Computers and the Humanities **31:** 483–502, 1998.

HiSQL: A Front-end Query System for Historical Relational Databases

RAMAZAN ACUN[1], RACHID ANANE[2] and SUSAN LAFLIN[3]
[1]*Hacettepe University (acun@eti.cc.hun.edu.tr);* [2]*Business Information Systems, ELBS (R.Anane@uel.ac.uk);* [3]*University of Birmingham (S.Laflin@cs.bham.ac.uk)*

Key words: defters, historical processing, model design, relational databases, spatial processing, S.Q.L., temporal processing

Abstract. The paper describes a prototype system called HiSQL (Historical SQL) which extends the functionality of SQL in manipulating historical data, by providing functions for spatial and temporal processing. Conceptually the paper is divided into three parts: the first part deals with the design and architecture of the system; the second part introduces a case study (the defters); and the third part describes specific functions for spatial and temporal processing of serial documents. The paper concludes with a comparison between HiSQL and SQL and suggestions for further work.

1. Introduction

Despite the appearance of new database management systems (such as object management systems and deductive databases) in the market place over the last decade, relational database management systems continue to be widely used in many areas, including historical research. The query language SQL has become an industry standard and is likely to remain so for some time to come. However, it is well known that SQL has some serious limitations when it comes to manipulating data with temporal and spatial properties in general and historical data in particular.

In historical processing, the problem is that an intuitively simple historical question may require the writing of long and difficult multi-step queries when expressed in SQL. This lack of transparency is a cause of concern to the historian because it interferes with his/her main task – that is interacting with his/her data. A more expressive system is needed in order to support historical processing effectively.

In this paper we describe a prototype query system called HiSQL (Historical SQL) which extends the functionality of SQL in manipulating historical data. HiSQL may be applied to a relational database (in Oracle) containing information such as census or tax records, where the format of the original data implies an hierarchical organisation of the spatial data. This organisation is explicitly coded into the database (together with information about the time points (or dates) over which the hierarchies are valid) and then HiSQL allows queries of interest to

historians to be expressed in a very simple form. HiSQL is based on the conceptualisation not only of the nature of historical data but also of the general characteristics of the historian's task.

We shall first describe the architecture of the system and then present several examples of HiSQL queries and compare them with equivalent queries expressed in standard SQL. The example chosen to develop the system was a database of Ottoman tax registers (defters) between 1485 and 1569. It has also been tested with a small sample of nineteenth century census data from England and we believe it will be equally useful for processing similar data from elsewhere. Finally we shall discuss possible extensions to the system and indicate our future plans for it.

2. Design of HiSQL

This section describes the design and implementation of HiSQL, the front-end query system. The overall specification of the system was to provide an expression level language better suited to historical analysis (i.e. the interaction between the historian and historical data).

The design of the HiSQL was based on the notion of *explicit-model design*, which is defined as "a structure composed of knowledge in the form of an explicit representation of the properties of some entity" (Edwards and Mason, 1988, p. 316). This approach encourages modular design in the high-level components of the system. Consequently, this type of design provides great flexibility and is able to accommodate many different types of tasks and users. In addition, because of their use of modularity, explicit models make the design more understandable and easier to modify (Edwards and Mason, 1988, p. 317).

The various types of knowledge required for system functionality were organised into three explicit models, each capable of communicating with the other two. These were the **Dialogue Model**, the **Task Model** and the **User Model**.

2.1 TASK MODEL

The task model contains an explicit representation of spatial divisions, and their associated operations. In the database, spatial divisions are stored in a hierarchical (recursive) relation or tree-structure. In reality because of changes over time, the relationships between spatial divisions develop into a graph structure and the graph data structure contained in this model is a directed labelled graph. In this graph, the vertices (nodes) represent spatial divisions and the arcs represent connections between them. Each arc is labelled with a time interval over which the respective connection holds valid. A time interval is defined by a start-date and an end-date (which may be the same). For an example, see (Acun, 1994, p. 136).

Four operations for searching this graph have been defined: "**Stree**" (spatial tree), "**Schange**" (spatial change), "**Scontinuity**" (spatial continuity), and "**Vspace**" (virtual space). They are described (with examples) in section 4. These

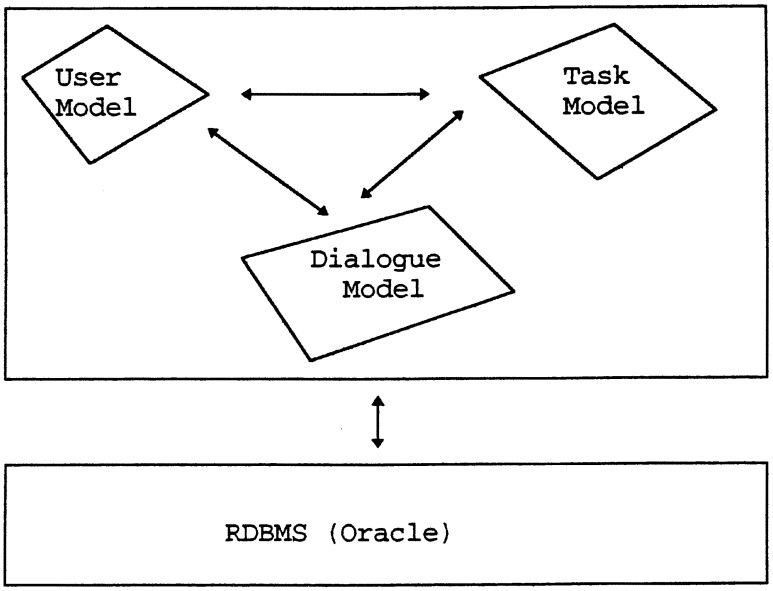

Figure 1. Architecture for HiSQL.

operations were designed either to be used alone or to be embedded in SQL statements to facilitate an easy retrieval of spatially oriented data for the purpose of spatial comparisons.

