"For All" Queries (1)

• Given the database schema:

```
Student(Sid,Sname,Addr)
Course(Cid,Cname,Credits)
Enrolled (Sid,Cid)
```

Consider the following query:

"Find the students who are enrolled in **all** courses."

Solution in RA?

```
\pi_{Sid} (Enrolled)
```

- No, this returns the students enrolled in some courses.
- Any ideas about what is wrong?

"For All" Queries (2)

- A solution strategy would be to:
 - start with all students mentioned in the Enrolled relation (all guys), from which we then subtract those not enrolled in all courses (bad guys)
- That is, to find all the "good guys", we need to determine "all guys" and "bad guys", i.e.,

Answer = Good guys = All guys - Bad guys

"For All" Queries (3)

Set of all students we need to consider:

```
A \leftarrow \pi_{Sid} (Enrolled)
```

- Set of all students not enrolled in all courses
 - 1. Create all possible "student-course" pairs:

```
\pi_{Sid} (Enrolled) \times \pi_{Cid} (Course)
```

- 2. Create all "actual" pairs made of one student and one course: which is **Enrolled**
- 3. Students who are not enrolled in all courses: $B \leftarrow \pi_{Sid}(\pi_{Sid}(Enrolled) \times \pi_{Cid}(Course) Enrolled)$
- The query: A B

The Division Operation (+)

- The previous query can be conveniently expressed in RA using the division operator
 - Divide Enrolled by π_{Cid} (Course)
 - Here, Enrolled $\div \pi_{Cid}$ (Course)
 - Schema of the result is {Sid,Cid} {Cid}
- r ÷ s requires that all attributes of S to be a subset of attributes of R.
 - − The schema of the result would be R − S

Example: Enrolled(student,sport)



Jim	Hockey
Jim	Soccer
Joe	Hockey
Joe	Soccer
Sue	Hockey
Sue	Soccer

Jim	Hockey
Joe	Soccer
Jim	Soccer
Sue	Hockey

Joe	Hockey
Sue	Soccer

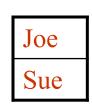
 $\pi_{student}$ (Enrolled) —

 $\pi_{student}$ (Enrolled) $\times \pi_{sport}$ (Enrolled) – Enrolled)

All

Bad

Jim Joe Sue





Jim is the only student enrolled in all sports

Another Example

- Given:
 - Customer(<u>id</u>, name)
 - Branch(bid, district)
 - Account(<u>cid</u>, <u>bid</u>)
- Query: "Find the names of all customers who have an account in every branch located in Westmount area"
- Solution?
 - π_{name} (Customer $\triangleright \triangleleft$ Account $\triangleright \triangleleft$ ($\sigma_{district = "Westmount"}$ Branch)) ?
- No, this returns the names of all customers who have an account at some branch in Westmount, not all such brances.

Solution

Customer(cid, name), Branch(bid, district), Account(cid, bid)

- We can apply the division operator
 - Find all pairs (cid, bid) for which the customer has an account at a branch $\pi_{cid,bid}$ (customer $\triangleright \triangleleft$ account)
 - Divide it by all bid's of branches in Westmount π_{bid} ($\sigma_{district} = "Westmount"$ branch))
- π_{name} ((customer $\triangleright \triangleleft$ account) ÷ π_{bid} ($\sigma_{district = "Westmount"}$ branch))
- So, we define $r \div s$ as follows:

$$r \div s = \pi_{R-S}(r) - \pi_{R-S}((\pi_{R-S}(r) \times s) - \pi_{R-S,S}(r))$$