Formal Representation of a Conceptual Data Model for the Patient-Based Medical Record

A. Gouveia-Oliveira, M.D.(1), Luísa Lopes, B.S. (2).
(1) Dept. Biomatematica, Faculdade de Medicina de Lisboa, 1700 Lisboa, Portugal
(2) Companhia IBM Portuguesa SA, Pr. Alvalade 7, 1799 Lisboa Codex, Portugal

We present a general architecture for the patient-based medical record as it is being developed for the SAMS, a private social security system. The conceptual data model is described in a convenient formal notation, the entity-relationship diagram. Although following the original formulation of the problem-oriented medical record (POMR), the data model was designed with a level of generalization that, functionally, makes structural differences between conventional and POMR no longer apparent. The main features of this model are its adaptability to individual work practices and its problem-oriented structure, including the representation of problems' evolution. This structure will enable physicians to organize the data, mostly collected elsewhere, by explicitly relating the facts that constitute a particular patient record, which is a simple way to store context information and clinical knowledge that is not part of patient data.

INTRODUCTION

The SAMS is a private social security system managed by the syndicate for banking personnel from southern Portugal that provides health care for over 120,000 affiliates. Its structure is being reorganized, in order to integrate a busy outpatient clinic, a recently built 137-bed hospital and a health-home for long-term care. The new health system will be organized around the outpatient clinic, which will be the major health-care center. The hospital is seen as a back-up health unit where patients may be referred to for special forms of treatment and/or diagnosis that are beyond the possibilities offered in the outpatient clinic. In a way, the hospital will be offering services primarily to health-care providers, not to patients, in contrast to the usual concept of conventional health care systems that give the central role to hospitals.

This organizational structure is being implemented in coordination with the development of a Clinical Information System (CIS). The SAMS' CIS is based on Medsolution 400/BASE, a commercial applicational offering by IBM for IBM AS/400 computers, whose major functions are patient management, order management, reporting, and admission-discharge-transfer. Previously existing custom systems for accounting and billing and laboratory management are being interfaced with the base system, and new modules, to be integrated with Medsolution, are under development. These include several subsystems that will manage multiple operational needs, like appointing and scheduling, prescriptions, operating rooms, etc.

It is the purpose of this project to make patient information available to physicians and to the administration, for patient care and for resources management. However, in order to achieve both goals, the system will have to gather the information stored throughout the system and organize it centered on the patient. In other words, patient data must be brought to a working space where the user may have easy access to it. The logical solution would be to create a model of the medical record, as this is the long established method for the storage and organization of clinical data and for expressing the clinical reasonings and the decisions taken upon a specific case.

Although research is very active on the development of clinical data models for specific domains, implementation of those models into relational databases and integration of the few available standard data models into CIS, reports on the literature about models for patient-based medical records are scarce. In general, papers have been mostly concerned with the functional aspects. The purpose of this paper is to introduce a formal representation of a general model for the medical record that was based on the theoretical and practical considerations that make the foundations of present-day medical records' structure. This model is now being implemented on the SAMS' CIS.

METHODS

The entity-relationship (ER) model [1] was selected as the primary designing tool. This data model is easily implemented into relational database systems, and a well standardized diagrammatic technique, the ER diagram, is available to express the information that is being modeled and the structure of the database. In fact, most CASE (Computer-Assisted Software

Engeneering) tools for designing relational schemes require that data models are specified in terms of ER model concepts. The formal notation that was adopted here is based on the Information Engineering Methodology [2], which is widely known to the industry and, probably, one of the most widely used.

In the ER model, the construct primitives are entities and associations. In the ER diagram, entities are shown as rectangles and associations as arcs connecting entities. Both entities and associations have an identifier (a name). Some properties of the associations are also represented in the diagram: entities may be forced to exist in an association (which is represented by a single dash on the side of the entity involved) or not (represented by a circle); a second symbol, nearer the end of the arc, represents the degree of the association: associations may be one-to-one (represented by an arc ending with a single dash) or one-to-many (represented by a three-branched arc). Many-to-many associations and complex associations (involving more than two entities) are replaced by two one-to-many associations plus a new entity, represented by a dashed rectangle.

RESULTS

Figure 1 shows the principal entities and the relationships among them. The model starts with the concept of a person's clinical file, which is the identification of one person in terms of clinical information. Although this information is actually split among several files, the PERSON's clinical file is the reference for all the clinical data concerning that person.

