

Relational Calculus

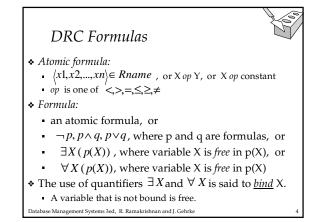


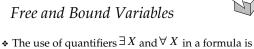
- Comes in two flavors: <u>Tuple relational calculus</u> (TRC) and <u>Domain relational calculus</u> (DRC).
- Calculus has variables, constants, comparison ops, logical connectives and quantifiers.
 - <u>*TRC*</u>: Variables range over (i.e., get bound to) *tuples*.
 - <u>DRC</u>: Variables range over *domain elements* (= field values).
 Both TRC and DRC are simple subsets of first-order logic.
- Expressions in the calculus are called *formulas*. An answer tuple is essentially an assignment of constants to variables that make the formula evaluate to *true*.

Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke

Domain Relational Calculus \diamond Query has the form: $[\langle x1, x2, ..., xn \rangle | p[\langle x1, x2, ..., xn \rangle]]$ \diamond Answer includes all tuples $\langle x1, x2, ..., xn \rangle$ that make the formula $p[\langle x1, x2, ..., xn \rangle]$ be true. \diamond Formula is recursively defined, starting with simple atomic formulas (getting tuples from relations or making comparisons of values), and building bigger and better formulas using the logical connectives.

Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke





- said to <u>bind</u> X.
 A variable that is not bound is <u>free</u>.
- Let us revisit the definition of a query:

Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke

$$|\langle x1, x2, \dots, xn \rangle| p[\langle x1, x2, \dots, xn \rangle]$$

There is an important restriction: the variables x1, ..., xn that appear to the left of `|' must be the *only* free variables in the formula p(...).

Find all sailors with a rating above \mathcal{F} $|\langle I,N,T,A \rangle| \langle I,N,T,A \rangle \in Sailors \land T > 7 \rangle$

- ◆ The condition $\langle I, N, T, A \rangle \in Sailors$ ensures that the domain variables *I*, *N*, *T* and *A* are bound to fields of the same Sailors tuple.
- ◆ The term $\langle I, N, T, A \rangle$ to the left of `|' (which should be read as *such that*) says that every tuple $\langle I, N, T, A \rangle$ that satisfies *T*>7 is in the answer.
- * Modify this query to answer:
- Find sailors who are older than 18 or have a rating under 9, and are called 'Joe'.
 Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke



 $|\langle I, N, T, A \rangle| \langle I, N, T, A \rangle \in Sailors \land T > 7 \land$

Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke

- $\exists Ir, Br, D ((Ir, Br, D) \in \operatorname{Re} serves \land Ir = I \land Br = 103)$
- ★ We have used $\exists Ir, Br, D(...)$ as a shorthand for $\exists Ir(\exists Br(\exists D(...)))$
- \clubsuit Note the use of $\ \exists$ to find a tuple in Reserves that `joins with' the Sailors tuple under consideration.

Find sailors rated > 7 who've reserved a red boar $\langle (I,N,T,A) | \langle I,N,T,A \rangle \in Sailors \land T > 7 \land$ $\exists Ir, Br, D \langle (Ir, Br, D) \in \text{Reserves} \land Ir = I \land$ $\exists B, BN, C \langle (B, BN, C) \in Boats \land B = Br \land C = 'red') \rangle$ \diamond Observe how the parentheses control the scope of each quantifier's binding.

 This may look cumbersome, but with a good user interface, it is very intuitive. (MS Access, QBE)

Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke

Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke

Find sailors who've reserved all boats $[\langle I,N,T,A \rangle | \langle I,N,T,A \rangle \in Sailors \land$ $\forall B,BN,C [\neg [\langle B,BN,C \rangle \in Boats] \lor$ $(\exists Ir,Br,D [\langle Ir,Br,D \rangle \in \text{Reserves} \land I = Ir \land Br = B]]]$ * Find all sailors I such that for each 3-tuple $\langle B,BN,C \rangle$ either it is not a tuple in Boats or there is a tuple in Reserves showing that sailor I has reserved it. Find sailors who've reserved all boats (again) $\{ \langle I, N, T, A \rangle | \langle I, N, T, A \rangle \in Sailors \land \land \land \langle B, BN, C \rangle \in Boats \\ [\exists \langle Ir, Br, D \rangle \in \text{Reserves}[I = Ir \land Br = B]] \}$ * Simpler notation, same query. (Much clearer!) * To find sailors who've reserved all red boats: $\{C \neq 'red' \lor \exists \langle Ir, Br, D \rangle \in \text{Reserves}[I = Ir \land Br = B]] \}$

Unsafe Queries, Expressive Power

 It is possible to write syntactically correct calculus queries that have an infinite number of answers! Such queries are called <u>unsafe</u>.

e.g.,
$$|S| \neg (S \in Sailors)$$

Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke

- It is known that every query that can be expressed in relational algebra can be expressed as a safe query in DRC / TRC; the converse is also true.
- Relational Completeness: Query language (e.g., SQL) can express every query that is expressible in relational algebra/calculus. tabase Management Systems 3ed, R. Ramakrishnan and J. Gehrke

Summary



- Relational calculus is non-operational, and users define queries in terms of what they want, not in terms of how to compute it. (Declarativeness.)
- Algebra and safe calculus have same expressive power, leading to the notion of relational completeness.

Database Management Systems 3ed, R. Ramakrishnan and J. Gehrke