All these operations have the same two parameters (assuming that there is only one graph to be searched): **division** and **time-interval**. Division contains the vertex from which the search starts. Time-interval contains (implicitly) one or more time points and is used to check the validity of a connection which exists between any given pair of vertices. A detailed description of these operations can be found in (Acun, 1994, pp. 132–133).

2.2 The User Model

The user model contains information about the users. In particular their right to access to database objects and their area and period of interest. The latter is used to restrict the amount of information to be held in the main memory for a given user. This is required by the system supervisor, but has little other effect on the way the system is used.

2.3 The Dialogue Model

The Dialogue Model is responsible for communication between the user and the system i.e. getting the input from, and presenting the output back to the user. This model also handles much of the communication with the Oracle RDBMS. In particular it contains a two level SQL parser. Examples will be included to show

the HiSQL queries and compare them with the full SQL queries. The parser enables the system to translate from HiSQL to SQL automatically and so the user does not need to know the details of the database implementation or the full complexity of SQL.

3. The Sample Data

The data used to illustrate HiSQL is taken from the "*tahrir defters*", the tax records of the Ottoman empire from 1431 (the earliest extant defter) until the early seventeenth century. The survey which produced the defters was carried out at intervals ranging from ten to thirty years throughout this period. The form of the documents implies a hierarchical structure for the spatial data, with the empire consisting of provinces with a separate document or defter for each province. Each province(or *sancak*) is made up of regions (*nahiye*), which in turn consist of villages (*karye*), which may contain hamlets (*mezraa*). The defters contain details of the taxes due from each area and the person or institution to whom the taxes have been allocated.

In a defterological study, the administrative structure of an area may be of interest to a researcher in two cases. In the first case, this structure is the main purpose of the study. In the second and the commonest case, this structure is used as a framework to investigate other aspects of life in that area. In either case, the evolution of this structure over the period is of great interest. For examples see (AcunF, 1993; Oz, 1990; Ozel, 1993).

The sample data to be used in this paper are extracted from the Tahrir Database developed for the Ottoman tax registers of the 15th and 16th centuries (specifically the defters of 1485 (TD37), 1520 (TD387), 1547 (TD255), and 1569 (TD478) for the province of Karahisar in northeast Anatolia) see (AcunF, 1993; Acun et al., 1994a,b; Acun, 1994). A listing of this data is included in Appendix A and consists of two tables, "Test_areal_division" and "Test_tax".

Since the operations deal with the spatial dimension of the data over time, elaborating a little more on the entry of spatial information is in order. In the defters, the information about the administrative divisions and settlements consists of the following: the type of division, e.g. village, town, region; the name of the division; and an indication of the higher level division to which this one is connected e.g. a village may be connected to a region (nahiye) which in turn may be connected to a province etc. In the case of settlements i.e. towns, villages and hamlets, this is usually followed by a listing of inhabitants and a listing of taxes to be collected from this settlement.

When the information is entered into the database the spatial information (at the conceptual level) is stored in the table "areal_division" in the form shown below.

The first two columns, document number (Doc.No) and document date (D.date), identify the defter from which the information is taken. The third column (Div.id) contains a unique division identifier number assigned to this spatial division (this is not the record's key, which is a composite key consisting of Doc.No, D.date

Table I. Spatial divisions

Doc.no	D.date	Div.id	Up.div	Div.type	Div.name
TD37	1485	2	1	SANCAK	KARAHISAR
TD37	1485	2341	2	NAHIYE	SERIN
TD37	1485	53	2341	KARYE	SUNUGER
TD255	1547	2	1	SANCAK	KARAHISAR
TD255	1547	37	2	NAHIYE	SIRYAN
TD255	1547	53	37	KARYE	SUNUGER
TD478	1569	2	1	SANCAK	KARAHISAR
TD478	1569	37	2	NAHIYE	SIRYAN
TD478	1569	53	37	KARYE	SUNUGER

and Div.id). Values of Div.id are generated on the first occasion a settlement is identified in a defter. In subsequent defters, the same value is used if the same settlement can be identified. New values of Div.id are generated only when new settlements are identified. The column upper division (Up.div) contains the values of the identifiers of the spatial divisions to which each division is connected. If a change occurs in the spatial organisation of the area, the values of Up.div for this settlement are adjusted accordingly. These two columns describe a recursive relationship. In this particular extract this relationship is shown at three points of time.

The time information is stored in the Doc.date columns. Changes are not explicitly coded into the database, but are discovered by using the schange (spatial change) operation within HiSQL. We record the explicit states of the hierarchies only at the dates for which data are available (i.e. the dates of the defters).

At the first point, the year 1485, the spatial division 53 (the *karye* or village of Sunuger) is connected to the spatial division 2341 (the *nahiye* or region of Serin) which in turn is connected to division 2 (the *sancak* or province of Karahisar). In more complex cases, there may be many such levels.

At the second time point (the year 1547) there are some changes in these relationships. The spatial division 2341 (Serin) no longer exists and a new division is created (the *nahiye* of Siryan) to which the identifier 37 has been assigned. At this date, the village of Sunuger is connected to this newly created division. Siryan is connected to the province of Karahisar and so Sunugar remains in this province. At the third time point, the year 1569, the situation is unchanged from that in 1547.

