Univ of Helsinki
23.2.2004
Introduction to Databases, s 2004
Relational algebra -revision

Relational algebra

- Union
- Difference
- Cross product
- Intersection
- Projection
- Selection
- Joins
  - Normal join (theta join, equijoin)
  - Natural join
  - Outer join

Cardinality of results: R op S

- $c(R)$= cardinality (number of rows) of R
- Union: rows of both relations, common ones only once
  - $c(R \cup S) = c(R) + c(S) - c(R \cap S) \geq max(c(R), c(S))$
- Difference: Rows of R that are not in S
  - $c(R-S) = c(R) - c(R \cap S) \leq c(R)$
- Intersection: Rows common to R and S
  - $0 \leq c(R \cap S) \leq min(c(R), c(S))$

Cardinality of results: R op S

- Product: each row of R connected to each row of S
  - $c(R \times S) = c(R) * c(S)$
- Projection: pick up distinct values appearing in a column collection
  - $c(\pi_a (R)) \leq c(R)$
  - $c(\pi_a (R)) \leq c(\pi_{a,b} (R))$
  - if a contains a key: $c(\pi_a (R)) = c(R)$
  - $c(\pi_a (R)) = 0$ only if $c(R)=0$

Cardinality of results: R op S

- Selection: pick up rows that conform to the selection criteria
  - $0 \leq c(\sigma_e (R))$
  - $0 \leq c(\sigma_{k=constant} (R)) \leq 1$, when k is a key of R
  - $c(R)-1 \leq c(\sigma_{k \neq constant} (R)) \leq c(R)$, when k is a key of R
  - $c(\sigma_{e \text{ and } f} (R)) \leq min(c(\sigma_e (R)), c(\sigma_f (R)))$
  - $c(\sigma_{e \text{ or } f} (R)) \geq max(c(\sigma_e (R)), c(\sigma_f (R)))$

Cardinality of results: R op S

- Join: row of R and S are connected based on connection criteriai
- $0 \leq c(R \bowtie_e S) \leq c(R \times S)$

- Let k be the key of R and v a foreing key in S that refers to R
- $c(R \bowtie_{k=v} S) \leq c(S)$, if null values are allowed for v
- $c(R \bowtie_{k=v} S) = c(S)$, if null values are not allowed for v

Cardinality of results: R op S

- $c(R \bowtie_{k \neq v} S) = c(R \times S) - c(S)$, if null values are not allowed for v

Natural join:

- $c(R * S) = c(R \times S)$, if R and S do not have columns with common name
- mostly natural joins carry out a join based on the equality of foreing key and the corresponding primary key
- If both R and S have a common schema, natural join equals to intersection

Cardinality of results: R op S

- outer join: Rows that drop out in normal join will also be included with 'empty' pair
- $c(R \rhd\!\!\lhd_{a=b} S) \geq c(R)$