

Select

Prof. Hasso Plattner

Select

- Fundamental operation of Codd's relational algebra on one table
- The traditional selection (restriction) is a unary operation written as
 - a and b are expressions:
attribute names, constant values
 - θ is a binary operation ($<, >, =, \dots$)
 - R is the relation
- DBMS allow more operations and concatenation of operations

$$\sigma_{a\theta b}(R)$$

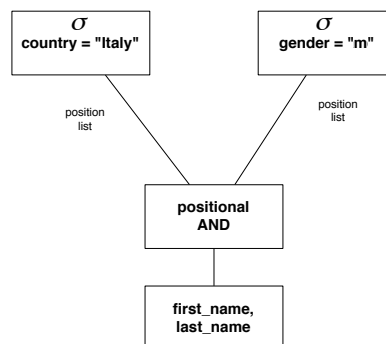
SELECT Example

```
SELECT first_name, last_name FROM world_population  
WHERE country="Italy" and gender="m"
```

id	fname	lname	country	gender
2394	Gianluigi	Buffon	Italy	m
3010	Lena	Gercke	Germany	f
3040	Mario	Balotelli	Italy	m
3949	Manuel	Neuer	Germany	m
4902	Lukas	Podolski	Germany	m
20102	Klaas-Jan	Huntelaar	Netherlands	m

3

Query Plan



- ❑ Predicates are evaluated and generate position lists
- ❑ Intermediate position lists are logically combined
- ❑ Final position list is used for materialization

4

Query Execution

Value ID	Dictionary for "country"
0	Algeria
1	France
2	Germany
3	Italy
4	Netherlands
...	

id	fname	lname	country	gender
2394	Gianluigi	Buffon	3	1
3010	Lena	Gercke	2	0
3040	Mario	Balotelli	3	1
3949	Manuel	Neuer	2	1
4902	Lukas	Podolski	2	1
20102	Klaas-Jan	Huntelaar	4	1

Value ID	Dictionary for "gender"
0	f
1	m

gender = 1 ("m")

country = 3 ("Italy")

Position
0
2

AND

Position
0
2
3
4
5

Position
0
2

5

Observations

- Leverage sequential scan over large data region
- Comparison is based on encoded values from the dictionary, dictionary lookup is only executed once
- In main memory systems, branching can become expensive

6