A Queries

The queries on the TPC-H schema are the following:

1. \textit{TPCH-1}: The query finds information on customers with a given account balance that have placed orders with a given total price, as well as prices of the order lines. The correlation between attributes \( o\text{-}totalprice \) and \( l\text{-}extendedprice \) causes the selectivity of the join \( L \bowtie O \) to be very high.

   \[
   \begin{align*}
   \text{select } & c\text{-}name, c\text{-}address \\
   \text{from } & \text{orders, lineitem, customer} \\
   \text{where } & o\text{-}orderkey=1\text{-}orderkey \quad \text{and} \\
   & o\text{-}totalprice=x \quad \text{and} \\
   & o\text{-}custkey=c\text{-}custkey \quad \text{and} \\
   & c\text{-}acctbal=z; \\
   \end{align*}
   \]

2. \textit{TPCH-2}: The query finds information on customers with a given account balance, that have placed orders with a given ship date and order date. The strong correlation between \( o\text{-}orderdate \) and \( l\text{-}shipdate \) causes the selectivity of the join \( L \bowtie O \) to be very high.

   \[
   \begin{align*}
   \text{select } & c\text{-}name, c\text{-}address \\
   \text{from } & \text{lineitem, orders, customer} \\
   \text{where } & l\text{-}orderkey=o\text{-}orderkey \quad \text{and} \\
   & l\text{-}shipdate=x \quad \text{and} \\
   & o\text{-}orderdate=y \quad \text{and} \\
   & o\text{-}custkey=c\text{-}custkey \quad \text{and} \\
   & c\text{-}acctbal=z; \\
   \end{align*}
   \]

3. \textit{TPCH-3}: The query finds information on suppliers with a given account balance that have participated in finished orders with a given ship date. The strong correlation between \( o\text{-}orderstatus \) and \( l\text{-}shipdate \) causes the selectivity of the join \( L \bowtie O \) to be very low.

   \[
   \begin{align*}
   \text{select } & s\text{-}name, s\text{-}address \\
   \text{from } & \text{orders, lineitem, supplier} \\
   \text{where } & o\text{-}orderkey=l\text{-}orderkey \quad \text{and} \\
   & o\text{-}shipdate=x \quad \text{and} \\
   & l\text{-}shipdate=y \quad \text{and} \\
   & l\text{-}shipdate=s\text{-}shipkey \quad \text{and} \\
   & s\text{-}acctbal=z; \\
   \end{align*}
   \]

4. \textit{TPCH-4}: The query finds information on pricing and cost of orders that were shipped after a given date, and ordered before a given date, with a further selection on available quantity. The correlation between \( o\text{-}orderdate \) and \( l\text{-}shipdate \) causes the selectivity of the join \( L \bowtie O \) to be very low.

   \[
   \begin{align*}
   \text{select } & o\text{-}totalprice, ps\text{-}suppcolyst \\
   \text{from } & \text{lineitem, orders, partsupp} \\
   \text{where } & l\text{-}shipdate=x \quad \text{and} \\
   & l\text{-}suppkey=s\text{-}suppkey \quad \text{and} \\
   & ps\text{-}availqt=y; \\
   \end{align*}
   \]

The queries on the IMDB schema were the following:

1. \textit{IMDB-1}: The query finds titles produced in 2009, for which the distributing company was US-based.

   \[
   \begin{align*}
   \text{select } & T\text{-}title \\
   \text{from } & \text{title as } T, \\
   & \text{movie_companies as } MC, \\
   & \text{company_name as } CN \\
   \text{where } & MC\text{-}company_id=CN\text{-}id \quad \text{and} \\
   & MC\text{-}country_code='[us]' \quad \text{and} \\
   & T\text{-}id=MC\text{-}movie_id \quad \text{and} \\
   & T\text{-}production_year=2009; \\
   \end{align*}
   \]

2. \textit{IMDB-2}: The query finds movie titles produced in 2009 for which the distributing company is known.

   \[
   \begin{align*}
   \text{select } & T\text{-}title \\
   \text{from } & \text{title as } T, \\
   & \text{movie_companies as } MC, \\
   & \text{company_name as } CN \\
   \text{where } & MC\text{-}company_id=CN\text{-}id \quad \text{and} \\
   & MC\text{-}company_type_id=1 \quad \text{and} \\
   & T\text{-}kind_id=1 \quad \text{and} \\
   & T\text{-}id=MC\text{-}movie_id \quad \text{and} \\
   & T\text{-}production_year=2009; \\
   \end{align*}
   \]

3. \textit{IMDB-3}: The query retrieves titles together with trivia on actors that have played in the movie.

   \[
   \begin{align*}
   \text{select } & T\text{-}title, CI\text{-}info \\
   \text{from } & \text{cast_info as } CI, \\
   & \text{person_info as } PI, \\
   & \text{title as } T \\
   \text{where } & CI\text{-}person_id=CI\text{-}person_id \quad \text{and} \\
   & CI\text{-}role_id=1 \quad \text{and} \\
   & PI\text{-}info_type_id=17 \quad \text{and} \\
   & CI\text{-}movie_id=T\text{-}id; \\
   \end{align*}
   \]

4. \textit{IMDB-4}: The query retrieves titles together with character names and trivia on actors that have played in the movie.

   \[
   \begin{align*}
   \text{select } & T\text{-}title, CN\text{-}name, CI\text{-}info \\
   \text{from } & \text{cast_info as } CI, \\
   & \text{person_info as } PI, \\
   & \text{title as } T, \\
   & \text{char_name as } CN \\
   \text{where } & CI\text{-}person_id=PI\text{-}person_id \quad \text{and} \\
   & CI\text{-}role_id=1 \quad \text{and} \\
   & PI\text{-}info_type_id=17 \quad \text{and} \\
   & CI\text{-}movie_id=T\text{-}id \quad \text{and} \\
   & CI\text{-}person_role_id=CN\text{-}id; \\
   \end{align*}
   \]