GJCRT : Global Journal of Computer Science Research & Technology
ISSN (online): 2349- 0608

Dimension- Based Generalization In Object Relational Databases Using Object Cube Model

Dr. Surender Kumar
S.G.T.B Khalsa college
Anandpur Sahib
drsurender.sgtb@gmail.com

Abstract

The increasing popularity of object relational database systems in advanced database applications. In this paper the work is proposed for class based generalization with an object cube model on relational database. The study shows that a dimension- based class generalization mechanism can be developed for object cube construction. Data analysis applications typically aggregate data across many dimensions looking for anomalies of unusual patterns. The SQL aggregate functions and the Group By operator produce Zero – dimensional or one – dimensional aggregates. Applications need the N- dimensional generalization of these operators. This operator is called the data Cube. The cube operator can be embedded in more complex non- procedural data analysis programs. The cube operator treats each of the N aggregation attributes as a dimensions of N- space.

INTRODUCTION

The paper" Dimension- Based Generalization In Object Relational Database Using object Cube Model" deals with an object cube model for the class- generalization in object- oriented Databases . A multi- dimensional database is called a data cube which is a database consisting of huge set of facts or multi -dimensional points and a relatively small set of dimensions with respect to which data analyzed . A data cube allows data to be viewed and modeled in multiple dimensions. In Data Warehouse a cube is N- dimensional . Data cube is a metaphor for multi- dimensional data storage. With the availability of operators in object generalization. We can explore the class- based generalization of a set of relevant objects. Generalization of a class of objects can be viewed as a sequence of set- oriented generalization processes, each transforming a class into a relatively more generalized class. The class fed into a generalization processes is called is the initial Working class. The class obtained by the application of a generalization operator is called the Resulting class.

RELATED WORKS

Operators for generalization of attributes, methods and Complex Components:- An object oriented database organizes a large set of complex data objects into classes which are in turn organized into class/ subclass hierarchies. Each object in a class is associated with,

(a). an object identifies.

(b). A set of a attributes which may contain sophisticated data structures, set- or- list valued data, class composition hierarchies, multimedia data.

(c). A set of methods which specify the computational routines or roles associated with the object class.

To facilitate the development of knowledge discovery mechanisms in object oriented databases , it is important to implement efficiently a relatively small set of generalization operators on which large set of possible generalization operations can be developed. The generalization of a components P of an object Oi can be written in an abstract way of Gen (Oi P), Where Gen is an abstract object, generalization operator which can be transformed into a concrete operation based on the role of the component and the specific learning requirement. The generalized component of an object can be further generalized by applying Gen. again , which could be the same of different generalization operators compared with the one applied in the last generation. If the generalization is performed by applying the same Sequence of generalization operators on a component P of Oi.

Then Gen^{n-1} (Gen(Oip))= Gen (Gen^{n-1} (OiP)).

Class generalization

With the availability for generalization, one can explore the class- based generalization of a set of relevant objects. Since a set of object may share many attributes and methods and the generalization of each attribute and method may apply a sequence of generalization operators.

The major issue becomes how to cooperate the generalization processes among different attributes and methods. Generalization of a class of objects can be viewed as a sequence of set- orient generalization processes, each transforming a class into relatively more generalized class. The class fed into a generalization Process is called the working class w, with the initial one called the initial working class Wo . The class generated by the application of generalization operator is called the resulting class R.

An object cube model

A popular conceptual model that influences the design of data warehouse and development of every query engines and front- end tools for on-line analytical processing is the multi-dimensional model of data in the data warehouse. A multidimensional database called as a data cube, is a database consisting of large set of facts or multidimensional points and a relatively small set of dimensions with respect to which data is analyzed . The dimensional tables are where the textual description of the dimensions of the business is stored . Each dimensions has a set of attributes. The facts are numerical measures collected in the corresponding dimensions and provided in a data warehouse for data analysis, e.g sales amount, budget etc. The fact table stores the key of multiple dimensions and the numerical measures of the business . The object cube model is the further development of the data cube model for data warehouse in object oriented databases.

An object cube is a multidimensional database constructed on top of a generalized class in an object- oriented database. The generalized class contains a set of generalized attributes serving as dimensions of the object cube, and one or a set of attributes collecting aggregate values or other extents and serving as measures of the cube.

The computation of an object cube can be performed as follows.

1. A class of object interested in the study is collected by object – oriented query processing which could be specified by an object- oriented query or simply done by a collection of potentially interested data sets.
2. Object generalization models are applied to the collected class, which generalizes. The data collected to a minimally generalized class which fits a multidimensional data model, which generalized attribute correspond to dimensions and

values aggregated into the cube cell as measured based on the measure computation methods defined in the object cube. This minimally generalized class forms the base cuboid of the object cube. Further generalization on the base cuboid derives other cuboids, which could either be Precomputed before a query is submitted to the system and stored for future use, or be performed on the fly to answer a particular mining query. The computed object cube, form the base for on- line analytical processing and data mining in object- oriented databases.

