Subquery Elimination: A Complete Unnesting Algorithm
For An Extended Relational Algebra

Pedro Celis, Hans Zeller
Tandem (celis_pedro@tandem.com, zeller_hansjorg@tandem.com)

Research in the area of subquery unnesting algorithms [1,2,3,4,5,6] has mostly focused on the problem of making queries more efficient at run-time by transforming subqueries into joins. Unnesting rules describe a transformation of a nested query tree or a nested SQL query into an equivalent tree or SQL query that is no longer nested. However, it is not possible to express all nested queries in a non-nested form, unless the used language (relational algebra or ISO/ANSI SQL) is extended. This means that a database system must continue to have the ability to process subqueries.

When working on a new optimizer and executor design for NonStop SQL, our development team was faced with a slightly different problem: we wanted to eliminate the need for optimization and execution of nested queries altogether and were looking for a complete subquery unnesting process. Such a process would allow us to develop a query optimizer and executor that do not need to process subqueries.

Our goal was to make use of the existing unnesting algorithms and to extend them in a way that does not necessarily improve or change the execution characteristics of nested queries, but that leads to complete unnesting of all forms of nested queries, as defined by the “full” level of the ISO/ANSI SQL92 standard. To indicate this different approach we will call it “subquery elimination” rather than “subquery unnesting”