Each person has a set of problems, the PERSON'S PROBLEM LIST. The problems are identified by the physician, who is given the freedom to name them at his/her own will. At least one problem must exist for each person with a clinical file, and if none is declared it will be assumed that the person has an Unspecified Health Problem. Problems do not have to be defined or structured previously, and in fact most of them will not. However, there may exist a subset of problems for which a systematized study would be desirable, due to some particular interest from the organization or a group of doctors. These problems will be PREDEFINED PROBLEMS, and for them an Initial Plan will be established. The Initial Plan is a set of structured data, defined by the organization or a group of specialists, that must be recorded whenever such problem is identified in a person. The data of the initial plan for a given problem of a specific person is the PERSON'S PROBLEM INITIAL PLAN.

The evolution of a PERSON'S PROBLEM is recorded through its EPISODES, the set of which is called the PERSON PROBLEM'S DIARY. In an outpatient clinic environment, each episode of a problem corresponds to the occurrence of an encounter; in each ENCOUNTER several problems can be addressed, hence several PROBLEM'S EPISODES can be created. For each PROBLEM'S EPISODE several pieces of information that correspond to the proposed SOAP (Subjective, Objective, Assessment and Plan) structure for progress notes can, optionally, be recorded or created: a) The EPISODE'S DATA, an entity with the attributes patient-originated data, physical examination and comments or assessment, and b) the PRESCRIPTIONS, which is the Plan component of SOAP and includes any clinical event other than the actual encounter: MEDICATION, patient EDUCATION and ORDERS (for diagnostic examinations, treatments, referrals, hospitalization, etc.). The ORDERS, whether they are or not permormed during the encounter, will then generate RESULTS, which are linked to the originating requisitions and problems. In the actual implementation, the results will be sent to the system by the performing departments and automatically referenced to that person problem's episode.

Besides the evolution of a person's problem considered as a succession of episodes, problems themselves are evolutionary concepts. This was reproduced in the model as far as the problems' type (active, inactive) and modifications (resolved, dropped, modified or grouped) are concerned. In particular, the latter two changes affect their designation. Additionally, problems may be related to other problems, which is represented in the diagram by a recursive relationship in the PERSON'S PROBLEM entity. These relationships are explicit and of types SECONDARY-TO, CAUSED-BY and ASSOCIATED-WITH. The evolution of problems and their relathionship to other problems is recorded in the PROBLEM'S HISTORICAL, where all types of changes, their date and the physician who made them, will be kept.

Three other entities were included in the model: the CLINICAL HISTORY, the DISEASE code and the PROTOCOL. The CLINICAL HISTORY is stored as free text and is optional, although it is recommended to be created in the first encounter. It is a recording of the person's health status in a particular moment and is not to be modified or updated. The DISEASE CODE, in the form of ICD-9-CM codes, is a common requirement in medical records. The ICD classification is independent of the identification of the problems but may be related to one or more problems. This classification is left as optional, at least in the first stage of implementation, in order to avoid an excessive workload on users not yet acquainted

