TempR-PDM: A Conceptual Temporal Relational Model for Managing Patient Data

AQIL BURNEY[1], NADEEM MAHMOOD[1], KAMRAN AHSAN[2]
[1]Department of Computer Science (UBIT)
University of Karachi
Karachi-75270, Karachi
PAKISTAN
burney@uok.edu.pk http://www.drburney.net
nmahmood@uok.edu.pk
[2]Faculty of Computer Engineering and Technology & Health Sciences Department
Staffordshire University
UNITED KINGDOM
k.ahsan@staffs.ac.uk

Abstract: - Time is one of the most difficult aspects to handle in real world applications, especially in database systems. Relational database management systems proposed by Codd offer very little built-in support for managing time varying data and theory of temporal semantics. Many temporal extensions of the relational model have been proposed and some of them are also implemented. The proposed system is based on an ontology of health Care terminology including patients vital data, activities and as well as the treatment data. We will propose a conceptual temporal model for handling time varying attributes in the relational database model with minimal temporal attributes. The proposed model is easy to define, manage and incorporates the important and relevant features in the target temporal relational model. Furthermore we have illustrated implementation of the model on patient database and also present the requirement analysis: work flow for managing patients in hospital environment. In response to rising health care costs, reliability, privacy, security and changing expectations concerning the quality of health care, data management is very important in improving health care services.

Key-Words: - Relational database, temporal semantics, ontology, temporal relational model, health care management.

1 Introduction

The relational model [11] is based on a brand of mathematics called relational algebra. Codd used that concept to manage huge amounts of data very effectively. Codd and others have extended the notion to apply to database design. Thus they were able to take advantage of the power of mathematical abstraction and the expressiveness of mathematical notation to develop a simple but powerful structure for databases [20].

A relation has to be in first normal form (FNF), meaning that the domains of the attributes in its schema may only be of scalar data types. In other words, a relation can be considered as a subset of the Cartesian product of all the attribute domains contained in its schema.

The relational data model only support functionality to access a single state of the real world, called a snapshot. The transition from one database state to another (updates) thereby giving up the old state. There exist, however, many application domains which need to have access not only to the most recent state, but also to past and even future states, and the notion of data consistency must be extended to cover all of these database states. Due to the FNF assumption in the relational model, there is a restriction in expressing the data structures. To overcome this drawback, the relational model has been extended [13] to support, non first normal form (NFNF) or nested relations.

Efforts to incorporate the temporal domain into database management system have been ongoing for more than a decade and dozens of temporal models have been proposed [6], [8], [10], [17] and a few of them have been implemented [1], [9], [29], [30].

Designing effective, secure[28] and useful healthcare information systems which handles temporal data is a great challenge for software engineers. It includes very complex information that evolves with time. Relational model is a very powerful model and well accepted model among the vendors. There are number of extensions to this model which incorporates time varying data. In this paper we investigate the patient data management (PDM) with respect to the time varying nature of data and propose a conceptual temporal relational schema for PDM.

This paper is organized as follows:
Section2, discussion on various temporal relational models has been made. Section 3 deals with the ontology and a proposed requirement analysis: work flow for patient data management. Section 4 describes the proposed conceptual data model [5], (TempR-PDM) with its logical schema.

2 Discussion

Many data models are introduced so far to capture the semantics of temporal data keeping the traditional entity relationship model (ERM) approach. Traditional ERM is not capable for capturing the whole temporal aspects. Many extensions [19] have been proposed to extend the ERM to capture time varying information in one way or the other. Unified modeling language (UML) is also used as a tool to develop the logical and conceptual schema of the mini world.
The other important point is how this new conceptual model [5] will be incorporated into a relational database. One way of doing this is to develop a temporal layer and this layer is responsible for translating the temporal queries to traditional SQL statements. The other approach is to design a complete temporal query language [6] which not only supports all SQL statements but incorporate new operators based on temporal relational algebra [27]. There are many solutions for this problem and few successful implementations are also summarized in table 1.
Since most of the work in the research area of temporal databases has been done with respect to the relational data model because of its strong structure. Table 1 mentions some of the most important models with respect to the work presented in this paper. Temporal data models can be categorized as tuple time stamped or attribute time stamped, FNF or NFNF, valid time or transaction time [21].

