

QUERY REWRITING IN A RELATIONAL DATA HARMONIZATION FRAMEWORK

Priority Claim

[01] This application claims priority to U.S. non-provisional application serial number 15/227,605, filed 3-August-2016, and to U.S. provisional application serial number 62/359,547, filed 7-July-2016.

Technical Field

[02] This disclosure relates to database systems, and to queries executed in database systems.

Background

[03] The processing power, memory capacity, available disk space, and other resources available to processing systems have increased exponentially in recent years. Database systems in particular have grown in capacity and capability to power extremely complex and sophisticated analyses on immense datasets that discover useful information, suggest conclusions, and support decision-making. Improvements in database systems will further advance database capabilities.

Summary

[03a] According to a first aspect of the present invention, there is provided a method including: receiving an initial model query, the initial model query including query statements configured to interrogate a predetermined semantic model of a relational database; obtaining the predetermined semantic model of the relational database; and executing query translation rules with a query rewriting processor to analyze the initial model query and the predetermined semantic model of the relational database to: generate an enriched model query by enriching the initial model query with knowledge of database interactions determined from the analyzed initial model query and predetermined semantic model; and translate the enriched model query to a relational database query configured for execution against the relational database to obtain database data responsive to the query statements.

[03b] According to a second aspect of the present invention, there is provided a system including: a query source interface configured to: receive an initial model query, the initial model query including query statements configured to interrogate a predetermined semantic model of a relational database; query rewriting circuitry in communication with the query source interface, the query rewriting circuitry configured to: obtain the initial model query; obtain the predetermined semantic model of the relational database; and execute query translation rules with a query rewriting processor to analyze the initial model query and the predetermined semantic model of the relational database to: generate an enriched model query by enriching the initial model query with knowledge of database interactions determined from the analyzed initial model query and predetermined semantic model; and translate the enriched model query to a relational database query configured for execution against the relational database to obtain database data responsive to the query statements.

[03c] According to a third aspect of the present invention, there is provided a system including: a query source interface configured to: receive an initial model query including <subject, predicate, object> query statements written to interrogate a specific semantic model of a corresponding relational database; and query rewriting circuitry including: a semantic model database including a library of semantic models including the specific semantic model; pre-defined query translation rules including at least: a 1 -to-many translation rule; a many-to-many translation rule; and a literal translation rule; and translation circuitry configured to: obtain the initial model query from the query source interface; obtain the specific semantic model of the relational database from the library of semantic models; and execute the query translation rules to analyze the initial model query and the specific semantic model of the relational database to: determine an enriched model query by enriching the initial model query with knowledge of database interactions determined from the analyzed initial model query and specific semantic model; and translate the enriched model query into a relational database query configured for execution against the relational database to obtain database data from the relational database.

Brief Description of the Drawings

- [04]** Figure 1 shows a query rewriting architecture and query rewriting logic.
- [05]** Figure 2 illustrates a semantic model for an example database.
- [06]** Figure 3 shows another example database.
- [07]** Figure 4 shows a corresponding semantic model for the database in Figure 3.

CLAIMS

What is claimed is:

1. A method including:
 - receiving an initial model query, the initial model query including query statements configured to interrogate a predetermined semantic model of a relational database;
 - obtaining the predetermined semantic model of the relational database; and
 - executing query translation rules with a query rewriting processor to analyze the initial model query and the predetermined semantic model of the relational database to:
 - generate an enriched model query by enriching the initial model query with knowledge of database interactions determined from the analyzed initial model query and predetermined semantic model; and
 - translate the enriched model query to a relational database query configured for execution against the relational database to obtain database data responsive to the query statements.
2. A method according to claim 1, where:
 - the predetermined semantic model represents tables and table relationships that are present in the relational database, but does not include the database data that is stored in the relational database.
3. A method according to either claim 1 or claim 2, where:
 - executing query translation rules includes:
 - identifying database tables referenced by the query statements.
4. A method according to claim 3, where:
 - executing query translation rules includes:
 - determining whether a one-to-many database table relationship exists among the database tables.
5. A method according to claim 4, further including:

modifying the initial model query to obtain an enriched model query that captures the one-to-many database table relationship.