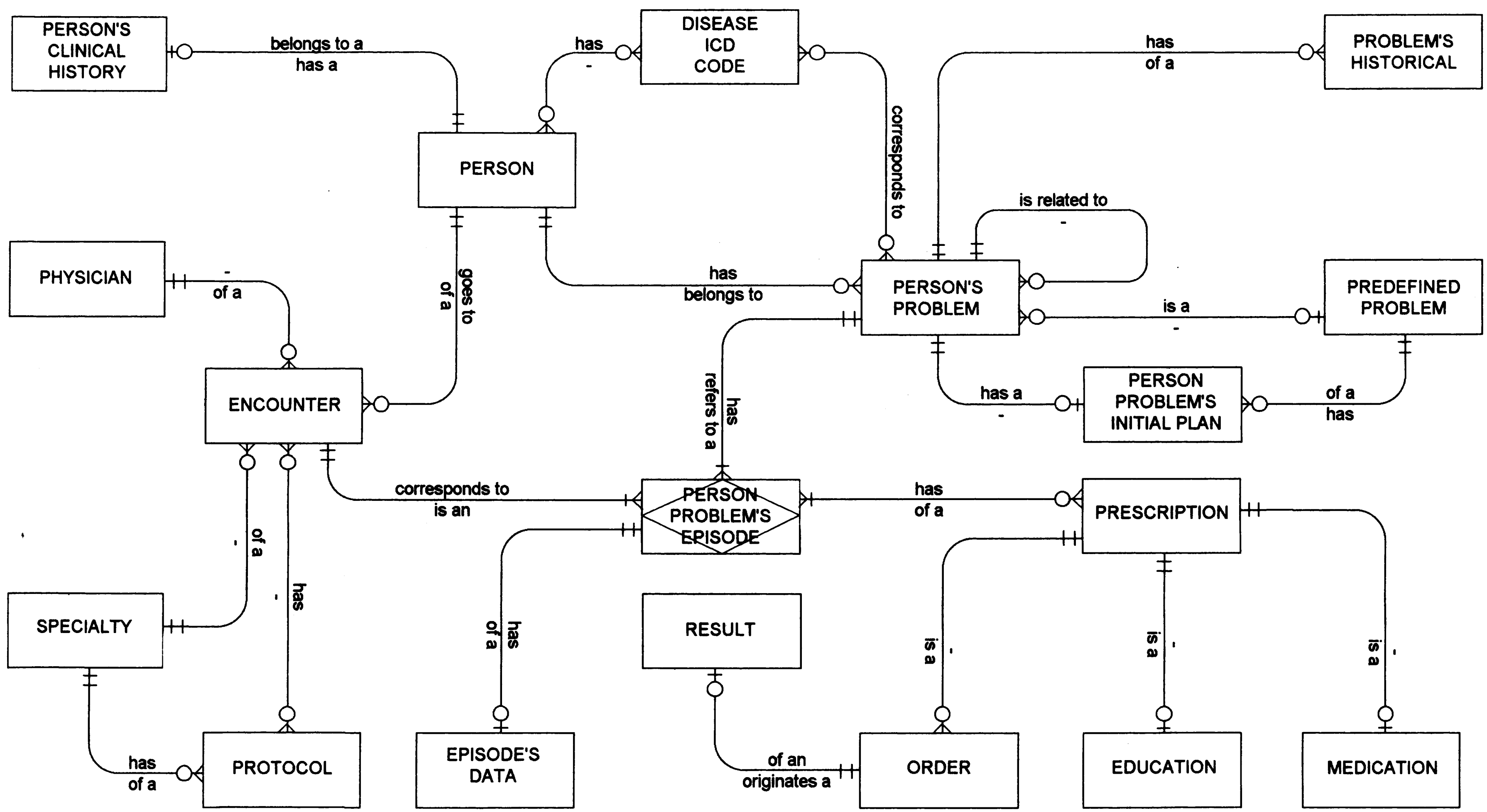

Figure 1. Entity-Relationships diagram of the patient-based medical record data model

with the system. A PROTOCOL is a set of user-defined structured data that has to be completed for every episode of a specific clinical specialty. As with the initial plans of problems, the protocols and their contents are defined by the institution or groups of physicians. Unlike initial plans, which are created only in the first occurrence of a problem, protocols are associated with every encounter of a clinical specialty. Examples of protocols are pregnancy, child development and hemodialysis.

DISCUSSION

Medical records store more than subjective and objective patient data. They also store information from general medical knowledge, for instance contextual information about some medical entity or event, or the clinical arguments justifying a particular medical event. Basically, this information consists mainly on the description of the medical entities and the relationships among them. Hence, although a large amount of patient data that is sent to a CIS originates in clinical and pathology laboratories and imagiology, the medical record is the entity that integrates all that information, and the physician is the only person with the habilitation to establish relationships among the data.

Data from medical records is also useful for purposes other than clinical care, but this may be accomplished reliably only if that data permits unambiguous interpretation. It may be the case, however, that there are ways of extracting useful information without requiring the data to be complete and explicit [3], and a system that stored explicit specification of the relationships among medical entities and medical events could increase that possibility.

The problem-oriented medical record (POMR) [4,5] addresses this issues by proposing that patient data be structured according to a series of problems identified by all the health-care providers concerned in his/her treatment, rather than on a structure based on the source of the data, as conventional medical records do [6]. In POMR, decisions taken during patient work-up are explicitly related to problems. Several improvements have been introduced in the problem-oriented practice method [7] but, despite numerous reports on its successful implementation in many diverse practices and the widely recognized interest of this approach, the implementation of POMR is still very limited. Difficulties encountered have been its redundancy, lack of criteria for problem formulation, compartmentalization of patient health status, time-consuming maintenance and difficulties in the process of information retrieval [6,8]. We believe these insufficiencies arise mostly from limitations of the paper platform, which is not at all adapted to cope with the evolving nature of the problems, an essential feature of the POMR approach. Implementing the POMR on a computer database system might help alleviate many of these problems and, in turn, that approach could provide more diverse views of the data and more accurate queries.

The development of patient-based medical records has generally been a built-up of experiences and improvements that eventually ended on a structure that may not correspond to the fundamental principles underlying that basic clinical tool. It is important to compare the proposed models with the theoretic aspects on which the clinical record is based [10,11]. For example, most reports on computer-based medical records state that they include a problem list or that they support SOAP. However, it may be inferred from those papers that, in most cases, those are only new elements that were added to the structure of a conventional medical and, actually, the basic structure was not built centred on the patient's problems. Further, the problem list is very often taken as a summary of the patient's active problems, which is distinct from the concept developed on the POMR's formulation. The active medication list is another new element that is commonly included, but it is not clear whether it should exist as an entity of its own or if it is merely one possible view of the data. Interpretation of current state-of-art is difficult because scientific papers do not present data models with objectivity. Proved methods for formal presentation of data models are available, however, and we believe that these methods could make an important contribution, by providing an objective basis for the communication of the accumulated experience and knowledge.