Since every social event occurs at some point in time and in some place, other events (or objects) stored in the database must be related to the spatial divisions. For example, the taxes collected from settlements are stored in the table “tax”, which contains columns for “Div.id”, “Doc.No” and “D.date” (containing the same information as for the table “areal_division”) as well as columns for information

Table II. Taxes of one village in two dates

Doc.No	D.date	Div.id	Tax_id	Tax_name	Amount
TD255	1547	53	1746	AGNAM	7
TD255	1547	53	1766	ARUSIYE	7
TD255	1547	53	1758	HINTA	1500
TD255	1547	53	1759	SAIR	800
TD255	1547	53	1762	KOVAN	50
TD478	1569	53	1746	AGNAM	1500
TD478	1569	53	1766	ARUSIYE	300
TD478	1569	53	1758	HINTA	3300
TD478	1569	53	1759	SAIR	1500
TD478	1569	53	1762	KOVAN	60
TD478	1569	53	1729	ISPENCE	1275
TD478	1569	53	4004	KETAN	150
TD478	1569	53	4006	NIM	227
TD478	1569	53	597	BENNAK	72
TD478	1569	53	611	CABA	91

about the taxes. Table 2 lists the taxes for the village of Sunuger in 1547 and 1569.

It may be observed further that these taxes continue to be collected at the same rate until the next survey in 1569. In the register compiled at that date there are some new taxes added to the list as well as increases in the amounts of taxes previously collected. See (Acun, 1994) for the complete conceptual model for the registers as well as methodological reasons which guided the modelling decisions.

4. Task-Oriented Retrieval Using HiSQL

This section presents a brief description of the contexts in which the operations can be used together with queries performed on the sample data presented in Appendix A.

4.1 "STREE" OR SPATIAL TREE

The operation **stree** is designed primarily to help in cases where the administrative structure of an area is of interest. This operation may also be used to view the hierarchy of spatial divisions at a single point in time. This is achieved by specifying identical values for the start date (tbegin) and the end date (tend). Other operations described below can also be used in conjunction with **stree**. The following examples show the use of this operation when applied to the sample data.

Query 1. Create the hierarchy of the spatial divisions for the SANCAK of KARAHISAR in the year 1485.

In HiSQL, this becomes:

/stree[division_name = KARAHISAR, division_type = SANCAK, tbegin = 1485, tend = 1485]

This results in the output:

```
SANCAK    KARAHISAR
  NAHIYE    GEZENGER
    KARYE    CENGERIS
    KARYE    GERSUN MIYANE
  NAHIYE    GUDUL
    KARYE    ASSAGI KINIK
    KARYE    SARAYCIK
    KARYE    TARABUL
  NAHIYE    MENKUFE
    KARYE    TARU BUKU
    KARYE    CANDIL
    KARYE    CANDIR
    KARYE    BES KILISA
    KARYE    CANAKCI
    KARYE    KUM
  NAHIYE    GAVEZIT
    KARYE    KARA AGAC
    KARYE    KARA SEHINSAH
    KARYE    YUKARI KINIK
    KARYE    ASSAGI GUVERCINLIK
    KARYE    YUKARI GUVERCINLIK
    KARYE    YUMURCAK TAS
  NAHIYE    SERIN
    KARYE    GERSUN BALA
      MEZRAA    GERSUN MIYANE
    KARYE    RIBAT
      MEZRAA    SARUCLU
    KARYE    SEYDI BABA
    KARYE    SUDEREK
    KARYE    SUNUGER
```

Query 1: can be expressed in SQL by the following code.

```
1 SELECT DIVISION_TYPE, DIVISION_NAME
2 FROM test_area
3 WHERE DDATE = 1485
4 CONNECT BY PRIOR DIVISION_ID = UPPER_DIVISION
```

```
5 AND DDATE = 1485
6 START WITH DIVISION_TYPE = 'SANCAK' AND DIVISION_NAME =
  'KARAHISAR'
7 AND DDATE = 1485
```

This is not an easy query for a naive user to write because it requires knowledge of the structure of the table and the relationships within this structure. In particular, the user needs to know about the recursive relationship between columns Div.id and Up.div and the way in which SQL creates the required hierarchy from this relationship. The corresponding HiSQL constructs require none of these; the user only needs to know the name and type of the division for which he/she wishes to construct the hierarchy and the date of it.

There is one way in which this SQL query could be simplified. That is to remove the recursive relationship from the table by adding a new column in the table for each level of the hierarchy. In this case, it would require the addition of two new columns called province and county which would store the names of the relevant province and county for each record. This violates the normalisation rules by introducing unwanted redundancy. Note that in some applications there may be many such levels. In the defters, for example, it is possible to come across the organisation of spatial divisions up to six levels.

Query 2. Create the hierarchy of the spatial divisions for the NAHIYE of GAVEZIT in the year 1485. This example shows that a subset of the hierarchy can be selected.

```
/stree[division_name = GAVEZIT,division_type = NAHIYE, tbegin = 1485,
tend = 1485]
```

```
NAHIYE   GAVEZIT
  KARYE   KARA AGAC
  KARYE   KARA SEHINSAH
  KARYE   YUKARI KINIK
  KARYE   ASSAGI GUVERCINLIK
  KARYE   YUKARI GUVERCINLIK
  KARYE   YUMURCAK TAS
```

Query 3. List the taxes collected from the NAHIYE of GAVEZIT in the year 1485. This illustrates the case when the objective is to study other aspects of the area using its current administrative structure. Here the operation is embedded in SQL statements to retrieve the relevant data.

```
select * from test_tax where stree[division_name = GAVEZIT, division_type =
NAHIYE, tbegin = 1485, tend = 1485] and ddate = 1485
```

