GLOBAL QUERY PROCESSING IN DBMS

S.rajadharani,Student
L.prabha, Student,
M.swetha, student,
Department of commerce.

ABSTRACT:

The huge number of autonomous and heterogeneous data repositories accessible on the "global information infrastructure" makes it impossible for users to be aware of the locations, structure/organization, query languages and semantics of the data in various repositories. There is a critical need to complement current browsing, navigational and information retrieval techniques with a strategy that focuses on information content and semantics. In any strategy that focuses on information content, the most critical problem is that of different vocabularies used to describe similar information across domains. We discuss a scalable approach for vocabulary sharing. The objects in the repositories are represented as intensional descriptions by pre-existing ontologies expressed in Description Logics characterizing information in different domains. User queries are rewritten by using interontology relationships to obtain semantics preserving translations across the ontologies

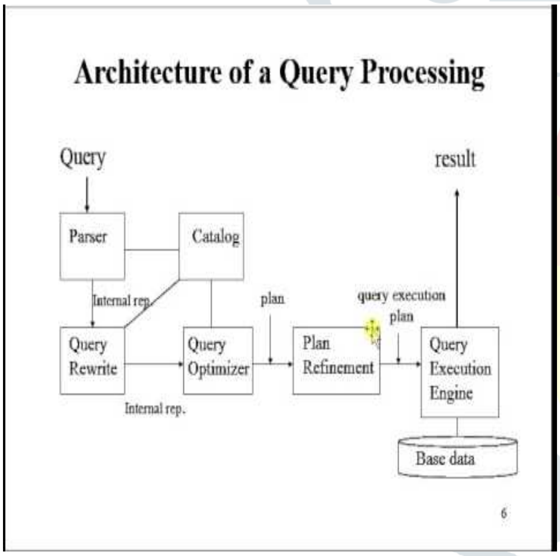
INTRODUCTION:

- We extend or build upon some of the above approaches by using metadata to capture the information content of the repositories. We represent intensional descriptions to abstract from the structure and organization of the individual repositories as intensional metadata .
- Our query processing approach enables the user to subscribe to the vocabulary he is familiar with. Even when the user poses a query using terms from one ontology, relevant concepts may be described in other ontologies, and the corresponding information may be accessible through semantic relationships associated with those ontologies.
- This requires a strategy that expands a query on one ontology to other relevant ontologies. In this context, we need a strategy to support this query expansion, but restrict it appropriately to manage the quality of information returned.
- The most critical problem in characterizing the information content is that of different vocabularies used to describe similar information across domains.
- This leads to different terms and constraints being used to characterize similar information. Interoperation across ontologies is achieved by traversing semantic relationships defined between terms across ontologies.
- User queries are rewritten in a semantics preserving manner by replacing them with synonym terms from different ontologies; hyponym and hypernym terms can also be used and the loss of information measured.
- The key objective of our approach is to reduce the problem of knowing the structure and semantics of data in the huge number of repositories in a global information system to the significantly smaller problem of knowing the synonym relationships between terms across ontologies.
- We have developed a prototype system which supports querying of real-world repositories.
- Some data has been down-loaded into local databases, some are in plain files and others are remote repositories accessed on-line through WWW supported forms.
- Determining the relevance of an information source before accessing the underlying data. Representation of ontologies in a knowledge representation system can enable the use of the underlying inference mechanisms for determination of relevance.
- Supporting wider accessibility of data via multiple ontologies and interoperation across them based on semantic relationships such as synonyms, hyponyms and hypernyms.

QUERY PROCESSING ARCHITECTURE:

- A CIS query processor architecture, has been proposed 291 to partition the connectivity problems. This architecture encapsulates the first-order issues into the local query processors (LQPs). Each LQP is responsible for the physical connection to a particular remote machine and the required database accesses.