Table. 1 Temporal Data Models: Brief Survey

Temporal Model	*NF*	*Time Stamping*	*Time*	*Features / Query Language*
Tansel's Model [32]	NINF	Attribute	VT	Non homogeneous
Clifford & Croker's Model [7], [8]	NINF	Attribute	VT	HDBM Inhomogeneous
McKenzie's Model[25][26]	NINF	Attribute	VT & TT	Extension of Snapshot algebra Historical algebra
Gadia's model [15][16]	NINF	Attribute	VT	Homogeneous HRQUEL
Ben-Zvi's Model [3]		Tuple	VT & TT	Time Relational model Extension of snapshot algebra
Snodgrass's Model [29]	NINF	Tuple	VT & TT	TQUEL
Lorentzos's Model [24]	1NF	Attribute	VT	Interval Relational Model (IRM) Extended Relational Model (XRM)
Jensen & Snodgrass's Model [22][23]	1NF	Tuple	VT & TT	Bi-temporal Conceptual Model (BCDM)
Ariav's model [2].	1NF	Tuple	VT & TT	Time stamps are based on time points
Gadia & Yeung's Model [17]		Attribute	VT & TT	Heterogeneous model
TSQL2 [30]	1NF	Tuple	VT, TT & User Defined Time	Homogeneous model

3 Requirement Analysis: Work Flow for Patient Data Management

3.1 Ontology

For a conceptual development an ontology which covers medical and hospital information is required to develop the PDM system [33]. Development of PDM-Requirement analysis: work flow in a hospital/clinical system includes the integration ontological knowledge base with the existing information system. This requires the creation of ontological categories based on the information in patient databases in a hospital. For details on the ontology in healthcare see [33].
Temporal aspect is significantly important in various fields and it can be very helpful in analyzing and understanding domain performance. We have proposed a requirement analysis work flow for patient data management in a hospital. The purpose of this model is to analyze the importance of patient's data to improve healthcare organizations service.
Fig.2 gives a work flow of patient data management by highlighting and categorizing the important and useful patient information in a hospital environment. The basic idea is to model a situation where we can distinguish temporal and non temporal data components. Treatment data is represented with a clock in the model, highlighting the fact that the treatment data is temporal in nature. Treatment may change with the passage of time and we must have to manage and retrieve that data when required, because time is one of the most important features while we want to improve health care process.