Doc.No	D.date	Div.id	Tax_id	Tax_name	Amount
TD37	1485	400	596	CIFT	42
TD37	1485	400	597	BENNAK	13
TD37	1485	400	616	CIFT VE BENNAK	20
TD37	1485	400	612	BAD-I HAVA	5
TD37	1485	400	598	KENDUM	240
TD37	1485	400	599	CEV	300
TD37	1485	407	611	CABA	8
TD37	1485	407	623	BENNAK	5
TD37	1485	407	612	BAD-I HAVA	3
TD37	1485	441	598	KENDUM	100
TD37	1485	441	599	CEV	250

4.2 "SCONTINUITY" OR SPATIAL CONTINUITY

In many cases, researchers need to be able to identify administratively stable areas in order to make noise free observations over the period of interest. In studies in historical demography, for example, when measuring the population growth in an area, it is considered as methodologically sound to select only those settlements which have existed throughout the period concerned. The settlements which do not conform to this criterion, are completely ignored. This is also the case in studies involving economic growth (see Cook, 1972; Erder and Faroqhi, 1979).

Query 4. Show the administrative structure of NAHIYE of GAVEZIT which remained stable between 1485 and 1569 (ignore those divisions appearing/ disappearing between these dates).

/scontinuity[division_type = NAHIYE, division_name = GAVEZIT, tbegin = 1485, tend = 1569]

NAHIYE GAVEZIT
  KARYE YUKARI KINIK
  KARYE ASSAGI GUVERCINLIK
KARYE YUMURCAK TAS

Query 5. Get the total value of taxes levied on the grain production for each point between 1485 and 1569 for the those areas obtained by Query 4. This requires a mixture of SQL and HiSQL.

select ddate, sum(amount) from test_tax where name in ('CEV', 'SAIR', 'HINTA') and scontinuity[division_type = NAHIYE, division_name = GAVEZIT, tbegin = 1485, tend = 1569] group by ddate

D.date	SUM(Amount)
1485	300
1547	4000
1569	7800

4.3 "SCHANGE" OR SPATIAL CHANGE

In a complete contrast to the previous case, administratively volatile areas such as border regions (which frequently change hands or become abandoned as a result of wars) or areas subject to administrative restructuring are sometimes of particular interest to researchers. A research hypothesis may predict a different result for such an area than that obtained for otherwise similar but administratively stable areas.

For example, in the first case the researcher may be expecting to find, say, a much smaller increase in the grain production. In the latter case, the study of aspects of such an area may reveal the underlying reason why that change took place in the first place (Hutteroth, 1985; Kaldy-Nagy, 1963).

The operation, **schange** was developed to assist in these and other similar situations where it is necessary to be able to identify and refer to such an area. The following query illustrates this.

Query 6. What change occurred in the administrative structure of NAHIYE of GAVEZIT between 1485 and 1569.

/schange[division_type = NAHIYE, division_name = GAVEZIT, tbegin = 1485, tend = 1569]

421	KARYE	OVACIK
441	KARYE	KARA AGAC
4202	KARYE	KARA SEHINSAH
408	KARYE	YUKARI GUVERCINLIK
423	MEZRAA	DELICE
387	KARYE	SARAYCIK
424	KARYE	AHURCIK
433	KARYE	TARABUL

This query has returned the list of the villages which were not connected to Nahiye of Gavezit continuously (i.e. from start to end) during the period 1485–1569. Obviously, these areas can also be used, as in the previous examples, as a basis for further queries.

The solution for Query 6 in SQL in a single step would be the following code.

```
 1 SELECT DISTINCT DIVISION_ID, DIVISION_TYPE, DIVISION_NAME
 2 FROM test_area
 3 where UPPER_DIVISION IN
 4    (SELECT DIVISION_ID FROM TEST_AREA
 5    CONNECT BY PRIOR DIVISION_ID = UPPER_DIVISION
 6    START WITH DIVISION_TYPE = 'NAHIYE' AND DIVISION_NAME =
      'GAVEZIT')
 7 AND DIVISION_ID NOT IN
 8    (SELECT DIVISION_ID
 9    FROM test_area WHERE DDATE = 1485
10       CONNECT BY PRIOR DIVISION_ID = UPPER_DIVISION
```

```
11      AND DDATE = 1485
12      START WITH DIVISION_TYPE = 'NAHIYE' AND DIVISION_NAME =
        'GAVEZIT'
13      AND DDATE = 1485
14        intersect
15        SELECT DIVISION_ID
16        FROM test_area WHERE DDATE = 1547
17        CONNECT BY PRIOR DIVISION_ID = UPPER_DIVISION
18        AND DDATE = 1547
19        START WITH DIVISION_TYPE = 'NAHIYE' AND DIVISION_NAME =
          'GAVEZIT'
20        AND DDATE = 1547
21        intersect
22        SELECT DIVISION_ID
23        FROM test_area WHERE DDATE = 1569
24        CONNECT BY PRIOR DIVISION_ID = UPPER_DIVISION
25        AND DDATE = 1569
26        START WITH DIVISION_TYPE = 'NAHIYE' AND DIVISION_NAME =
          'GAVEZIT'
27*       AND DDATE = 1569)
```

As is seen, this piece of code is considerably larger than **schange** command of HiSQL. In fact this code will get even larger if more time points are considered; an additional 7 lines (the block of the code between and including one of the *intersect* operators) need to be written for each such point. So for example, if this query is amended to consider a period which includes, say, six survey points, which is quite common in studies dealing with serial documents, this piece would grow to an amazing 48 lines whereas **schange** command will remain as a single line.

This query requires even more sophisticated knowledge of SQL. The other important point is the execution time. Compared with the almost instantaneous execution of **schange**, the SQL code takes several minutes to produce the result, even for this small size sample.

