Object-relational Features for modeling and analysis of Spatio-temporal Data

A.Z.M Kamruzzaman, Mohd Taib Wahid
Department of Information Systems, Faculty of Computer Science & Information Systems
University Technology of Malaysia, K.B. 791
81310 Skudai, Johor, Malaysia
Tel: (607)-5532422, Fax: (607) 5565044
akzaman@siswa.utm.my, taib@fsksm.utm.my

Abstract

The age of discovery, description, and classification of biodiversity is entering a new phase. In responding to the importance and uses of the large number of biodiversity data of Malaysia make it essential for their biodiversity to be conserved. In order to manage large number of biodiversity data, it is essential to develop a data model which is a new integrated and support spatio-temporal data. Recently, we propose a conceptual spatio-temporal object-relational biodiversity data model and which is being tested with Malaysian biodiversity data and has been loaded to oracle database. The main objectives of this paper are new aspects of using object-relational feature of the underlying database management system Oracle to improve the extensibility and flexibility of data models, enhance interoperability and analyzing potential as well as to ensure consistency by defining standards based on abstract data types. The object-relational extension of data models can be combined with data pyramid framework techniques to analyze present, past and future data. In addition to this, by using this model forest department, researchers, as well as natural recourses will be helpful.

Keywords: Biodiversity, Spatio-temporal, Object-relational, Oracle 9i

1. Introduction

Biodiversity, or biological diversity, refers to the variety of life on earth and the most important factors Influencing the stability and health of ecosystems. Information about biodiversity and ecology is vital to a wide range of scientific, educational, industrial, governmental activities. Malaysia is undoubtedly on the richest rain forests in the world with diverse floristic composition and complex ecosystem. It is not only rich for plant biodiversity data also rich in fauna and peat swamp. Malaysia has been loosing much of its natural resources including plants and animal species through ecosystem and degeneration.

Malaysia has been endowed with vast amount of natural resources including luxuriant tropical forest which is one of the most diverse and complex ecosystems of the world. Malaysia is rated as one of the world's 12 "mega diversity" countries. It boasts over 150,000 species of invertebrates, 286 mammal species, 736 bird species and 15,000 types of flowering plant. Forest resources have been one of the major sources of revenue in the Malaysian economy. However, it is decreasing every year. The change of land covered by forests has been reduced from 65.9% in 1990 to 59.5% in 2002 [1].

Different types of forests can be found in the Peninsular, Sabah and Sarawak. From the management perspective, forests of Peninsular Malaysia can be classified into dipterocarp, peat swamp and mangrove forests. Among which approximately 95% is covered by dipterocarp forests, 3.34% by peat swamps and 1.84% by mangrove forests (UNEP, 2002)**.** In Sabah, there is a gradual succession of forest vegetation from the coastal beach forests and mangrove forests to lowland dipterocarp forest and eventually montane forests. In Sarawak, five types of natural forest types are abundant namely: Hill Mixed Dipterocarp Forest, Peat Swamp Forest, Mangrove Forest, *Kerangas* Forest (Heat Forest) and Montane Forest (UNEP 2002). Currently there are total 610,606 hectares of mangrove forest estimated in Malaysia. The largest concentration are found in Sabah (56%) followed by Sarawak (27.5%) and the remaining in Peninsular Malaysia (16.5%) [2]. However, this estimation is still on process and areas in Sabah and Sarawak have not been actively surveyed botanically. Based conservation, management, data analysis, data accessing, monitoring and data complexity of biodiversity data, data model is required to protect biodiversity data.

After almost two decades research, representation of space and time in databases and functional applications are still problematic [8]. Problem with previous data models is that spatial and temporal aspects of databases are modeled separately [9], [10]. Spatial database focuses on supporting geometries [11], while temporal databases focus on the past state [12].