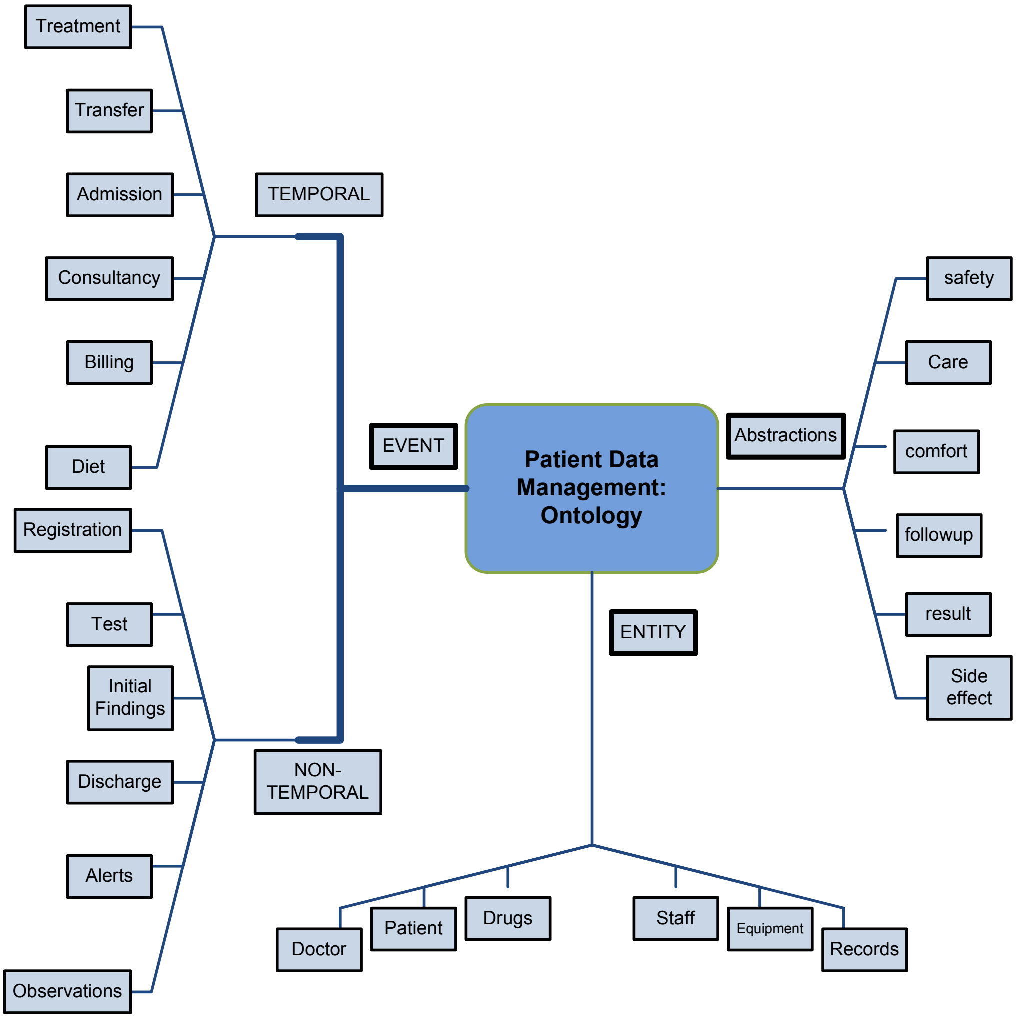

Fig. 1 Patient Data Management (PDM) Ontology

Similarly the data regarding the doctors who are looking after the patient and the para-medical staff who is performing duties in the wards has to be kept in the database. These relationships are represented with the dotted line and with the clock symbol. These relationships are called as temporal relationships [21]. Admission of patients in the wards is also an important factor, because when the patient came to the ward and when he/she is shifted to another ward and related information. The relationship is also classified as temporal relationship. The model depicts the information flow in a hospital and also represents the important work stations and information flow to the patient database and servers, with the objective to incorporate those features which best suit to patient database model [18].

4 Proposed Conceptual Model: TempR-PDM

By now the relational model is the most effective method for organizing huge amounts of data and still the widely accepted technology amongst vendors. enterprises all around the world. Temporal data has its own semantics and organization of such data requires some modifications in the relational model [4]. Attributes in a relation can be time varying or non-time varying attributes. Treatment relation is a temporal relation, because treatment changes with time and forms a treatment history, contrary doctor relation is a non temporal relation because it deals with the attribute which are non time varying.