4.4 "VSPACE" OR VIRTUAL SPACE

It was mentioned on several occasions that the representation of space changes over time. The same geographical area may have been represented under different names and may have become part of different administrative/political structures at different times. Data about this same geographical area may have been organised under each of the new representations. In many studies, it is necessary to relate these different representations of the same geographical area to each other in order to be able to do temporal comparisons or to do searches over a certain time period. The **vspace** operation is illustrated in the next query.

Query 7. Get the name and value of those taxes collected from the NAHIYE of SERIN in 1569 which were not being collected from the NAHIYE of GUDUL at any point. Consider the period from 1485 to 1569 inclusive for the area searches.

select name, amount from test_tax where ddate = 1569 and vspace[division_type = NAHIYE, division_name = SERIN, tbegin = 1485, tend = 1569] and tax_id not in (select tax_id from test_tax where vspace[division_type = NAHIYE, division_name = GUDUL, tbegin = 1485, tend = 1569])

(The SQL code for this query occupies a full 48 lines.)

Tax_name	Amount
KETAN	150
NIM	227
ISPENCE	1275

In this query the user only knows that there must be nahiyes called SERIN and GUDUL in the period from 1485 to 1569 but he/she neither knows, nor wishes to know, the exact extent of these. Despite the fact that there was no nahiye called SERIN in 1569, he/she was able to refer to it through the **vspace** construct and retrieve the appropriate data (compare this problem with the one described in (Thaller, 1989, p. 4)).

To a defterologist, the data retrieved by this query would reveal that as far as the sample data can tell, a nahiye in the Eastern part of the SANCAK of KARAHISAR was radically different from a nahiye in its central part in terms of its crops (as indicated by the first tax), availability of a relatively larger amount of land for cultivation by its peasants (as indicated by the second tax), and the religious composition of its population (as indicated by the third tax).

5. A Comparison of the HiSQL with SQL

SQL code for some of the HiSQL queries has been included to provide a basis for the evaluation of HiSQL. Two points are emphasized:

1) the ease with which the queries can be formulated (e.g. the number of lines of code required) and
2) the amount of time required.

The work presented here shows that HiSQL improves on the SQL's capability to handle historical data. Because of the limited space we cannot go on to compare HiSQL with other extensions of SQL. For this see (Acun, 1994).

HiSQL has been compared with Oracle's implementation of SQL, which is a superset of the standard SQL and has facilities for querying hierarchical relationships. The standard SQL (Date and John, 1993) has no such capability. From these examples, one might conclude that SQL is not at all suitable for dealing with historical data. However, the issue of time and space is also a problem for other paradigms such as object-oriented and deductive database systems and these two dimensions are subject to extensive research in those areas as well, see for example

(Rose and Segev, 1991; Wuu and Dayal, 1993; Kafer et al., 1990; Baudinet et al., 1993).

In the database modelling, temporal and spatial dimensions of the real world have usually been treated separately. There is a substantial amount of work on both aspects separately, but there are relatively few works which attempt to unify these two aspects in a seamless model. Only recently have the issues concerning the simultaneous support of both space and time in databases been considered. The majority of the works concerning the spatial aspect deals with the physical implementation of spatial databases such as developing access methods for spatial data. By comparison abstract modelling and querying of spatial data have received less attention (Gadia and Nair, 1993).

6. Generalisability to Other Sources

HiSQL is a general system which should be applicable to any other database which includes the same hierarchy of spatial divisions at one or more points in time. This implementation has been tested with a small sample of data from the English census records of the nineteenth century. These imply spatial division into counties, hundreds, parishes and also list hamlets and individual farmsteads within the parish.

The English census data contains no information about taxes and so this cannot be extracted from the database. Instead these records can be queried to give information on population, family size, occupations and movement of population (by comparing birthplace with present place of residence). A small sample of this data has been entered into HiSQL and queries have shown that the schange and scontinuity operations make the queries much easier to use. A larger sample would be needed to make real use of the other operations, but we have checked enough queries to establish that this works correctly. Other census or similar data could also be entered into our system and the same advantages would become apparent.

7. Conclusion and Future Research

This paper has described HiSQL, the front-end query system for historical relational databases and has compared it with the standard SQL. The comparison has shown clearly that HiSQL is considerably easier to use for the manipulation of historical data. This is because HiSQL incorporates some knowledge of the task of the historian in the form of its operations on data. Since the system has a modular design, it can easily be extended by adding new operations formulated along the lines of the operations described here.

There are, however some limitations of the system. Firstly the current implementation only works for the Oracle RDBMS. Secondly it needs a large amount of memory in order to store the data about the spatial divisions (the larger the data, the larger the amount of memory required). Thirdly reading in the spatial data and

building the graph structure in the memory is a slow process, although once this has been set up, the later queries using this structure are very fast.

One obvious improvement would be the implementation of this system to run with Access on a PC system. This is high on our list of priorities and we hope to report on progress in due course.

References

Acun, R. *Modelling and Retrieval of Historical Data*. PhD thesis, The University of Birmingham, 1994.

Acun, F. *Ottoman Administration in the Sancak of Karahisar-i Sarki (1485–1569): An Analysis Based on the Tahrir Defters*. PhD thesis, The University of Birmingham, 1993.

Acun, R., R. Anane and S. Laflin. "Database Design for Ottoman Tax Registers". In *Yesterday, Proceedings of the 6th AHC Conference at Odense*, 1994a, pp. 109–122.

Acun, R., R. Anane and S. Laflin. "User Interface Design for Historical Databases". In *Storia & Multimedia, Proceedings of 7th AHC Conference at Bologna*, 1994b, pp. 534–547.

