Database Systems Notes

1 General Database Patter

Relations ↔ Tables

1.1 Relational Algebra

Relation algebra can express the same statements as SQL `SELECT-FROM-WHERE` statements.

π	Projection
σ	Selection
$\times$	(*Cartesian/Cross*) Product
ρ	Renaming
$\bigcup$	Union
$\bigcap$	Intersection
$\setminus$	Difference
-	Difference
$\bowtie$	Natural Join

1.1.1 π: Projection

Projection is a *vertical* operation, allowing you to choose some *columns.*

Syntax:

$\pi_{\text{Set of Attributes}}(\text{relation})$

Any set of attributes not mentioned by the projection are discarded.

1.1.2 σ: Selection

Selection is a *horizontal* operation, allowing you to choose *rows that satisfy a condition.*

Syntax:

$\sigma_{\text{Condition}}(\text{Relation})$

This provides users with a *view* of data, hiding rows that do not satisfy the condition.

Consecutive selections are the same as a conjunction of the conditions in one selection; however, they can bring around different levels of performance. Consecutive selections are performed sequentially and eliminate rows that do not meet the criteria, whereas the conjunction means that the selection is performed once with a stricter condition.

1.1.3 $\times$: Cartesian Product

Where each row of two relations is *concatenated* to produce a new relation.

Syntax:

A Relation $\times$ Another Relation

A Cartesian product is the product of the two relations' sizes; meaning that Cartesian products can balloon in size.

1.1.4 ρ: Renaming

A useful tool that *aids* Cartesian products where the two relations have columns that go by the same name.

For example, a bank may have a table for all of a customer's accounts, and another table for all accounts open at some branch. These two tables could have a selection performed upon them to find all of a customer's accounts across all branches, but they both have an account name column. Renaming one allows for them to be distinguishable.

1.1.5 $\bowtie$: Natural Join

Joining two tables on *common attributes.*

This however can be expressed using a combination of projection, selection, renaming and the Cartesian product. It is more efficient to use a natural join however, similar to sequential selections.

2 SQL

All queries have the general format:

```
SELECT list, of, attributes
FROM list, of, relations
WHERE conditions
```

WHERE statements allow for tables to be **joined**.

2.1 Union Compatibility

Sometimes straightforward queries in English are a little more complicated than they are in relational algebra and in SQL.