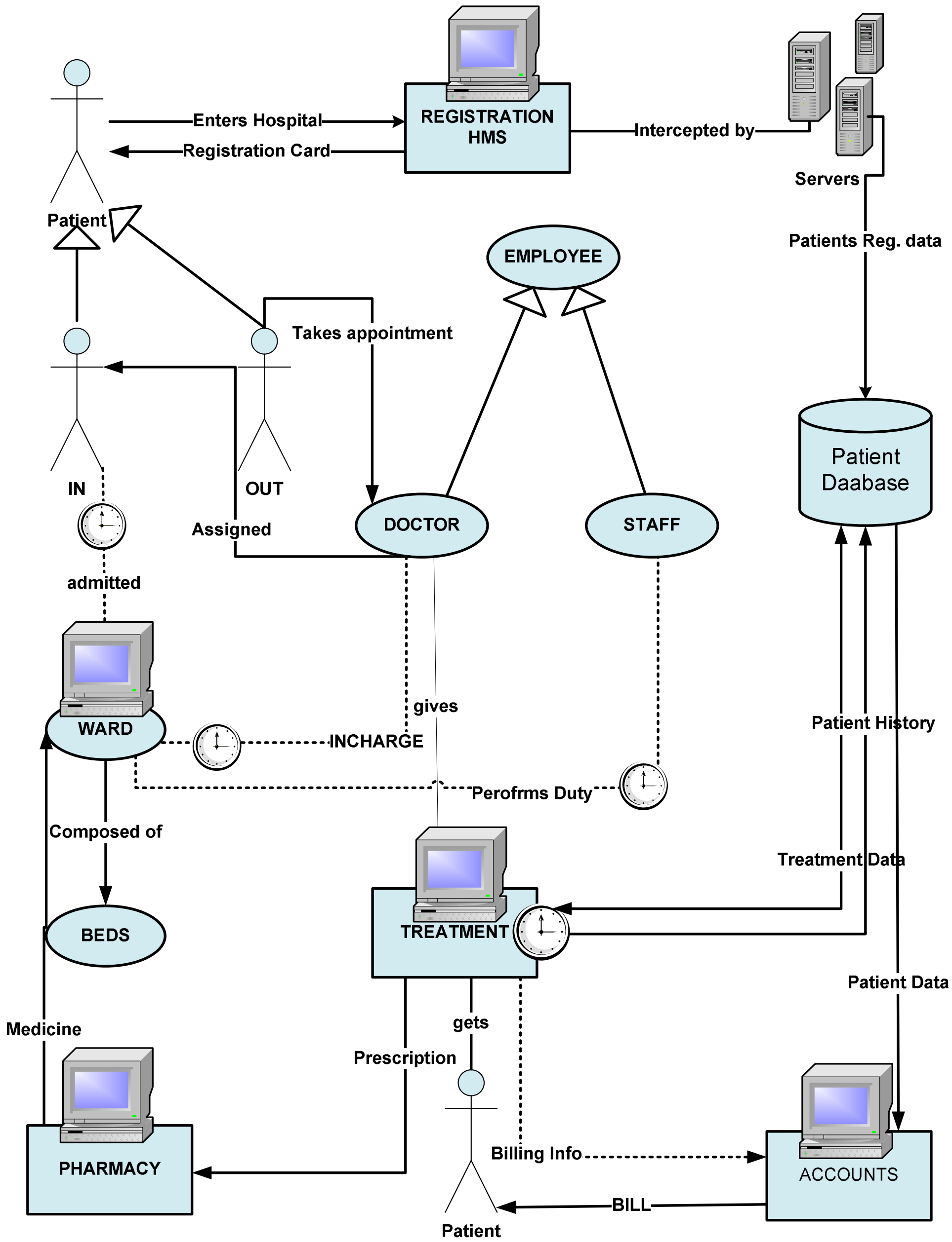

Fig. 2 Requirement analysis: Work flow for Patient Data Management

4.1 Structure of Time

We will take the Structure of the set Time to be < T,< >, where T is some countable set and < is a linear order on T, i.e. for any two times t1 and t2, either t1 = t2, t1 ≤ t2, or t2 ≤ t1. In our model we treat time as discrete and isomorphic to the natural numbers because any practical domain that we might define for time attributes in our proposed model would have at most an infinite countable set of names for time moments or time intervals.

Tuple time-stamping approach has been adopted to define a temporal model. The reason for this is the simplicity and to keep the 1NF assumption and the essence of the relational model.

The proposed logical temporal data model (fig. 3) comprises of three main constructs namely, the entity (temporal or no temporal), attributes (time-varying, non-time-varying and partial time varying attributes) and thirdly the relationship type amongst the entities are n-ary and categorized as temporal and non-temporal.

Following are some constraints to ensure the consistency of the conceptual schema:

- Fully temporal attribute will be handled by a separate entity called temporal entity.
- Entity is categorized as a temporal and non-temporal entity
- Temporal entity type must has a combinational primary key composed of time-varying and non-

time-varying attributes. The activation start time is the part of the key.

- The n-ary relationship determines weather it is temporal or non-temporal relationship.