Algorithm:- Generalization in an object oriented database by dimension- based induction based on user's generalization request.

Method

The dimension- based induction is performed in the following steps:

1. Init class:- Execute DB query to derive the initial working class W_o. ie. collect the set of task- relevant data by an object-oriented database query based on the generalization requests.
2. Base Gen:- Generalize the initial working class W_o . minimally to form a minimally generalized base cuboid cube 0.
   a. collect the distinct values for each attribute ai in the initial working class and count the number distinct values. For numerical values, generate minimal intervals for registration of the number of intervals.
   b. Taking a group of hierarchical attributes as one dimension and count, sum etc as measures, construct the base cuboid.
3. **PreGen :-** Prepare for prime cube generation
   a. compute the desired level Li for each dimensions a_i based on its dimension threshold Ti.
   b. Determine the mapping pairs (v, vi) for each dimensions a_i and vi is its corresponding generalized value at level Li.
4. Prime Gen: Derive the prime class cube p.

RESULT AND DISCUSSION

Generalization- based data mining in object- oriented database using an object cube yields a generalized form of the database according to the query given as the input. Initial stage of this generalization executes the database query and an initial working class, W_o is derived. This initial working class is then generalized minimally to form the minimally generalized base cuboid cube 0 from this base cuboid computing the desired level for each dimension prepared the prime cube. In the final stage class, cube P is derived.

Generalization mechanism are developed for object cube construction using a dimension-based class generalization mechanism knowledge discovery in object-oriented data bases shares many attributes with the knowledge discovery in relational database, especially when object-oriented data are generalized into prime generalized classes, which are often in the form of generalized tables or cuboids.

SYSTEM MODEL

Objective:- Implement an object cube model for dimension-based class generalization algorithm in object-Relational databases . The proposed system uses object relational databases is suitable for class generalization in an efficient way.

ORDBMS An object relational database (ORD) or object relational data base management system (ORDBMS) is a relational database management system that allows developers to integrate the data base with their own custom

data types and methods.The term object-relational database is sometimes used to describe external software products running over traditional DBMS to provide similar features; these systems are correctly referred to as object-relational mapping systems.ORDBMS employ a data model that attempts to incorporate object oriented features into RDBMS.All database information is stored in tables. but some of the tabular entries may have riches data structure, termed abstract data types (ADTS). An ORDBMS supported an extended form of SQL called SQL3 that is still in the development stages.

The extensions are needed because ORDBMS have to support ADTS. The ORDBMS has the relational model in it because the data is stored in the form of tables having rows and columns and SQL is used as the query language and the result of a query is also tables or tuples (rows) . But the relational model has to be drastically modified in order to support the classic features of object –oriented programming.

Hence the characteristics of an ORDBMDS are:-Basic data type extension, support complex objects , Inheritance, and rule systems ORDBMS allows users to define data types , functions and operators as a result, the functionality of the ORDBMS increases along with their performance. An object relational database (ORD) or object relational database management system that allows developers to integrate the database with their own custom data types and methods. The term object relational data base is sometimes used to describe external software products running over traditional DBMS to provide similar features; these systems are more correctly referred to as object-relational mapping systems.Where as RDBMS or SQL-DBMS products focused on the efficient management of data drawn form a limited set of data types (defined by the relevant language standards) an object-relational DBMS allows software developers to integrate their own types and methods that apply to them into the DBMS. The goal of ORDBMS technology is to allow developers to raise the level of abstraction at which they view the problem domain.Object- relational data bases add to the relational model the possibility of working with new and complex data and applications. One widely accepted mechanism

for assuring the quality of are object-relational; data base is the use of metrics and it is important to formalize the metrics for having a better understanding of their definitions. Metrics formulization assures the reliable repetition of their computation and facilitates the automation of metrics collection.

In this paper I present the formulization of a set of metrics defined for object- relational database described using ontology. Cube model is to apply the object cube base generalization techniques on object relational database. The main drawback of OORDBMS has been poor performance. ORDBM is basically a complex multidimensional data model where data processing techniques are required to reduce complexity and to improve the quality of data in order to improve the accuracy and efficiency of the subsequent mining Process.We can improve the efficiency of the ORDBMS by using user Defined functions and other techniques

REFERENCES

1. V. Harinarayanan , A.Rajaraman, J. D. Vllman "Implementing data cube efficiently".
2. Jiawari Han, Micheline Kamber "Data mining: Concepts and techniques".
3. Jim Gray, Surjit Chaudhari, "Data Mining and knowledge discovery".
4. M. Brown P. 1999 'Object- Relational DBMS Tracing The Next GreatWave'.
5. Jim Gray, Surajit Choudhuri ata, " Cube : A Relational Aggregation operator Generalizing Group- by Cross- Tab, and Sub-Totals".
6. J. Hanj. J Chiang, S. Chee, M. Kamber. DB Miner:- A system for data mining in relational databases and data warehouse