent Access to Remote and Local Databases, AMIA, 1990.

[Koto89].   Koton, P. A. A method for improving the efficiency of model-based reasoning systems. *Applied Artificial Intelligence*, 3(2-3):357-366, 1989.

[Koto89a].   Koton, P. A. Using experience in learning and problem solving. Technical Report 441, MIT-LCS, Cambridge, MA, March 1989.

[Kuip86] Kuipers, K. Qualitative simulation. *Artificial Intelligence*, 29:289-338, 1986.

[Lask93]. Lasker, R. D. The diabetes control and complications trial: Implications for policy and practice (Editorial). *New England Journal of Medicine* 1993, *329*, 1035-1036.

[Lawr86]. Lawrence, M. A computer generated patient carried health check card. *Journal of the Royal College of General Practitioners*, 1986. **36**(291): p. 458-60.

[Leon94]. Leong, T.-Y. An Integrated Approach to Dynamic Decision Making under Uncertainty. PhD thesis, MIT, Cambridge, MA, May 1994. In preparation.

[Lisk90]. Liskov, B. "Implementation of Thor," Mercury Design Note 46, MIT Lab for Comp Science, Cambridge, MA, 1990.

[Long92]. Long, W. J., Naimi, S., Criscitiello, M. G. Development of a Knowledge Base for Diagnostic Reasoning in Cardiology. *Computers and Biomedical Research* 25:292-311, 1992.

[Long92a]. Long, W. J., Naimi, S., Criscitiello, M. G., Adusumilli, R. K.``Validation of a causal probabilistic medical knowledge base for diagnostic reasoning,'' In Deep Models for Medical Knowledge Engineering, E. Keravnou, ed., Elsevier Sciences Publishers B.V., pp. 147-160, 1992.

[Long94]. Long, W. J., Naimi, S., Criscitiello, M. G. ``Evaluation of a New Method for Cardiovascular Reasoning,'' to be published by *Jour. Amer. Med. Info. Assoc.*

[McDo90]. McDonald, C. J., Tierney, W. M., Martin, D. K., Overhage, M., Day, Z., Adams, P., Blevins, L., Cassidy, P., Glazener, T., Lee, M., Lemmon, L., Mamlin, B., Meeks-Johnson, J., Porterfield, B., Warvel, J. A., Warvel, J. S. The Regenstrief Medical Record: 1990. In: Proceedings Symposium on Computer Applications in Medical Care. R. A. Millers, Eds., Washington, DC: IEEE Computer Press, 1990:934-938.

[McNe82]. McNeil, B. J., Pauker, S. G., Sox, Jr., H. C., Tversky, A. On the elicitation of preferences for alternate therapies. *New England Journal of Medicine*, 306:1259-1262, 1982.

[Mill84]. Miller, P. L., Black, H. R. Medical plan-analysis by computer: Critiquing the pharmacologic management of essential hypertension. Computers and Biomedical Research 17:554-569, 1984.

[Muse89]. Musen, M. A. Automated Generation of Model-Based Knowledge-Acquisition Tools, Morgan-Kaufmann, San Mateo, CA, 1989.

[NIH93]. *Report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure*. NIH Publication 93-1088, National Institutes of Health, 1993.

[Pati81]. Patil, R. S. Causal representation of patient illness for electrolyte and acid-base diagnosis. TR 267, Massachusetts Institute of Technology, Laboratory for Computer Science, 545 Technology Square, Cambridge, MA, 02139, October 1981.

[Pauk88]. Pauker, S. G. Clinical decision making. In Wyngaarden, J. B., Smith, L. H., editors, Cecil Textbook of Medicine, pages 74--79. Saunders, Philadelphia, 1988.

[Pauk75]. Pauker, S. G., Kassirer, J. P. Therapeutic decision making: A cost-benefit analysis. *New England Journal of Medicine*, 293:229--234, 1975.

[Pauk76]. Pauker, S. G., Gorry, G. A., Kassirer, J. P., Schwartz, W. B. Towards the Simulation of Clinical Cognition: Taking a Present Illness by Computer. *American J. Med.* 60:981-996, 1976.

[Pepp92]. Peppiatt, R. Eliciting patients' views of the cause of their problem: a practical strategy for GPs. *Fam Pract*, **9**(3), 295-8, 1992.

[PHS91]. "Healthy People 2000: National health promotion and disease prevention objectives," (PHS) 91-50213, Department of Health & Human Services, 1991.

[Pick87]. Pickering, T.G., Devereux, R. B. Ambulatory monitoring of blood pressure as a predictor of cardiovascular risk. *American Heart Journal*, 1987. **114**: p. 925-928.

[Pryo83]. Pyor, T. A., Gardner, R. M., Clayton, P. D., Warner, H. R. The HELP system. *Journal of Medical Systems* 1983, *7*:87.

[Quin93]. Quinlan, J. R. C4.5: *Programs for Machine Learning*, Morgan Kaufman, San Mateo, CA, 1993.