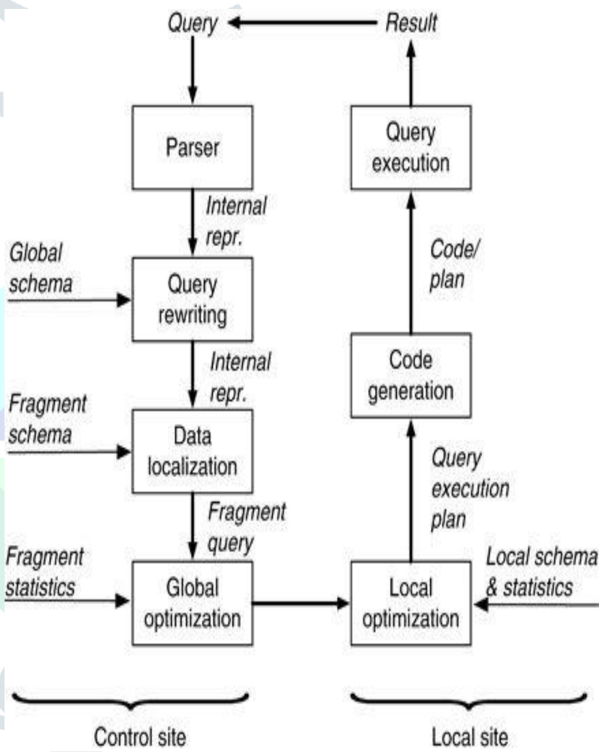
- An important design choice for the CIS query processor architecture is the data model to be used for the representation of the data in the LQPs, the Global Query Processor (GQP), and the Application Query Processors (AQPs)
- Although many semantic data models and schema integration methodologies have been proposed and discussed in depth in the literature our experiences in developing CIS have led us to a combination of the Entity Relation and relational models: The relational model is theoretically mature, information structure oriented, and commercially successful.



- The ER diagram is better suited for representing the semantics of the underlying model. It specifies the relationship between entities explicitly
- We have, therefore, chosen to extend the relational model in our work, and to incorporate the useful constructs developed in the semantic data
- models, such as the ER diagram where relevant global query processing in CIS.

GLOBAL QUERY PROCESSING EXAMPLE:

An implication of the relational/ER approach is that a relational interface is required for each LQP. In representing the local database, the relationship between entities is explicitly defined as link entities. By requiring a relational interface for each LQP as well as for the GQP, a coherent CIS query processing mechanism based on the relational theory can be developed. The mechanism is exemplified Below using a simplified placement assistant system. A global ontology-based approach which supports expression of complex constraints as a part of the query. This involves development and integration of domain-specific ontologies into a common global ontology and partitioning it into microtheories. This approach transfers the burden of information correlation and filtering on the query processing system. However it can be very difficult to support because of the complexity involved in integrating the ontologies and maintaining consistency across concepts (originally)from different ontologies.



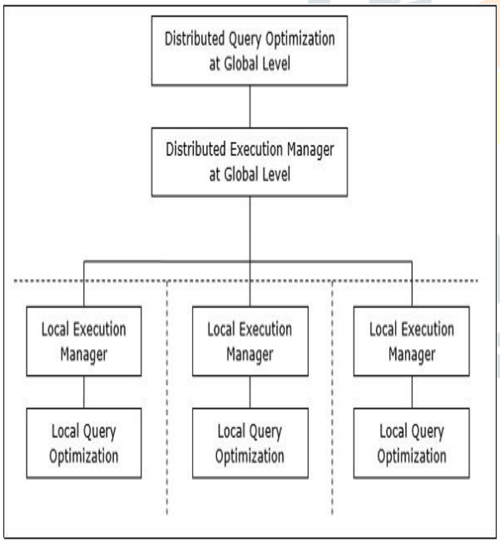
It takes as input a user query expressed in DLs using terms from a chosen user ontology. The query processor navigates other component ontologies of the global information system and translates terms in the user query into the component ontologies preserving the semantics of the user query. Our focus in this paper is supporting “semantically rich” queries in an environment where different vocabularies are used.

GLOBAL QUERY PARSING:

Consider the query "Find all the names and degrees of the alumni who are CEOs of corporations" for the query to the global schema. An SQL expression and the corresponding relational algebra for the CIS query. The parse tree for the relational algebra where n denotes project, a denotes select, and denotes join.

GLOBAL QUERY OPTIMZER:

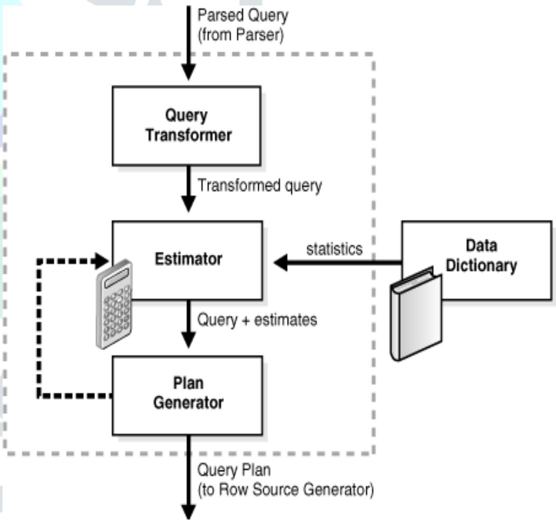
- we discuss how to incorporate the query optimization techniques introduced in the last section into a global query optimizer designed for the CORDSMDBS. We describe the query processing approach. procedure and components in the system .
- Note that not all components have been fully implemented. Mixed Compilation and Interpretation Approach There are two approaches for processing queries One is the compilation approach The other is the in terpretation approach .
- The former performs comprehensive query optimization at compile time to generate an e client execution plan .The same execution plan can be repeatedly invoked at run time.
- This approach is suitable for stored or embedded queries which usually need to be executed many times. The interpretation approach on the other hand performs query optimization on the while a query is executed. No execution plan is stored .
-



