Relational Algebra: Basic Operators

- **Projection (π):**

  $\pi_{id,\ name}(Students)$ → SELECT id, name FROM <SQL for Students>

- **Selection (σ):**

  $\sigma_{id\ =\ 100}(Students)$ → SELECT * FROM <SQL for Students> WHERE id = 100

- **Join (⋈):**

  Students $\bowtie_{Student.id\ =\ Grades.id}$ Grades →

  SELECT * FROM <SQL for Students>, <SQL for Grades>
  WHERE Student.id = Grades.id

Relational Algebra: More operators

- **Set operations (∪, ∩, −):** Same as in SQL (UNION, INTERSECT, MINUS)
- **Renaming (ρ):**

  $\rho_{S(i,n)}$(Students) → (SELECT id AS i, name AS n FROM Students) AS S
- **Removing duplicates (δ):**

  δ(R) → SELECT DISTINCT * FROM <SQL for R>
- **Grouping/aggregation (γ):**

  $\gamma_{\text{id, AVG(grade)} \rightarrow \text{average}}$(Grades) →

  SELECT id, AVG(grade) AS average FROM <SQL for Grades>
  GROUP BY id

Relational Algebra: Sanity check

- Every condition, projection, etc. should **ONLY** mention attributes that exists in their operands.

```
Students(idnr, name)
Grades(student, course, grade)
   student -> Students.idnr
```

“Select the students with at least two passed courses with a grade of at least 3”

$$\pi_{\text{name}}\left(\sigma_{\text{passed>=2 AND idnr=student AND grade>=3}}\left(\text{Students} \bowtie \gamma_{\text{student, COUNT(*)} \rightarrow \text{passed}}(\text{Grades})\right)\right)$$

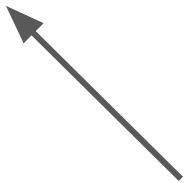
Relational Algebra: Sanity check

- Every condition, projection, etc. should **ONLY** mention attributes that exists in their operands.

```
Students(idnr, name)
Grades(student, course, grade)
    student -> Students.idnr
```

"Select the students with at least two passed courses with a grade of at least 3"

$$\pi_{\text{name}}\left(\sigma_{\text{passed>=2 AND idnr=student AND grade>=3}}\left(\text{Students} \bowtie \gamma_{\text{student, COUNT(*)} \rightarrow \text{passed}}(\text{Grades})\right)\right)$$



Can not use grade here!

Relational Algebra: Sanity check

- Every condition, projection, etc. should **ONLY** mention attributes that exists in their operands.

```
Students(idnr, name)
Grades(student, course, grade)
    student -> Students.idnr
```

"Select the students with at least two passed courses with a grade of at least 3"

$$\pi_{\text{name}}\left(\sigma_{\text{passed}>=2 \text{ AND } \text{idnr}=\text{student}}\left(\text{Students} \bowtie \gamma_{\text{student, COUNT}(*) \rightarrow \text{passed}}\left(\sigma_{\text{grade}>=3}(\text{Grades})\right)\right)\right)$$

We need to filter by grade before the aggregating our data!