Another problem is representation of data should be natural to human. The structures of space and time are identified as essential for the realization of cognitive systems [13]. According to Donna J. Peuquet and her group [14], models of spatial temporal data in geographical database representations must incorporate human cognitive principles. Human knowledge of the dynamic geographical world comprises of three different (and interrelated) subsystems that handle *what*, *where* and *when* aspects of object properties [15].

Recent years have seen an explosion of digitally available information about biological diversity [3]. Data management tools and software need more sophisticated facilities to face new requirements from emerging application areas and non-traditional user interactions. In particular, better concepts and tools for manipulating spatio-temporal data are needed. Major DBMS tools are incorporating facilities for spatial or temporal data management (e.g., Oracle's Spatial Cartridges and Informix's Databases). Temporal systems are still somehow behind, with no generic products on the marketplace, just a few ad hoc systems or application-specific developments (e.g., for time series management). However, current tools do not match the user perception of and reasoning about the application data

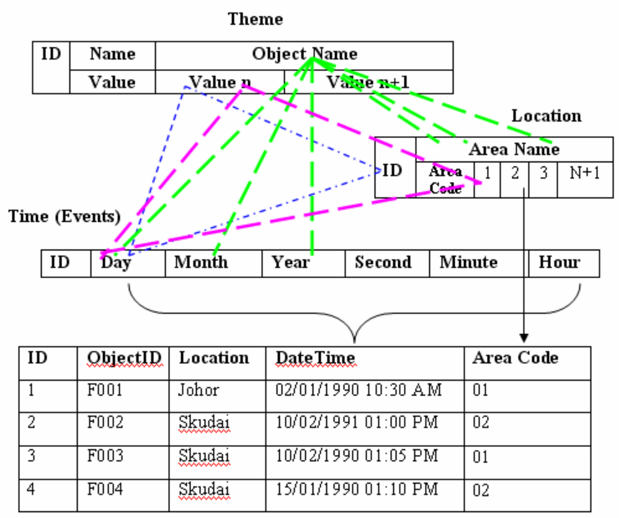
So far no data model for analysis of biodiversity data has been performed using GIS and database technology, particularly in regard to new possibilities provided through object-relational features. Since database models get closer to real world objects and events by allowing the setup of an integrated and adjusted representation based on user-defined types. A complete description of spatial objects in database systems can be provided for the first time. Therefore it is obvious to make use of object-relational structures provided by the underlying DBMS to develop a new approach to develop biodiversity data model. In the following a brief introduction of object-relational modeling concepts is given. As the implementation is based on an ORACLE 9i database, the corresponding SQL-dialect is followed. Object relational approach with its characteristic of inheritance and aggregation is capable of capturing the various notions of space and time and reflecting them into a single framework extensible to different applications.

2. Spatio-temporal Data Modeling Approaches

2.1 Pyramid framework and event based approach

In the design of object-relational biodiversity data model to support spatio-temporal data, pyramid frame can make enhancement to support a biodiversity data model design. A conceptual framework (also known as pyramid framework) was designed which guide the implementation of the semantic geographical information system (GIS) data model [23]. Conceptual frame work (pyramid) is interrelated with two separate parts one is data component and another is knowledge component. Data component can be divided into three categories: location (position in the spatial three dimensions), time (position alone a time line) and theme (a set of observations, measurements, or attribute values associated with a particular location and time). Data components stores spatio-temporally referenced observational data such as *spectral, climate*, *vegetation* and *other environmental attributes* that maybe queried and visualized to reveal embedded spatio-temporal pattern and relationships. Data component can be computerized as a multi-dimensional, spatio-temporal referenced 'hyper-cube' of observational data that is similar to the 'feature space' concept commonly cited in analysis of remote sensing imagery.

