Polaris: A System for Query, Analysis, and Visualization of Multidimensional Relational Databases
C. Stolte, D. Tang, P. Hanrahan
IEEE Transactions on Visualization and Computer Graphics

Abstract

In the last several years, large multidimensional databases have become common in a variety of applications such as data warehousing and scientific computing. Analysis and exploration tasks place significant demands on the interfaces to these databases. Because of the size of the data sets, dense graphical representations are more effective for exploration than spreadsheets and charts. Furthermore, because of the exploratory nature of the analysis, it must be possible for the analysts to change visualizations rapidly as they pursue a cycle involving first hypothesis and then experimentation. In this paper we present Polaris, an interface for exploring large multi-dimensional databases that extends the well-known Pivot Table interface. The novel features of Polaris include an interface for constructing visual specifications of table-based graphical displays and the ability to generate a precise set of relational queries from the visual specifications. The visual specifications can be rapidly and incrementally developed, giving the analyst visual feedback as they construct complex queries and visualizations.
1 Concept

Polaris enables interactive exploratory analysis of multidimensional databases. It extends the Pivot Table interface to directly generate a rich, expressive set of graphical displays. Polaris builds tables using an algebraic formalism involving the fields of the database.

2 Implementation

2.1 Graphic Generation

Polaris provides an interface for rapidly and incrementally generating table-based displays. In Polaris, a table consists of a number of rows, columns, and layers. Each table axis may contain multiple nested dimensions. Each table entry, or pane, contains a set of records that are visually encoded as a set of marks to create a graphic.

Polaris’s Visual Specification consists of 3 parts:

1. The specification of the different table configurations
2. The type of graphic inside of each pane
3. The details of the visual encodings

The specification of the table configuration is done using an algebra. A complete table configuration consists of three separate expression in this table algebra. Two of the expressions define the configuration of the $x$ and $y$ axes of the table, partitioning the table into rows and columns. The third expression defines the $z$ axis of the table, which partitions the display into layers. The operands in this table algebra are the names of the ordinal and quantitative fields of the database. The algebra supports the following operators:

**Concatenation** Performs an ordered union of the sets
Cross Performs a Cartesian product of the sets

Nest Similar to the Cross operator but only creates set entries for which there exist records with those domain values.

Polaris allows analysts to combine multiple data sources in a single visualization. When multiple data sources are imported, each data source is mapped to a distinct layer. All layers share the same configuration of tables axes but each data source can have a different expression for partitioning its data into layers.

Polaris aims to simplify the selection of graphic types by categorizing graphic types based on the type of fields assigned to their axes:

**ordinal-ordinal** table

**ordinal-quantitative** bar chart, dot plot, Gantt chart

**quantitative-quantitative** scatter plot, map

The retinal properties shape, size, orientation and color may also be set in Polaris and bound to data values.

### 2.2 Data Transformations

Polaris allows the generated graphics and their driving data to be manipulated/transformed by the user. Four techniques are supported.

1. Deriving Additional Fields
   
   Five methods of deriving additional fields are supported: simple aggregation of quantitative measures, counting of distinct values in ordinal dimensions, discrete partitioning of quantitative measures, ad hoc grouping within ordinal dimensions, and threshold aggregation.

2. Filtering
   
   Allows the user to choose which values to display so more screen real estate can be dedicated to areas of interest

3. Sorting
   
   Allows the user to uncover hidden patterns and trends and to group together interesting values.

4. Brushing
   
   Allows the user to choose a set of interesting data points by drawing a rubber-band around them.

5. Tooltips
   
   Allows the user to view additional details on demand.

6. Undo and Redo
   
   Allows the user to undo and redo an unlimited number of actions.
3 Related Work


Polaris’s graphical specification was inspired by and are extensions of Wilkinson’s efforts to develop a grammar for statistical graphics. Wilkinson’s grammar encapsulates both the statistical transformation of datasets and their mapping to graphic presentations.


4 Data Characteristics

Polaris was designed to work specifically with relational databases. It converts its visualization specification into SQL queries.

5 Notes

Polaris has been commercialized as Tableau Software.