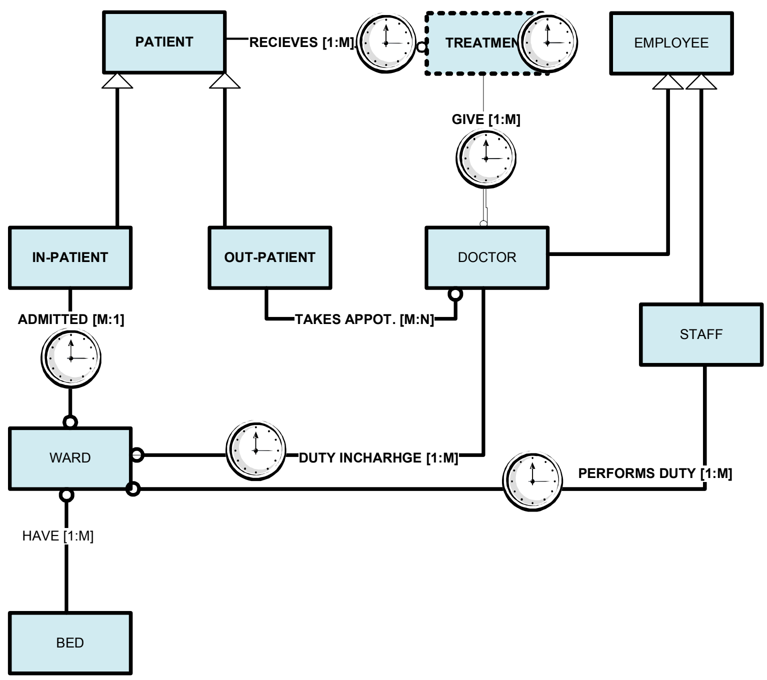

Fig. 3 TempR-PDM: Conceptual Temporal Relational Model for Patient Data Management.

4.2 Representing Temporal Characteristics:

We have introduced a new term as activation_start and activation_end rather than valid time because validity itself is of many types and it creates complexity while defining an entity.

- It represents temporal facts both at time points and as well as on intervals.
- Update time is introduced instead of transaction time. An update refers to change in data (tuple) of any sort (insert, delete or change).
- Activation time can be represented with different time granularities [12] such as year, month, week, day, hour, minute and second and even beyond that. Conversion from one time granularity [31] to another is accomplished by the conversion functions.

4.3 Temporal Relation Schema

A temporal relation schema TR =< A, K > is an ordered pair consisting of finite set of A = {A1, A2, . . . , An, activation_start, activation_end, updatetime}. Each temporal relation has a current state, and the changed states (future and past). Logical schema of a patient database are as follows.

PATIENT: {**Regno**, Pname, Paddress, Pcontact, Gender, Age, Date of Birth}

- Conventional (non temporal relation) no temporal functional dependency.

Note: Although address and contact are the attributes which can also change with the passage of time. This is not very common, usually these attributes donot change very often. For the simplicity of the problem we consider these attributes as weak temporal attributes.

DOCTOR:
{**Doctorid**, Dname, Department, Type, Pager, Cell

- Conventional (non temporal relation) no temporal functional dependency.

IN: {**Pregno**, **Wardno**, **Bedno**, Date/time, activation_start,activation_end, update_time, consultingdoctor

OUT: {**Pregno**, Date/time, consulting_doctor

WARD: {**Wardno**, **activation_start**, **Staff**, Wardname, ,activation_end, update_time.}

WARDDOCTOR:
{**Wardno**, **activation_start**, doctorid, activation_end, update_time.

BED: {**Bedno, Wardno**, Type}

STAFF: {**ID**, Duty hours, responsibility}
EMPLOYEE:
{**Empno**, Eaddress, Econtact,
EType, HireDate, BirthDate, designation, salary}
TREATMENT:
{**treatmentno**, **patient id, doctor id**, **activation_start**, treatment date, Provisional diagnosis, Remarks, Test Reports required, Next visit, , activation end, update time}.

The non temporal facts of employees are stored in a relation patient. Although the address attribute is also a time varying attribute, patient do not change address so often so it can be managed in the same relation, introducing a separate relation over here will increase overhead. .
Current time is a function which returns the current time. Minute is the default time granularity [46] for defining these relations but it can be changed by invoking conversion functions if required. The complete patient history of the employee can easily be recovered from the treatment relation and also from the other relations.
Update operations (insert, delete and change) result in insertion of a new tuple that contains the new values associated with a new timestamp. The new attribute value does not replace the previous (old) value of the attribute in case of an update such as deletion or insertion. Similarly, the deletion of tuples does not cause an actual deletion from database tables.