Knowledge component stores information about higher-level semantic 'objects' the geographic entities or process that are describe by the data. Information concerning on object's location, time and composition. All the objects are also placed within two hierarchical relationship structures central to cognitive knowledge representational and object-oriented modeling, plant taxonomy (generalization) and partonomy (aggregation). The taxonomy structures groups similar objects within a category and stores a rule-base that describes how those objects may be identified within the data space. These rules may be derived from expert knowledge or from inductive analysis of the observation data. In an object-relational data model consists of a set of *object classes* (of different *types* or *schemas*). Each object class has an associated set of *objects*; each object has a number of *attributes* with values drawn from certain *domains* or *atomic data types*. Of course, there are additional features, such as object valued attributes (Oid), methods, object class hierarchies, etc. Besides objects, attributes describing *geometries* including *tim*e are of particular interest. Hence we would like to define collections of *abstract data types*, or in fact *many-sorted algebras* containing several related types and their operations, for spatial values changing over time. This section presents a simple and expressive system of abstract data types, comprising data types and encapsulating operations, which may be integrated into a query language, to yield a powerful language for querying spatio-temporal data.



ID	ObjectID	Location	DateTime	Area Code
1	F001	Johor	02/01/1990 10:30 AM	01
2	F002	Skudai	10/02/1991 01:00 PM	02
3	F003	Skudai	10/02/1990 01:05 PM	01
4	F004	Skudai	15/01/1990 01:10 PM	02

Figure 1: Relationship between theme, space and time

Theme attributes record the information of "what is this object" and other related property information. For example, if the theme of this object is "Forest", a Forest must have attributes such as Area or Province it belongs to, Forest Size, and Forest Type, etc. Theme can be inherited according to certain classification hierarchy. For example, theme can be classified as natural resources (forest, earth quake, and rainfall), transportation (rail and highway), settlement (country, city), water system (river, lake, and pool), and cadastre etc.

Location attributes record space information of the object, such as whether it is a point, line, or polygon, and its location. Location attributes can be inherited into subclasses such as the point class, line class and polygon class. Time attributes are represented through timestamps associated with theme and location information. When properties or forest information of an object changed, a new timestamp will be recorded for that object. That means only initial information and changed information will be recorded. Each timestamp is an instant. Interval is the time between two timestamps

In this model, theme, spatial, time attributes are recorded as 3 tables. Each attributes use one list. For example, a flora object has theme attributes as *"Name"*, *"FloraID"*, *"FloraType"* and spatial attributes as *"Area"*. During different season and time of a year, the volume of the lake will change, and the border of the lake expands or contracts. Whenever the volume or the border of the lake changes, a timestamp will be recorded. Thus, as shown in Figure 1, each object has a name list, a value (*1 to N+1*), a Location has area code (*1, 2, 3 and n+1*) and a timestamp or events (*Day, Month, Year, Second, minute, hour)*. Data can be collected same day same time from different area by different data collectors. Sometime data collects from different forests for several days, months and years. Time stores as an events records as EventID, StartTime, Duration and Attribues (types of data). Diagram bellow indicates fields that can be used to associate objects between tables.

3. Object-relational Features

As we know that within the classic relational database model there are only scalar but no complex data types. With the introduction of object types the definition and composition of abstract data types is possible. An abstract data type can be comprised of a multitude of scalar types and again of user-defined complex types and enhances consistency when creating database models.

```
create or replace type Address_TY as object (
        City VARCHAR2 (30),
        Sate VARCHAR2 (30),
        Area_code NUMBER (8),
        Country VARCHAR2 (30));
```

Here create or replace the Address_TY data type after that create Org_Address_TY by using Address_TY data type.

```
create or replace type Org_Address_TY as
   object (
            Org_name VARCHAR2 (30),
            Address  Address_TY,
            Org_url VARCHAR2 (50),
            Phone NUMBER (12),
            Email VARCHAR2 (30));
```

Data types only represent descriptions of data structures. To ensure persistency a table must be bound to the data type. A relation with column objects can be created that is a set of Observer_ID with Observer_name, Observer_org and Observer_Address as:

```
create or replace type Observer as object (
            Observer_ID VARCHAR2 (15),
            Observer_name   VARCHAR2 (30),
            Observer_org VARCHAR2 (50),
            Observer_Address Org_Address_TY );
```

Object views allow user to treat relationally data as they allow synthesizing object from data that continues to be stored in relational tables. Object views have similar functionality like object tables. They can have methods, belongs to collections, reference one another, have object identity and can be accessed from SQL. In addition tables are assigned to object views can be uploaded by using special instead of triggers. Any instance (row) of an object class contains a unique ID called Object_ID (ID). Generally OIDs are system-generated but can also be derived from a primary key column or can be user-defined.