- Since the time for query optimization the response time for a query only simple query optimization is usually performed in this approach The interpretation approach is suitable for ad hoc queries. Since stored and embedded queries are assumed in our environment it appears that the compilation approach should be adopted .
- However in accurate or incomplete information at the global level in an MDBS may cause the query optimizer to generate an in efficient execution plan for a query at compile time .
- To reduce the chance for performing expensive re optimization at run time some part of a query may be left un compiled in the system if desired information is not available at compile time.
- In other words. an incomplete execution plan may be generated. The un compiled part of a query is then optimized and interpreted at run time based newly available information.

MODULES IN THE GLOBAL QUERY OPTIMIZER :

The global query optimizer is a major subsystem of the CORDSMDBS. It consists of several smaller modules .It performs query optimization to generate a good execution plan for a query at compile time. It also performs adaptive query optimization to improve an execution plan or complete an incomplete execution plan at run time



OPTIMIZATIONO CONTROLLER IN QUERY PROCESSING:

- This is a control module of the global query optimizer At compile time. it accepts the query parse tree from the request coordinator It passes the tree to the semantic optimizer for improvement.
- It then invokes the processing strategy analyser to choose an client execution strategy for the improved query Finally the execution plan generator is invoked to generate an execution plan.

- according to the chosen strategy At run time optimization controller passes runtime information to the runtime optimizer to improver complete an execution plan adaptively
- Semantic query optimizer This module transforms a user query expressed in the form of parse tree into another more client equivalent query based on the semantic information provided by the multi database catalogue.
- Some heuristic rules such as per forming unary relational operations as early as possible are also used to rewrite a user query Processing strategy analyser.
- A query processing strategy describes the major decisions for processing a global query such as how to decompose the query where to perform local queries and how to integrate local results
- However such a processing strategy for a query is usually not unique Using the local cost parameters estimated by the query sampling method a module called processing strategies analyser alternative processing strategies for the given query and choose a good one among them .
- This module may also perform some probing queries on local DBSs to obtain update some information required for the analysis .

ISSUES AND APPRROACHES OF GLOBAL INFORMATION SYSTEM :

We now revisit the global query exemplified .Suppose that CEO is not an attribute of the Corporation entity in the alumni database. In this case, although the GSC table remains to be the same, the corresponding EIC table will be different. The GSC driver recognizes that the "CEO" attribute does not exist , but does exist in the company database. Therefore, the o4 action routine is invoked instead . It first inserts the tuple which retrieves the company entity from the company database without any restriction. In addition, register the company database as the location for query execution. Secondly, it inserts which joins company on the primary key of GCompany, i.e., join Corporation name with Company name. In addition, it registers the GQP as the location for query execution. Now that has all the attributes required for the a operation, the restriction is made in the GQP, and the result is retained in depicts the equivalent parse tree for the case.

CONCLUSION:

- We have described an architecture for Global Information Systems that is especially tailored to address the challenges discussed in Section 2. Our approach is based on: Use of intensional metadata descriptions to model and query the information content in various repositories, and Ontology based interoperation by navigating terminological relationships, to handle the vocabulary problem.
- Novel contributions in this paper include the representation of the synonym relationships between terms across ontologies, an algorithm for translating the intensional query expression into different ontologies, and an algorithm to combine the partial translations in different ontologies such that they satisfy the constraints in the original query.
- The heterogeneity in the values is managed by using transformer functions stored by the IRM. Unlike a regular thesaurus, the expressiveness of the DL systems allows using descriptions when a defined term has no translation.
- The methods described in this paper are implemented in a prototype system developed at the LSDIS lab, accessible .This prototype accesses information in real-world data repositories using pre-existing real-world ontologies in the domain of bibliographic information by using pre-existing real-world ontologies and real-world repositories, helps the user to observe a semantic conceptual view of a global information system by giving her/him the ability to browse multiple domain specific ontologies as opposed to individual heterogeneous repositories.
- Uses the CLASSIC system and demonstrates a practical use of DLs for interoperation across domain specific ontologies to support querying and information organization in a global information system.

REFERENCE:

1.Eduardo mena ”global query processing”.
2. V.kashyap “query information”
3. carl sagan “ontology based query processing”.

REFERENC LINKS:

1.https://www.researchgate.net

2. https://www.ijirae.com

3. https://www.slideshare.net