This model is actually centered on the idea that the most useful way of organizing clinical data is through the concept of problems and their evolution. The problem list is not only a means of bringing together information that is contained in a non-linear clinical record, but also a method for organizing it in a clinically meaningful way and to record its evolution.

Another feature of this model regards the representation of the dynamic nature of the problem list. Modeling the evolution of problems is very different from simply editing the problem list. We considered three ways whereby a problem may evolve: a) the successive recording of data regarding each problem as it is being studied over time, commonly known as progress notes, which is organized in the problem's diary; b) the modification of the entity identified as a problem, specifically the modification of its type (active, inactive, solved, discarded) or of its identification, as happens when a problem's name is changed or when two or more problems are merged in a single one; c) the modification of the relationships among problems. The problem's historical keeps all the information related to the evolution of the problems. Consequently, the concept of problem list, the kernel of this model, is essentially a dynamic one. This is an important feature of the model, if one wishes to guarantee that all the existing information will be displayed in answer to physicians' problem-directed queries.

It certainly has to be considered that many physicians may be reluctant to adopt the POMR structure for organizing their patient data. The design of the medical record system for an organization such as the SAMS requires flexibility enough to accommodate different work practices of over 200 physicians. Fortunately, the two methods are not conflicting, rather the conventional medical record may be seen as a POMR with a single patient problem, in which case the structure of both systems will be identical. We applied this concept to develop a data model with a level of generalization that makes it independent of each physician's work method.

Still on the account of the flexibility issue, we have decided in this particular implementation not to restrict a physician's selection of the appropriate name for a problem by resorting to a controlled vocabulary. This agrees with the views expressed and discussed by other authors [11,12], yet the general nature of the model will accommodate different opinions. It must be stressed, however, that one of the cornerstones of the POMR was the abandoning of the custom of expressing patient difficulties through disease nomenclatures, because the name of a disease was taken as an incomplete indication of all the prognostic and therapeutic distinctions that have to be made in the care of a patient [6]. Incidentally, this was also the reason why disease names were considered as a distinct entity in this model. Because problems are of a virtually infinite number, we anticipate severe difficulties when using an approach like a controlled vocabulary to impose standardization. We do plan to evolve to a controlled vocabulary of problem names for some restricted clinical domains and after some time of clinical utilization. The same is the case with the results of examinations, namely with imagiology reporting, which will be entered as free text, to be replaced in the future by structured data entry and automatic reporting.

The model presented in this paper is not only a conceptual one, but is also specified in accurate terms and, consequently, can be directly implemented in any Relational Database System. The kind of approach used for its description, using a formal methodology and notation for modeling the information has clear advantages as it enables a much greater objectivity in its specification. The Entity-Relationships model highlights the fundamental pieces of information that have to be stored and how they relate to each other. Specified in this formal way, the model can be used as an objective basis for discussion and, consequently, further enhancements and improvements may be proposed on solid grounds.

References

[1] Chen PPS. The entity-relationship model: towards a unified view of data. ACM Transactions on Database Systems 1976: 9-36.

[2] Martin J. Information Engineering. A Trilogy. (3 vols.). Prentice Hall, Englewood Cliffs, New Jersey 1989.

[3] Van Ginneken AM, Van der Lei J, Moorman PW. Towards unambiguous representation of patient data. Proc Annu Symp Comput Appl Med Care 1992: 69-73.

[4] Weed LL. Medical records that guide and teach. N Engl J Med 1968. 278: 593-600, 652-657.

[5] Weed LL. Medical records, medical education and patient care: the problem-oriented medical record as a basic tool. Press of Case Western University, Cleveland, Ohio 1969.

[6] Feinstein AR. The problems of the «Problem-Oriented Medical Record». Ann Int Med 1973. 78: 751-762.

[7] Hurst JW. Walker HK (editors). The problem-oriented system. Medcom press, New York 1972.

[8] Goldfinger SE. The problem-oriented record: a critique from a believer. N Engl J Med 1973. 288: 606-608.

[9] Bradbury AR. Computerized medical records: defining a standard without the computer. Med Inf Lond 1991. 16: 279-286.

[10] Rector AL, Nowlan WA, Kay S. Foundations for an electronic medical record. Methods Inf Med 1991. 30: 179-186.

[11] Safran C, Rury C, Rind DM, Taylor WC. Outpatient medical records for a teaching hospital. Beginning the physician-computer dialogue. Proc Annu Symp Comput Appl Med Care 1991: 114-118.

[12] Wilton R. Non-categorical problem lists in a primary-care information system. Proc Annu Symp Comput Appl Med Care 1991: 823-827.