5. Conclusion

The proposed model can be easily extended to incorporate time varying attributes. In our model each time varying attribute, identified in the mini world has a separate relation called as temporal relation. The concept of temporal relationship is introduced. We have presented some of the most important temporal models and discussed the semantics of our proposed temporal schema of the patient database.
The benefits of the proposed model are not restricted to reducing temporal attributes in the relation but it also keeps the essence of the relational model and causes less overhead. Due to the simplicity of the model, it will be easier to implement.
The examples presented in this paper do not represent a complete picture; however this can be revised and scaled to a larger problem domain easily. It decreases development time and highlight the most important features of the target temporal relational database schema.

6. Future work and Evaluation

There are still some points which have not been addressed in this paper. The overhead reduced by this model in the case of insert, update and change operations should formally be estimated and verified through experimental results. The fact that some events might occur in the modeled reality, but the reflection of these events to the database is delayed, is out of the scope of this paper. We defer the discussion of these points to a future work.
In future we will evaluate this model in healthcare mobile system for the analysis of patients' movements to improve healthcare organizations service. Our other partner is working on healthcare care application domain in which context based knowledge management system model will be implemented to improve service for hospital patients. They are trying to integrate knowledge factor in changing patients' movement scenario and trying to capture this through mobile technology. The role of our temporal model will be vital use for their database implementation and integration which will further evaluate our model in healthcare domain. We are also working on the rule based system for patient management using Petri nets [14] and then transforming it to the temporal relational data model.

References:

[1] Ahn, SQL+T: A Temporal Query Language. *Proceedings of the Infrastructure for Temporal Databases*, Arlington, 1993.

[2] G. A. Ariav, Temporally Oriented Data Model. *ACM Transactions on Database Systems*, 11(4), 1986, pp.499-527.

[3] Ben-Zvi, The Time Relational Model. *Ph.D. thesis*, Computer Science Department, UCLA, 1982.

[4] M. H. Böhlen, and C. S. Jensen, Seamless Integration of Time into SQL. *TR R-96-2049*, Aalborg University, Aalborg, 1996.

[5] Carlo Combi, et. al., Conceptual Modeling of Temporal Clinical Workflows, *14th International Symposium on Temporal Representation and Reasoning* (TIME'07), 2007.

[6] Chomicki, Temporal Query Languages: A Survey. In Gabbay, D. and Ohlbach, H., editors, *Temporal Logic, First International Conference*, Springer-Verlag, LNAI 827, 1994, pp. 506-534.

[7] J. Clifford, and A. Croker, The Historical Relational Data Model (HRDM) and Algebra Based on Lifespans. In *Proceedings of the International Conference on Data Engineering*. IEEE Computer Society Press, 1987, pp. 528-537.

[8] J. Clifford, and A. Croker, The Historical Relational Data Model (HRDM) Revisited. In A.

Tansel, J. Clifford, S. Gadia, S. Jajodia, A. Segev, and R. Snodgrass, editors, *Temporal Databases: Theory, Design, and Implementation*, Benjamin/Cummings Publishing Company, 1993, pp. 6-27.

[9] J. Clifford, A. Croker and A. Tuzhilin, On the Completeness of Query Languages for Grouped and Ungrouped Historical Data Models. In [Tansel et al., 1993], (1993), pp. 496-533.

[10] J. Clifford, A. Croker and A. Tuzhilin, On Completeness of Historical Relational Query Languages. *ACM Transactions on Database Systems*, 19(1), 1994, pp.64-116.

[11] E. F. Codd, A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, 13(6), 1970, pp. 377-387.

[12] C. Combi, M. Franceschet, and A. Peron, Representing and Reasoning about Temporal Granularities. *Journal of Logic and Co1mputation*, 14(1), 2004, pp. 51–77.

[13] C. J. Date, H. Darwen, and N. A. Lorentzos, *Temporal Data and the Relational Model*. Morgan Kaufmann Publishers, 2003.

[14] Dong-Her Shih, Hsiu-Sen Chiang, Binshan Lin, A Generalized Associative Petri Net for Reasoning, IEEE Transactions on Knowledge and Data Engineering, (vol. 19 no. 9), 2007, pp. 1241-1251.