[Rive78]. Rivest, R., Shamir, A., Adelman, L. A Method for Obtaining Digital Signatures and Public-Key Criptosystems. MIT-LCS, 1977.

[Russ90]. Russ, T. A. Using hindsight in medical decision making. Computer Methods and Programs in Biomedicine, 32(1):81-90, May 1990.

[Russ91]. Russ, T. A. Reasoning with Time Dependent Data. PhD thesis, MIT-EE/CS, Cambridge, MA, September 1991. MIT/LCS/TR-545.

[Sack90]. Sacks, E.P. Automatic qualitative analysis of ordinary differential equations using piecewise linear approximations. *Artificial Intelligence*, 41(3):313-364, January 1990.

[Schi94]. Schivers, O. BodyTalk and the BodyNet: The Architecture of a Personal Communications Infrastructure, MIT-LCS, 1993

[Shac94]. Shachter, R. D., Andersen, S. K., Szolovits, P. The equivalence of exact methods for probabilistic inference on belief networks. Submitted for publication.

[Smit92]. Smith, D. R. Transformational approach to scheduling. Technical Report KES.U.92.2, Kestrel Institute, November 1992.

[Stat76]. Statland, B. E., Winkel, P. *Limitations of Using Laboratory Data to Allocate Patients into Clinical Classes: Consideration of the Time-Course of a Disease and the Inter-Individual Variations in Set Points as Determined in the Basal State Free from Disease*. Computer-Assisted Decision Making Using Clinical and Paraclinical (Laboratory) Data, Mediad Inc., NY 1980.

[Stra85]. Strack, T., Bergeler, J., Hutten, H. Computer assisted conventional insulin therapy. *Life Support Sys*t. 1985, 1, 568-72.

[Suth91]. Sutherland, H. J., *et al.*, Communicating probabilistic information to cancer patients: is there 'noise' on the line? *Social Science & Medicine*, 1991. **32**(6): p. 725-31.

[Swar83]. Swartout, W. R. Xplain: A system for creating and explaining expert consulting programs. *Artificial Intelligence*, 21:285-325, 1983.

[Szol77]. Szolovits, P., Hawkinson, L., Martin, W. A. An overview of OWL, a language for knowledge representation. TM 86, MIT-LCS Cambridge, MA, 1977. Also in Rahmstorf, G., and Ferguson, M., Eds., Proceedings of the Workshop on Natural Language Interaction with Databases, International Institute for Applied Systems Analysis, Schloss Laxenburg, Austria, January 10, 1977.

[Szol81]. Szolovits, P., Martin, W. A. Brand X: Lisp support for semantic networks. In Proceedings of the Seventh International Joint Conference on Artificial Intelligence, pages 940-946, 1981.

[Szol92]. Szolovits, P., Pauker, S. Pedigree analysis for genetic counseling. In K. C. Lun et al., editor, MEDINFO 92: Proceedings of the Seventh Conference on Medical Informatics, pages 679-683, 1992.

[Tenn93]. Tennenhouse, D. et al. (40 authors). "I-95: The Information Market". Technical Report MIT/LCS/TR-577, Massachusetts Institute of Technology, Laboratory for Computer Science, August 1993.

[UMLS91]. UMLS Knowledge Sources: 2nd Experimental Edition, National Library of Medicine, 1991.

[Verm90] Verma, T. S., and Pearl, J. Equivalence and Synthesis of Causal Models. Sixth Conference on Uncertainty in Artificial Intelligence, 220-227, 1990.

[Warn72]. Warner, H. R., Olmstead, C. M. , Rutherford, B. D., "HELP -- A Program for Medical Decision-Making," *Computers in Biomedical Research* 5:65-74, 1972.

[Well90]. Wellman, M. P. Formulation of Tradeoffs in Planning Under Uncertainty. Pitman and Morgan Kaufmann, 1990.

[Well91]. Wellman, M. P., Doyle, J. Preferential semantics for goals. In Proceedings of the National Conference on Artificial Intelligence, pages 698-703, 1991.

[Well92]. Wellman, M. P., Doyle, J. Modular utility representation for decision-theoretic planning. In Proceedings of the First International Conference on AI Planning Systems, 1992.

[Well89]. Wellman, M. P., Eckman, M. H, Fleming, C., Marshall, S. L., Sonnenberg, F. A., Pauker, S. G. Automated critiquing of medical decision trees. *Medical Decision Making*, 9(4):272-284, 1989.

[Will92]. Williams, B. C., Miller, C. A. Preventive health care for young children. Findings from a 10-country study and directions for United States policy.*Pediatrics* 1992, *89*, 983-998.

[Will91]. Williams, B. T., Imrey, H., Williams, R. G. The lifespan personal health record*. Medical Decision Making*, 1991. .

[Wu92]. Wu, T. D. A decompositional search algorithm for efficient diagnosis of multiple disorders. PhD thesis, MIT-EE/CS, Cambridge, MA, June 1992.

[Yeh91]. Yeh, A. Automatic analysis of systems at steady state: Handling iterative dynamic systems and parameter uncertainty. TR 525, MIT-LCS Cambridge, MA, December 1991.