Relational Algebra and SQL

Chapter · January 2002

DOI: 10.1007/0-387-21688-X_3

Peter Revesz
University of Nebraska at Lincoln

3

Relational Algebra and SQL

In this chapter we will learn about two different query languages for relational databases. The first query language, *relational algebra*, has a small set of well-defined operators that can be composed to form query expressions. It is a *procedural* language, because the sequence of operators and the operators themselves can be evaluated using well-defined procedures. Relational algebra is a useful theoretical language that serves to define more complex languages.

The second language we consider, *Structured Query Language* or SQL, is a practical language that allows a high-level expression of queries. A user of SQL does not need to think procedurally about queries, but can rely, to a significant degree, on the meaning of higher-level keywords provided by SQL. However, most SQL queries can be translated to relational algebra queries, and such translation can elucidate the precise meaning of a SQL query.

3.1 Relational Algebra

Relational algebra is a query language composed of a number of operators described in Section 3.1.1. We give several example relational algebra queries in Section 3.1.2 and discuss relational algebra query composition in Section 3.1.3.

3.1.1 Relational Algebra Operators

Relational algebra is a query language composed of a number of operators, each of which takes in relations as arguments and returns a single relation as result. The following are the main operators in relational algebra.

Intersection: The intersection of two relations A and B, denoted $A \cap B$, is the set of points that belong to both A and B. The intersection operator can be applied only to operands that have the same set and order of attributes.

Union: The union of two relations A and B, denoted $A \cup B$, is the set of points that belong to A or B or both. The union operator can be applied only to operands that have the same set and order of attributes.

Difference: The difference of two relations A and B, denoted $A \setminus B$ is the set of points that belong to A but do not belong to B. The difference operator can be applied only to operands that have the same set and order of attributes.

Product: The product operator applied to an n-dimensional relation A and an m-dimensional relation B, denoted $A \times B$, returns a relation that contains all $(n + m)$-dimensional points whose first n components belong to A and last m components belong to B. The product operator can be applied only to operands that have no common attributes.

Project: This operator is used to reorder the columns of a relation or to eliminate some columns of a relation. The project operator from a relation A is denoted $\Pi_L A$, where L is a list $[l_1, \ldots, l_k]$ that specifies an ordering of a subset of the attributes of A. The project operator creates a new relation that contains in its ith column the column of A corresponding to attribute l_i.

Select: The select operator is used to select from a relation A those points that satisfy a certain logical formula F. The select operator has the form $\sigma_F A$. The logical formula F is a constraint formula, as described in Section 2.1.

Rename: The rename operator, $\rho_{B(X_1/Y_1,\ldots,X_n/Y_n)} A$ for any $n \geq 1$, changes the name of relation A to B and changes the attribute X_i to Y_i. If we only want to change the name of the relation, then the form $\rho_B A$ can be used.

Natural Join: The natural join operator applied to an n-dimensional relation A and an m-dimensional relation B that have k attributes in common is denoted $A \bowtie B$. The natural join operator returns a relation that contains all $(n + m - k)$-dimensional points whose projection onto the attributes of A belong to A and whose projection onto the attributes of B belong to B.

Example 3.1.1 We illustrate these relational algebra operators on spatial relations. Consider the set of points $A(x, y)$ and $B(x, y)$ shown in Figure 3.1. They both have dimension two, hence the operators of intersection, union, and difference can be applied.