Relationships between objects can be defined using reference types. A reference column stores OIDs of associated (row) objects since column objects do not have inherent OIDs and therefore cannot be referenced. Row objects that belong to a reference can be selected and dereferenced using the DEREF VALUE operator. Modeling object relationships with OIDs and REFs is often compared with foreign key relationships inside the relational model but implicates some benefits like the ability to distinguish between equal and identical objects. Objects are identical if they have one common OID. They are equal if they have different OIDs but coincide with their attributes and values.

In the classical Entity-Relationship-Model aggregations and compositions are modelled through master-detail-relationships. Object-relational dbms provide collection types that contain multiple elements and thus are suitable to express 1: N relationships directly. Each element or value for a collection has the same substitutable data type. The most popular collection types are varrays and nested tables.

A varying array contains a variable number of ordered elements. Varying array data types can be used as a column of a table or as an attribute of an abstract type. Named table types can be created in an Oracle database using SQL. These table types can be used as nested tables to provide the semantics of an unordered collection. As with varray a nested table type can be used as a column of a table or as an attribute of an object type.

Named table types can be created in an Oracle database using SQL. These table types can be used as nested tables to provide the semantics of an unordered collection. As with varray a nested table type can be used as a column of a table or as an attribute of an object type.

```
Create type Address_TY as table of
Org_Address_TY;
```

Multi-Level-Collections that lead to multiple nested tables can be realized if useful for applications but it’s up to the user to balance – a more intuitive representation of data vs. higher complexity of accessing the data.

An object type declaration can also include methods that are defined on values of that type. When using these objects types in tables their methods are also applied to the data of these tables. The method is declared in the *create type* command and the code for the function itself (the definition of the method) is in a separate *create type body* command.

```
create or replace type Taxon_TY as object (
Taxon_Lname VARCHAR2 (20),
Tanon_Ename VARCHAR2 (20),
Taxon_Species    Species_TY,
Taxon_family     Family_TY,
Taxon_area Area_TY,
Taxon_description VARCHAR2 (50),
Taxon_start_date DATE,
Taxon_end_date DATE,
Member function DURATION (start_date,
end_date in Date) return Number);
```

Names the function that is a "member" of the Taxon_TY datatype. Since Duration will return a value, it is a fuction. To define DURATION function, use the create body command, whose full syntax is shown in the Alphabetic Reference.

Let's create the Duration function as a member function within the Taxon_TY datatype. The DURATION function will return duration, in days, of the taxon.

```
create or replace type body Taxon_TY as
member function DURATION (start_date,
end_date DATE)
return NUMBER is
begin
return totaldays (start_date + end_date );
end;
```

The following query executes the duration method in the data type Taxon_TY and returns the duration of all collection objects in relation Observation. Values of components of an object (attributes and methods) are accessed with the dot notation.

```
select Taxon_ID, O.Taxon_Area, Duration
(O.start_date, O.end_date)
from Observation O ;
```

Object types can be mapped to the prevalent OO languages like C++ and Java. Thus object type instances in the database can be accessed and modified to and from C++, Java applications [4].

In the current version Oracle database 10g native data types for storing raster and persistent topology data were added to further extend possibilities to model real world objects based on object-relational database structures. The completeness and therefore the quality of spatial features' descriptions can be further enhanced with the introduction of a new system-defined datatype XMLType that can be used as the datatype for columns in tables and views to create, extract, and index XML data. A feature's position and spatial representation is then modeled by using vector resp. raster types, its characteristics are captured by (complex) attributes and any further information on the spatial object, that cannot be structured to be kept in the data types mentioned, is maintained in xml-documents, all stored natively within the database.