6. A method according to any one of claims 3 to 5, where:
executing query translation rules includes:
 - determining whether a many-to-many database table relationship exists among the database tables.
7. A method according to claim 6, further including:
modifying the initial model query to obtain an enriched model query that captures the many-to-many database table relationship.
8. A method according to claim 7, where:
modifying includes:
 - identifying a mutual table referenced by other tables in the many-to-many database table relationship; and
 - changing the initial model query to cause at least two joins between the mutual table and the other tables in the relational database query.
9. A method according to any one of the preceding claims, where:
executing query translation rules includes:
 - identifying a literal in the initial model query; and
 - creating a condition that includes the literal in a 'Where' clause in the relational database query.
10. A system including:
a query source interface configured to:
 - receive an initial model query, the initial model query including query statements configured to interrogate a predetermined semantic model of a relational database;
 - query rewriting circuitry in communication with the query source interface, the query rewriting circuitry configured to:
 - obtain the initial model query;
 - obtain the predetermined semantic model of the relational database; and

execute query translation rules with a query rewriting processor to analyze the initial model query and the predetermined semantic model of the relational database to:

generate an enriched model query by enriching the initial model query with knowledge of database interactions determined from the analyzed initial model query and predetermined semantic model; and

translate the enriched model query to a relational database query configured for execution against the relational database to obtain database data responsive to the query statements.

11. A system according to claim 10, where:

the predetermined semantic model includes table and table relationship metadata that represents the relational database, but does not include the database data itself.

12. A system according to either claim 10 or claim 11, where:

the query translation rules are configured to identify database tables referenced by the query statements.

13. A system according to claim 12, where:

the query translation rules are configured to determine whether a one-to-many database table relationship exists among the database tables.

14. A system according to claim 13, where:

the query rewriting circuitry is configured to modify the initial model query to obtain an enriched model query that captures the one-to-many database table relationship.

15. A system according to any one of claims 12 to 14, where:

the query translation rules are configured to determine whether a many-to-many database table relationship exists among the database tables.

16. A system according to claim 15, where:

the query rewriting circuitry is configured to modify the initial model query to obtain an enriched model query that captures the many-to-many database table relationship.

17. A system according to claim 16, where:
 - the query rewriting circuitry is configured to:
 - identify a mutual table referenced by other tables in the many-to-many database table relationship; and
 - change the initial model query to cause at least two joins between the mutual table and the other tables in the relational database query.
18. A system according to any one of claims 10 to 17, where:
 - the query translation rules are configured to identify a literal in the initial model query; and
 - the query rewriting circuitry is configured to create a condition that includes the literal in a 'Where' clause of the relational database query.
19. A system including:
 - a query source interface configured to:
 - receive an initial model query including <subject, predicate, object> query statements written to interrogate a specific semantic model of a corresponding relational database; and
 - query rewriting circuitry including:
 - a semantic model database including a library of semantic models including the specific semantic model;
 - pre-defined query translation rules including at least:
 - a 1 -to-many translation rule;
 - a many-to-many translation rule; and
 - a literal translation rule; and
 - translation circuitry configured to:
 - obtain the initial model query from the query source interface;
 - obtain the specific semantic model of the relational database from the library of semantic models; and

execute the query translation rules to analyze the initial model query and the specific semantic model of the relational database to:

determine an enriched model query by enriching the initial model query with knowledge of database interactions determined from the analyzed initial model query and specific semantic model; and

translate the enriched model query into a relational database query configured for execution against the relational database to obtain database data from the relational database.

20. A system according to claim 19, where:

the 1-to-many translation rule is configured to cause the query rewriting circuitry to:

identify a variable in the initial model query that is an <object> in a first query statement and also a <subject> in a second query statement;

identify a <predicate> table name in the first query; and determine that a 1-to-many relation exists when the first query is a correct model query;

the many-to-many translation rule is configured to cause the query rewriting circuitry to:

identify a many-to-many database table relationship including a mutual table between other tables; and

change the initial model query to explicitly cause at least two joins between the mutual table and the other tables in the relational database query; and

the literal translation rule is configured to cause the query rewriting circuitry to:

identify a literal in the initial model query; and

create a condition that includes the literal in a 'Where' clause of the relational database query.