Baudinet, M., J. Chomicki and P. Wolper. *Temporal Deductive Databases*, in (Tansel et al., 1993), 1993, pp. 294–320.

Cook, M.A. *Population Pressure in Rural Anatolia 1400–1600*. London, 1972.

Date, C.J. and C. John. *A Guide to the SQL Standard, a User's Guide to the Standard Relational Language SQL*. Addison-Wesley, 1993.

Edwards, J.L. and J.A. Mason. "Toward Intelligent Dialogue with ISIS". *Int. J. Man-Machine Studies*, 28 (1988), 309–342.

Erder, L. and S. Faroqhi. "Population Rise and Fall in Anatolia 1550–1620". *Middle East Studies*, 15 (1979), 328–345.

Gadia, S.K. and S.S. Nair. *Temporal Databases: A Prelude to Parametric Data*, in (Tansel et al. 1993), 1993, pp. 28–66.

Hutteroth, W. "Ottoman Administration of Desert Frontier in the Sixteenth Century". *Asian and African Studies*, 2 (1985), 145–155.

Kafer, W., N. Ritter and H. Schoning. "Support for Temporal Data by Complex Objects". *Proc. of the 16th Int. Conf. on Very Large Data Bases*. Brisbane, Australia, 1990.

Kaldy-Nagy, G. "The Administration of Sancak Registration in Hungary". *Acta Orientalia*, 13 (1963), 181–223.

Oz, M. *Population, Taxation and Regional Economy in the District of Canik (According to Ottoman Tahrir Defters, 1455–1576)*. Ph.D. thesis, Cambridge, 1990.

Ozel, O. *Changes in Settlement Patterns, Population and Society in Rural Anatolia: A Case Study of Amasya, 1576–1642*. Ph.D thesis, Manchester University, 1993.

Rose, E. and A. Segev. "TOODM-A Temporal Object Oriented Data Model with Temporal Constraints". *Proceedings of the 10th International Conference on the Entity-Relationship Approach*. 1991.

Tansel, A.U., J. Clifford, S.Gadia, S. Jajodia, A. Segev, and R. Snodgrass. *Temporal Databases, Theory, Design and Implementation*. Massachusetts, 1993.

Thaller, M. *The Need for a Theory of Historical Computing*, in *History & Computing II* Denley et al. 1989, pp. 2–11.

Wuu, G.T.J. and U. Dayal. *A Uniform Model for Temporal and Versioned Object-oriented Databases* in (Tansel et al. 1993). 1993, pp. 230–247.

Appendix: Sample Data

TEST_AREAL_DIVISION

DOC.NO	DDATE	DIV.ID	UPPER_DIV.	DIV.TYPE	DIV.NAME
TD37	1485	2	1	SANCAK	KARAHISAR
TD37	1485	38	2341	KARYE	RIBAT
TD37	1485	44	2360	KARYE	CANAKCI
TD37	1485	46	2360	KARYE	CANDIL
TD37	1485	47	2279	KARYE	CENGERIS
TD37	1485	53	2341	KARYE	SUNUGER
TD37	1485	71	2360	KARYE	CANDIR
TD37	1485	72	2360	KARYE	TARU BUKU
TD37	1485	85	2341	KARYE	GERSUN BALA
TD37	1485	86	2279	KARYE	GERSUN MIYANE
TD37	1485	101	2360	KARYE	BES KILISA
TD37	1485	109	2360	KARYE	KUM
TD37	1485	162	2341	KARYE	SEYDI BABA
TD37	1485	365	2	NAHIYE	GAVEZIT
TD37	1485	387	2232	KARYE	SARAYCIK
TD37	1485	400	365	KARYE	YUKARI KINIK
TD37	1485	407	365	KARYE	ASSAGI GUVERCINLIK
TD37	1485	408	365	KARYE	YUKARI GUVERCINLIK
TD37	1485	426	365	KARYE	YUMURCAK TAS
TD37	1485	433	2232	KARYE	TARABUL
TD37	1485	441	365	KARYE	KARA AGAC
TD37	1485	2232	2	NAHIYE	GUDUL
TD37	1485	2237	2232	KARYE	ASSAGI KINIK
TD37	1485	2279	2	NAHIYE	GEZENGER
TD37	1485	2341	2	NAHIYE	SERIN
TD37	1485	2348	2341	MEZRAA	SARUCLU
TD37	1485	2351	2341	MEZRAA	GERSUN MIYANE
TD37	1485	2360	2	NAHIYE	MENKUFE
TD37	1485	4202	365	KARYE	KARA SEHINSAH
TD37	1485	4584	2341	KARYE	SUDEREK
TD255	1547	2	1	SANCAK	KARAHISAR
TD255	1547	35	43	MEZRAA	CANAKCI
TD255	1547	37	2	NAHIYE	SIRYAN
TD255	1547	53	37	KARYE	SUNUGER
TD255	1547	43	37	KARYE	CERMIS
TD255	1547	46	37	KARYE	CANDIL
TD255	1547	47	37	KARYE	CENGERIS
TD255	1547	71	37	KARYE	CANDIR
TD255	1547	72	37	KARYE	TARU BUKU
TD255	1547	85	37	KARYE	GERSUN BALA
TD255	1547	86	37	KARYE	GERSUN MIYANE