For modeling highly dynamic changes of real world objects the SQL type *TIMESTAMP* can be used which is a high-precision time and date type that allows storing fractions of a second [5]. The timestamp values are basically points on a linear time axis. For slower changes in state we can use *DATE* or create own adapted abstract data types that can express discrete dates or times intervals. The data types can be flexibly extended to meet an application's accuracy requirements and provide attributes to express the fuzziness of a time record, for instance if a time interval cannot be determined as accurate as the data type could store [6].

```
create or replace type TIME_TY as object (
        year NUMBER (4,0),
        month NUMBER (2,0),
        day NUMBER (2,0),
        hour NUMBER (2,0));
```

```
create or replace type TIME as object (
        start_date TIME_TY,
        end_date TIME_TY,
        start_date_per NUMBER (1,0),
        end_date_per NUMBER (1,0));
```

4. Object-relational Application Development

In terms of Biodiversity data of Malaysia, object-relational database technology can be used to model different applications, e.g. the spatio-temporal process of flora or fauna based on its influencing factors or real time deer tracking.

For the development biodiversity data model some major influencing factors have been proved to be crucial. In the model we have judged like data should be natural to human. The structures of space and time are identified as essential for the realization of cognitive systems. Models of spatial temporal data in geographical database representations must incorporate human cognitive principles. Finally model support human knowledge of the dynamic geographical world comprises of three different (and interrelated) subsystems that handle *what*, *where* and *when* aspects of object properties.

```
create or replace type Position_TY as object (
Map VARCHAR2 (30),
Grid25 NUMBER (4),
Grid100 NUMBER (4).

Create or replace type SpatialData as object (
Forest_ID NUMBER (6),
Forest_name VARCHAR2 (30),
Forest_area AREA_TY,
Forest_position Position_TY);
```

Based on this type an object view is created and OIDs are assigned to the spatial datasets of the corresponding forest area:

```
create or replace view Forestdetails_OV
of Position_TY
with object identifier (ForestID) as
select Forest_name, position,
from Forest_area
where AreaName = Johor;
```

The columns of the relational base tables are now accessible as row objects through their corresponding object views. Any detail tables of the underlying relational model could now be simulated by further object views with references that could point to the row objects of the particular upper object view, to allow accessing the data in either way.

Projection operations on moving objects map either to their spatial or to their temporal aspect. Assume that we are interested in the geometric locations where the data was changed at the year January 20, 1990. These can be obtained by:

```
        SELECT Flora_ID, AreaName
        FROM Flora, Area
        WHERE (((Forest_area =#Johor# AND
        #01/20/1990#));
```

This operation computes the *spatial projection* of a spatio-temporal object for the Johor forest area. For an evolving region the trajectory operation returns an object of the spatial type *region* which results from projecting the union of the region values for the Date 20th January, 1990.

The following query inquires about the largest collection of flora areas.

```
        SELECT Area (max (Extent)) FROM
        ForestArea
        WHERE Type = "Flora"
```

The query demonstrates an example of a *spatio-temporal aggregation* operation max which is an extension of the well known aggregation operator in SQL of the same name. It is here applied to a collection of evolving regions contained in a relation column and computes a new evolving region. Internally, this operator is based on a binary function ***MAX$_{ST}$*** applied to two evolving regions ***R$_1$*** and ***R$_2$*** 'and yielding a new evolving region in the following way:

$$MAX_{ST} (R_{1,} R_2):= \{(t, r) |t \in \text{time} \quad r \wedge MAX_{geo} (R_1 (t), (R_2 (t))\}$$

This definition uses a function ***MAX$_{geo}$*** which is applied to two regions R_1 and R_2 and which returns larger of both regions.

$$MAX_{geo}(R_{1,} R_2)= \begin{cases} r_1 & \text{If area } (r_{1)} > \text{area } (r_2) \\ r_2 & \text{Otherwise} \end{cases}$$

Altogether this means that for 2 evolving regions $R_{1,\ldots}$, R_n we first compute the evolving region R = MAX_{st} (R_1,….MAX_{st}(R_{n-1}, R_n)..). Afterwards, we apply the raise area of R, which computes the area of R at all times as a temporal real number.