[15] S. K. Gadia, A Homogeneous Relational Model and Query Languages for Temporal Databases. *ACM Transactions on Database Systems*, 13(4), 1988, pp.418-448.

[16] S. K. Gadia, (1992) A Seamless Generic Extension of SQL for Querying Temporal Data. *Technical Report TR-92-02*, Computer Science Department, Iowa State University, 1992.

[17] S. K. Gadia, and C. S. Yeung, A Generalized Model for a Relational Temporal Database. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Chicago, IL, 1988, pp. 251-259.

[18] J. D. Gold, and M. J. Ball, The Health Record Banking imperative: A conceptual model, *IBM SYSTEMS Journal*, Vol. 46, NO 1, 2007.

[19] H. Gregersen, TimeERplus: A Temporal EER Model Supporting Schema Changes. In BNCOD, volume 3567 of *Lecture Notes in Computer Science*, Springer, 2005, pp. 41–59.

[20] J. A. Hoffer, F. R. McFadden and M. B. Prescott, *Modern Database Management*, 7th edition Upper Saddle River, NJ: Pearson/Prentice Hall, 2005.

[21] C. S. Jensen, and C. E. Dyreson, [editors]. The Consensus Glossary of Temporal Database Concepts - February 1998 Version. In *Temporal Databases: Research and Practice*, volume 1399 of Lecture Notes in Computer Science, Springer-Verlag, 1998, pp. 367–405.

[22] C. S. Jensen, and R. T. Snodgrass, The TEMPIS Project. Proposal for a Data Model for the Temporal Structured Query Language. *TEMPIS Technical Report No. 37*, Department of Computer Science, University of Arizona, Tuscon, 1992.

[23] C. S. Jensen and R. T. Snodgrass. Temporal Data Management. *IEEE Transactions on Knowledge and Data Engineering*, 11(1), 1999, pp.36-44.

[24] N. A. Lorentzos, The Interval-extended Relational Model and its Application to Validtime Databases. In A. Tansel, et. al., editors, *Temporal Databases: Theory, Design, and Implementation*, Benjamin / Cummings Publishing Company, chapter 3, 1993, pp. 67-91.

[25] J. E. McKenzie, An Algebraic Language for Query and Update of Temporal Databases. *Ph.D. thesis*, Computer Science Department, Univ. of North Carolina at Chapel Hill, 1988.

[26] J. E. McKenzie and R. T. Snodgrass. Supporting Valid Time in an Historical Relational Algebra: Proofs and Extensions. *Technical Report TR_91_15*, University of Arizona, Tucson, AZ., 1991.

[27] J. E. McKenzie and R. T. Snodgrass. Evaluation of Relational Algebras Incorporating the Time Dimension in Databases. *ACM Computing Surveys*, 23(4), 1991, pp. 501-543.

[28] Rafae Bhatti et. al. Engineering a Policy-Based System for Federated Healthcare databases, IEEE Transactions on Knowledge and Data Engineering, (vol. 19 no. 9), 2007, pp. 1288-1304.

[29] R. T. Snodgrass, The Temporal Query Language TQuel. *ACM Transactions on Database Systems*, 12(2), 1987, pp. 247-298

[30] R. T. Snodgrass, editor *The TSQL2 Temporal Query Language*. Kluwer Academic Publishers, 101 Philip Drive, Assinippi Park, Norwell, Massachusetts 02061, USA, 1995.

[31] S. Spranger, Calendars as Types – Data Modeling, Constraint Reasoning, and Type Checking with Calendars. *Ph.D. Thesis*. Herbert Utz Verlag, M¨unchen, 2006.

[32] Tansel, J. Clifford, S. Gadia, S. Jajodia, A. Segev, and Snodgrass, R. *Temporal Databases: Theory, Design, and Implementation*. Benjamin/Cummings Publishing Company, 1993.

[33] Velma Payne and P. Douglas P. Metzler, Hospital Care Watch (HCW): An Ontology and Rule-based Intelligent Patient Management Assistant, Proceedings of the 18th *IEEE Symposium on Computer-Based Medical Systems* (CBMS'05), 2005.