TD255	1547	90	37	KARYE	SUSUZ
TD255	1547	102	37	KARYE	MIYADIN
TD255	1547	109	37	KARYE	KUM
TD255	1547	132	37	KARYE	ILIKCI
TD255	1547	135	37	KARYE	IN OZI
TD255	1547	154	37	KARYE	HURMAN
TD255	1547	155	37	MEZRAA	HURMAN BALA
TD255	1547	365	2	NAHIYE	GAVEZIT
TD255	1547	387	365	KARYE	SARAYCIK
TD255	1547	400	365	KARYE	YUKARI KINIK
TD255	1547	407	365	KARYE	ASSAGI GUVERCINLIK
TD255	1547	421	365	KARYE	OVACIK
TD255	1547	423	365	MEZRAA	DELICE
TD255	1547	424	365	KARYE	AHURCIK
TD255	1547	426	365	KARYE	YUMURCAK TAS
TD255	1547	433	365	KARYE	TARABUL
TD478	1569	2	1	SANCAK	KARAHISAR
TD478	1569	37	2	NAHIYE	SIRYAN
TD478	1569	53	37	KARYE	SUNUGER
TD478	1569	43	37	KARYE	CERMIS
TD478	1569	46	37	KARYE	CANDIL
TD478	1569	47	37	KARYE	CENGERIS
TD478	1569	69	43	MEZRAA	CANAKCI
TD478	1569	71	37	KARYE	CANDIR
TD478	1569	72	37	KARYE	TARU BUKU
TD478	1569	85	37	KARYE	GERSUN BALA
TD478	1569	86	37	KARYE	GERSUN MIYANE
TD478	1569	90	37	KARYE	SUSUZ
TD478	1569	102	37	KARYE	MIYADIN
TD478	1569	109	37	KARYE	KUM
TD478	1569	132	37	KARYE	ILIKCI
TD478	1569	135	37	KARYE	IN OZI
TD478	1569	154	37	KARYE	HURMAN
TD478	1569	155	37	KARYE	HURMAN BALA
TD478	1569	365	2	NAHIYE	GAVEZIT
TD478	1569	387	365	KARYE	SARAYCIK
TD478	1569	400	365	KARYE	YUKARI KINIK
TD478	1569	407	365	KARYE	ASSAGI GUVERCINLIK
TD478	1569	421	365	KARYE	OVACIK
TD478	1569	423	365	KARYE	DELICE
TD478	1569	424	365	KARYE	AHURCIK
TD478	1569	426	365	KARYE	YUMURCAK TAS
TD478	1569	433	365	KARYE	TARABUL

TEXT_TAX

DOC.NO	DDATE	DIV._ID	TAX_ID	NAME	AMOUNT
TD37	1485	400	596	CIFT	42
TD37	1485	400	597	BENNAK	13
TD37	1485	400	616	CIFT VE BENNAK	20
TD37	1485	400	612	BAD-I HAVA	5
TD37	1485	400	598	KENDUM	240
TD37	1485	400	599	CEV	300
TD37	1485	407	611	CABA	8
TD37	1485	407	623	BENNAK	5
TD37	1485	407	612	BAD-I HAVA	3
TD37	1485	441	598	KENDUM	100
TD37	1485	441	599	CEV	250
TD37	1485	2263	612	BAD-I HAVA	2
TD37	1485	2264	598	KENDUM	800
TD37	1485	2264	599	CEV	180
TD255	1547	387	1757	RUSUM	96
TD255	1547	387	1758	HINTA	1000
TD255	1547	387	1764	DESTBANI	10
TD255	1547	387	1759	SAIR	800
TD255	1547	387	1766	BAD-I HAVA VE ARUSIYE	84
TD255	1547	387	1746	AGNAM	40
TD255	1547	400	1757	RUSUM	60
TD255	1547	400	1758	HINTA	1150
TD255	1547	400	1762	KOVAN	60
TD255	1547	400	1759	SAIR	1000
TD255	1547	400	1746	AGNAM	400
TD255	1547	400	1764	DESTBANI	60
TD255	1547	400	1743	YATAK	33
TD255	1547	400	1744	DUD	30
TD255	1547	400	1766	BAD-I HAVA VE ARUSIYE	200
TD255	1547	407	1746	AGNAM	152
TD255	1547	407	1766	BAD-I HAVA VE ARUSIYE	200
TD255	1547	407	1758	HINTA	750
TD255	1547	407	1757	RUSUM	162
TD255	1547	407	1743	YATAK	53
TD255	1547	407	1744	DUD	30
TD255	1547	407	1762	KOVAN	10
TD255	1547	407	1759	SAIR	600
TD255	1547	407	1764	DESTBANI	5
TD255	1547	421	1725	BOSTAN	20
TD255	1547	421	1744	DUD	10
TD255	1547	421	1743	YATAK	30
TD255	1547	421	1758	HINTA	500
TD255	1547	421	1762	KOVAN	30
TD255	1547	421	1759	SAIR	400