5. Conclusion

Object-relational technology in DBMS helps to model and analyze spatial-temporal events as it provides highly flexible means for storing, manipulating and validating diverse and complex data structures. As for applications in Malaysian biodiversity data, the real challenge is to understand the relationship of the most complex processes and find definite and describable entities that can be part of a database model. Thus it is often useful to extend the concept of objects from a conventional point of view to a more general sense, so to speak to describe real world scenes (instead of objects) that show on their part some correlation with a set of commonly used objects. To achieve a smoth interplay between the embedded language and an embedded system of abstract data type, a few interface facilities and notations are needed, expressible in one form or another in most object-oriented or object-relational query language.

6. References

[1] Thang, H.C. (2004). Forest Management and Development in Malaysia. Paper presented at the Malaysian Timber Council's Familiarisation Programme for European Trade Representatives, 12 July 2004, Kuala Lumpur, Malaysia. Unpublished. UNEP (2002) Malaysia-Study Area. http://www.rrcap. unep.org/lc/cd/

[2] Jeremy L, Mernnis, Donna J. Peuquet, Diansheng Guo, (2000), "A Semantic GIS Database Model for the Exploration of Spatio-temporal Environmental Data", GIS/EM4, Alberta, Canada, September 2-8, 2000

[3] Bisby, F.A. (2000), "The Quiet Revolution: Biodiversity Informatics and the Internet." Science 289, 2309-2312

[4] Lee, G., (2003), "Simple Strategic for complex Data: Oracle 9i Object-relational Technology: An Oracle Technical White Paper", October 2003.

[5] Qian, L., (2004), "Oracle database 10g-Development Spatial Application Using Oracle Spatial and Map Viewer: An Oracle technical White Paper, February 2004

[6] Plabst. S., (2001), "Object-relational Data models for Geo information System", Technical University Munchene

[7] Kevin Loney, George Koch, (2003), "Oracle9i: The complete reference", McGraw-Hill Publication, USA.

[8] Langran, G. (1992), "Time in Geographic Information Syetems:, Taylor & Francis, London

[9] Sellis, T. (1999). Research issues in Spatiotemporal Database Systems, Springer-Verlag Berlin Heidelberg.

[10] Martin Erwig, R. H. G., Markus Schneider, Michalis Vazirgiannis (1999). " Spatio-Temporal Data Types: An Approach to Modeling and Querying Moving Objects in Databases." Geoinformatica 3(3): 269-296.

[11] Guting, R. H. (1994). "An Introduction to spatial Database Systems." VLDB Journal 4: 357-399.

[12] Tansel, A. U. and J. Clifford (1993). Temporal Databases: Theory, Design, and Implementation, Benjamin/Cummings Publishing Company

[13]Freksa, C. (1998). Spatial and Temporal Structures in Cognitive Processes, Springer Verlag Heidelberg.

[14] Peuquet, D. J. (2001). "Making Space for Time: Issues in Space-Time Data Representation." GeoInformatica 5(1): 11- 32.

[15] Mennis, J. L., D. J. Peuquet, et al. (2000). "A Conceptual Framework for Incorporating Cognitive principles into geographical database representation." International Journal of Geographical Information System 14(6): 501-520

[16] M. Bouzeghoub, "Morse: a functional query language and its semantic data model", proc. of int. Conf. on Trends an applications of databases, IEEE-NBS, Gaithersburg, USA, 1994