

CSE 344 Introduction to Data Management

Review: RC, Datalog

TA: Yi-Shu Wei

Review sections

Topic	Date	Venue/Time	Who
RC/RA/Datalog	3/6 (Th) Sections	Sections	Yi-Shu
BCNF/ ERD	3/10 (M)	CSE 303, 4:30-5:30pm	Vaspol
Transaction	3/13 (Th) Sections	Sections	Yi-Shu
Parallel DB/MR	3/14 (F) Class	Class	Sudeepa

Problem

- The MOOC database
- Schemata:
 - Student(sid, name, univ)
 - Course(cid, name, year, duration)
 - Enrollment(sid, cid)
 - Instructor(tid, name, univ)
 - Teaches(tid, cid)
- Assume that more than one instructor can teach a class

- 1) Find the names of the students who took some course taught by at least one professor from UW in year 2012
- 2) Find the names of the students who *only* took courses taught by at least one professor from UW in year 2012
- 3) Find the names of the students who *only* took courses that are *only* taught by professors from UW in year 2012

If every course is taught by exactly one professor (i.e. cid is the *primary key* for Teaches), then 2) and 3) are the same!

Steps

- Translate English into logical forms (RC)
- Translate RC into Datalog^{NEG}
- Translate Datalog into SQL/Relational Algebra

- 1) Find the names of the students who took some course taught by at least one professor from UW in year 2012

$$\begin{aligned} \text{Ans}(n) = & \exists \text{sid} \exists \text{suniv} \exists \text{cid} \exists n2 \exists d \exists \text{tid} \exists n3 \\ & \text{Student}(\text{sid}, n, \text{suniv}) \wedge \text{Enrollment}(\text{sid}, \text{cid}) \wedge \text{Course}(\text{cid}, n2, 2012, d) \\ & \wedge \text{Teaches}(\text{tid}, \text{cid}) \wedge \text{Instructor}(\text{tid}, n3, \text{'UW'}) \end{aligned}$$

- 2) Find the names of the students who *only* took courses taught by at least one professor from UW in year 2012

Find student x such that all courses x took in 2012 are taught by some UW professor

$$\begin{aligned} \text{Ans}(n) = & \exists \text{sid} \exists \text{suniv} [\text{Student}(\text{sid}, n, \text{suniv}) \wedge \\ & (\forall \text{cid} \forall n2 \forall d \text{Enrollment}(\text{sid}, \text{cid}) \wedge \text{Course}(\text{cid}, n2, 2012, d) \\ & \Rightarrow \exists \text{tid} \exists n3 \text{Teaches}(\text{tid}, \text{cid}) \wedge \text{Instructor}(\text{tid}, n3, \text{'UW'}))] \end{aligned}$$

- 3) Find the names of the students who *only* took courses that are *only* taught by professors from UW in year 2012

Find student x such that all instructors of all courses x took in 2012 are from UW

$$\begin{aligned} \text{Ans}(n) = & \exists \text{sid} \exists \text{suniv} [\text{Student}(\text{sid}, n, \text{suniv}) \wedge \\ & (\forall \text{cid} \forall n2 \forall d \forall \text{tid} \text{Enrollment}(\text{sid}, \text{cid}) \wedge \text{Course}(\text{cid}, n2, 2012, d) \\ & \wedge \text{Teaches}(\text{tid}, \text{cid}) \Rightarrow \exists n3 \text{Instructor}(\text{tid}, n3, \text{'UW'}))] \end{aligned}$$

- 1) Find the names of the students who took some course taught by at least one professor from UW in year 2012

Ans(n) :- Student(sid, n, _), Enrollment(sid, cid), Course(cid, _, 2012, _),
Teaches(tid, cid), Instructor(tid, _, 'UW')

- 2) Find the names of the students who *only* took courses taught by at least one professor from UW in year 2012

UWCourse(cid) :- Teaches(tid, cid), Instructor(tid, _, 'UW')
NonAns(sid) :- Enrollment(sid, cid), Course(cid, _, 2012, _), NOT UWCourse(cid)
Ans(n) :- Student(sid, n, _), NOT NonAns(sid)

- 3) Find the names of the students who *only* took courses that are *only* taught by professors from UW in year 2012

UWInstructor(tid) :- Instructor(tid, _, 'UW')
NonAns(sid) :- Enrollment(sid, cid), Course(cid, _, 2012, _), Teaches(tid, cid),
NOT UWInstructor(tid)
Ans(n) :- Student(sid, n, _), NOT NonAns(sid)

Datalog queries should be SAFE!

- 1) Find the names of the students who took some course taught by at least one professor from UW in year 2012

```
SELECT Student.name
FROM Student, Enrollment, Course, Teaches, Instructor
WHERE Student.sid=Enrollment.sid AND Enrollment.cid=Course.cid
AND Course.cid=Teaches.cid AND Teaches.tid=Instructor.tid
AND Course.year=2012 AND Instructor.univ='UW'
```

- 2) Find the names of the students who *only* took courses taught by at least one professor from UW in year 2012

```
SELECT name FROM Student WHERE sid NOT IN
(SELECT sid FROM Enrollment, Course
WHERE Enrollment.cid=Course.cid AND Course.year=2012
AND Course.cid NOT IN
(SELECT cid FROM Teaches, Instructor
WHERE Teaches.tid=Instructor.tid AND Instructor.univ='UW'))
```

- 3) Find the names of the students who *only* took courses that are *only* taught by professors from UW in year 2012

```
SELECT name FROM Student WHERE sid NOT IN
(SELECT sid FROM Enrollment, Course, Teaches
WHERE Enrollment.cid=Course.cid AND Course.cid=Teaches.cid
AND Course.year=2012 AND Teaches.tid NOT IN
(SELECT tid FROM Instructor WHERE univ='UW'))
```