TD255	1547	421	1757	RUSUM	78
TD255	1547	421	1766	BAD-I HAVA VE ARUSIYE	70
TD255	1547	421	1746	ADET-I AGNAM	30
TD255	1547	421	1764	DESTBANI	30
TD255	1547	424	1757	RUSUM	56
TD255	1547	424	1766	BAD-I HAVA VE ARUSIYE	30
TD255	1547	424	1746	ADET-I AGNAM	30
TD255	1547	424	1764	DESTBANI	14
TD255	1547	424	1735	CIFTLIK-I HACI AHMED	200
TD255	1547	424	1725	BOSTAN	20
TD255	1547	424	1759	SAIR	120
TD255	1547	424	1758	HINTA	400
TD255	1547	426	1759	SAIR	200
TD255	1547	426	1758	HINTA	300
TD255	1547	426	1757	RUSUM	104
TD255	1547	426	1743	YATAK	23
TD255	1547	426	1762	KOVAN	23
TD255	1547	426	1746	ADET-I AGNAM	80
TD255	1547	426	1764	DESTBANI	20
TD255	1547	426	1744	DUD	20
TD255	1547	426	1766	BAD-I HAVA VE ARUSIYE	80
TD255	1547	433	1757	RUSUM	143
TD255	1547	433	1758	HINTA	500
TD255	1547	433	1759	SAIR	400
TD255	1547	433	1762	KOVAN	30
TD255	1547	433	1761	BAGAT VE BESATIN	120
TD255	1547	433	1760	. . .	20
TD255	1547	433	1766	BAD-I HAVA VE ARUSIYE	70
TD255	1547	433	1765	AGNAM	20
TD255	1547	433	1764	DESTBANI	36
TD255	1547	433	1763	PIYAZ	20
TD478	1569	387	596	CIFT	57
TD478	1569	387	1759	SAIR	1000
TD478	1569	387	1758	HINTA	1200
TD478	1569	387	1766	BAD-I HAVA VE ARUSIYE	84
TD478	1569	387	597	BENNAK	18
TD478	1569	387	1762	KOVAN	10
TD478	1569	387	4584	DONUM	50
TD478	1569	387	611	CABA	117
TD478	1569	387	1764	DESTBANI	100
TD478	1569	400	597	BENNAK	180
TD478	1569	400	1764	DESTBANI	273
TD478	1569	400	1766	BAD-I HAVA VE ARUSIYE	300
TD478	1569	400	611	CABA	117
TD478	1569	400	1743	YATAK	30
TD478	1569	400	1744	DUD	30

TD478	1569	400	4584	DONUM	20
TD478	1569	400	1762	KOVAN	150
TD478	1569	400	1759	SAIR	1500
TD478	1569	400	1758	HINTA	1800
TD478	1569	407	1725	BOSTAN	20
TD478	1569	407	1758	HINTA	1800
TD478	1569	407	1759	SAIR	450
TD478	1569	407	597	BENNAK	54
TD478	1569	407	1766	BAD-I HAVA VE ARUSIYE	1000
TD478	1569	407	1762	KOVAN	20
TD478	1569	407	1743	YATAK	50
TD478	1569	407	1729	ISPENCE	375
TD478	1569	407	1764	DESTBANI	150
TD478	1569	407	1674	DUHAN	30
TD478	1569	421	1725	BOSTAN	20
TD478	1569	421	1744	DUD	10
TD478	1569	421	597	BENNAK	188
TD478	1569	421	1743	YATAK	20
TD478	1569	421	1764	DESTBANI	100
TD478	1569	421	1766	BAD-I HAVA VE ARUSIYE	120
TD478	1569	421	1729	ISPENCE	325
TD478	1569	421	1758	HINTA	900
TD478	1569	421	1762	KOVAN	30
TD478	1569	421	1759	SAIR	500
TD478	1569	423	597	BENNAK	57
TD478	1569	423	1764	DESTBANI	23
TD478	1569	423	1766	BAD-I HAVA VE ARUSIYE	30
TD478	1569	423	1759	SAIR	250
TD478	1569	423	1758	HINTA	180
TD478	1569	423	4584	DONUM	50
TD478	1569	423	1762	KOVAN	7
TD478	1569	423	1725	BOSTAN	6
TD478	1569	424	597	BENNAK	90
TD478	1569	424	1764	DESTBANI	44
TD478	1569	424	1766	BAD-I HAVA VE ARUSIYE	30
TD478	1569	424	611	CABA	26
TD478	1569	424	1735	CIFTLIK-I HACI AHMED	200
TD478	1569	424	4584	DONUM	50
TD478	1569	424	1762	KOVAN	20
TD478	1569	424	1758	HINTA	480
TD478	1569	424	1725	BOSTAN	20
TD478	1569	424	1759	SAIR	150
TD478	1569	426	1759	SAIR	750
TD478	1569	426	1766	BAD-I HAVA VE ARUSIYE	172
TD478	1569	426	611	CABA	52
TD478	1569	426	1762	KOVAN	50
TD478	1569	426	1743	YATAK	38

TD478	1569	426	1764	DESTBANI	220
TD478	1569	426	597	BENNAK	198
TD478	1569	426	1758	HINTA	1500
TD478	1569	426	1744	DUD	20
TD478	1569	433	596	CIFT	57
TD478	1569	433	1763	PIYAZ	40
TD478	1569	433	1764	DESTBANI	200
TD478	1569	433	597	BENNAK	216
TD478	1569	433	1766	BAD-I HAVA VE ARUSIYE	269
TD478	1569	433	611	CABA	13
TD478	1569	433	1762	KOVAN	50
TD478	1569	433	1760	. . .	50
TD478	1569	433	1761	BAGAT VE BESATIN	400
TD478	1569	433	1759	SAIR	1750
TD478	1569	433	1758	HINTA	2400
TD255	1547	53	1746	AGNAM	7
TD255	1547	53	1766	BAD-I HAVA VE ARUSIYE	7
TD255	1547	53	1758	HINTA	1500
TD255	1547	53	1759	SAIR	800
TD255	1547	53	1762	KOVAN	50
TD478	1569	53	1746	AGNAM	1500
TD478	1569	53	1766	BAD-I HAVA VE ARUSIYE	300
TD478	1569	53	1758	HINTA	3300
TD478	1569	53	1759	SAIR	1500
TD478	1569	53	1762	KOVAN	60
TD478	1569	53	1729	ISPENCE	1275
TD478	1569	53	4004	KETAN	150
TD478	1569	53	4006	NIM	227
TD478	1569	53	597	BENNAK	72
TD478	1569	53	611	CABA	91