

Generelle Semantik

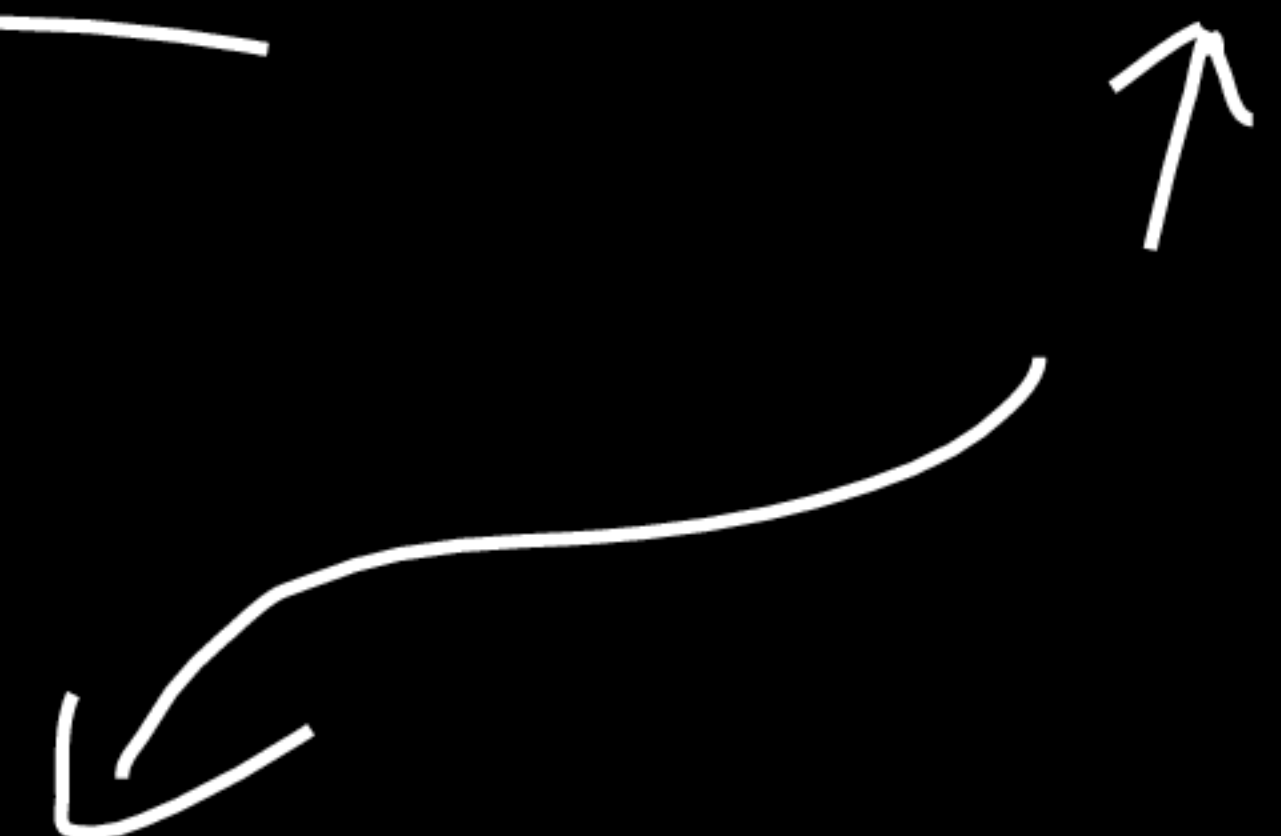
map(record) -> set of (ikey, ivalue)

42

42, 3

42, 4

42, 5



reduce(ikey, set of ivalue) -> output

↑
42, {3, 4, 5}

Beispiel: Web-Index Erstellung

map((docID, document)) -> set of (term, docID)

↑
pk

↑
html-string

↑
Word

↑
pk

reduce(term, set of docID) -> (term, (posting list of docID, count))

```
map( 44,  
     "This is text on a website!")  
)
```

```
->
```

```
{
```

```
  ("This", 44),
```

```
  ("is", 44),
```

```
  ("text", 44),
```

```
  ("on", 44),
```

```
  ("a", 44),
```

```
  ("website", 44)
```

```
}
```

```
map( 42,  
    "This is just another website!")
```

```
)
```

```
->
```

```
{
```

```
  ("This", 42),
```

```
  ("is", 42),
```

```
  ("just", 42),
```

```
  ("another", 42),
```

```
  ("website", 42)
```

```
}
```

```
map( 43,  
     "One more boring website!")
```

```
)
```

```
->
```

```
{
```

```
  ("One", 43),
```

```
  ("more", 43),
```

```
  ("boring", 43),
```

```
  ("website", 43)
```

```
}
```

```
reduce("This",  
      {42,  
       43}  
)
```

->

```
("This", ([42, 43], 2))
```



```
reduce("is",  
      {42,  
       43}  
)
```

->

```
("is", ([42, 43], 2))
```

```
reduce("boring",  
      {43}  
)
```

->

```
("boring", ([43], 1))
```

etc.

Beispiel: Gruppierung mit Aggregation

```
SELECT    erfahrung, count(*), max(gehalt)
FROM      mitarbeiter
WHERE     gehalt < 50000
GROUP BY  erfahrung
```

map() und reduce()

Mit z.B. `getSalary`

`map(record)` -> set of (ikey, ivalue):

if (`getSalary < 50000`)

`emit (erfahrung, getSalary);`

ikey

ivalue

`reduce(ikey, set of ivalue)` -> output:

`Ai` | `int maxSalary = -42;`
| For each value in ivalue:
| `maxSalary = max(maxSalary, value);`
| `emit (ikey, set.size(), maxSalary);`

Beispiel: HAVING

SELECT erfahrung, count(*), max(gehalt)

FROM mitarbeiter

WHERE gehalt < 50000

GROUP BY erfahrung

HAVING count(*) > 2

map() und reduce()

map(record) -> set of (ikey, ivalue):

wie oben

reduce(ikey, set of ivalue) -> output:

if (set.size() > 2)

A