CS 6016

Database Systems

Adv. Queries II

This week

Advanced Queries

- Teamwork
 - Project Phase 2 / Phase 3 / Phase 4

Ponder

•Find all people with the same phone number

Phones

CardNum	Phone
1	555-5555
2	666-6666
2	555-5555
3	777-7777
4	888-888
4	999-9999
5	777-7777

•Find all people with the same phone number

Phones

CardNum	Phone
ī	555-5555
2	666-6666
2	555-5555
3	777-7777
4	888-8888
5	777-7777

CardNum	Phone	CardNum	Phone
1	555-5555	1	555-5555
1	555-5555	2	666-6666
1	555-5555	2	555-5555
1	555-5555	3	777-7777
1	555-5555	4	888-888
1	555-5555	5	777-7777
2	666-6666	1	555-5555
2	666-6666	2	666-6666
2	666-6666	2	555-5555
2	666-6666	3	777-7777
2	666-6666	4	888-888
•••	•••	•••	•••

•Find all people with the same phone number

Phones

CardNum	Phone
1	555-5555
2	666-6666
2	555-5555
3	777-7777
4	888-888
5	777-7777

CardNum	Phone	CardNum	Phone
1	555-5555	1	555-5555
1	555-5555	2	666-6666
1	555-5555	2	555-5555
1	555-5555	3	777-7777
1	555-5555	4	888-888
1	555-5555	5	777-7777
2	666-6666	1	555-5555
2	666-6666	2	666-6666
2	666-6666	2	555-5555
2	666-6666	3	777-7777
2	666-6666	4	888-888
•••	•••	•••	•••

•Find all people with the same phone number

Phones

CardNum	Phone
1	555-5555
2	666-6666
2	555-5555
3	777-7777
4	888-8888
5	777-7777

CardNum	Phone	CardNum	Phone
1	555-5555	1	555-5555
1	555-5555	2	666-6666
1	555-5555	2	555-5555
1	555-5555	3	777-7777
1	555-5555	4	888-888
1	555-5555	5	777-7777
2	666-6666	1	555-5555
2	666-6666	2	666-6666
2	666-6666	2	555-5555
2	666-6666	3	777-7777
2	666-6666	4	888-888
•••	•••	•••	•••

•Find all people with the same phone number

Phones

CardNum	Phone
1	555-5555
2	666-6666
2	555-5555
3	777-7777
4	888-8888
5	777-7777

CardNum	Phone	CardNum	Phone
1	555-5555	1	555-5555
1	555-5555	2	666-6666
1	555-5555	2	555-5555
1	555-5555	3	777-7777
1	555-5555	4	888-888
1	555-5555	5	777-7777
2	666-6666	1	555-5555
2	666-6666	2	666-6666
2	666-6666	2	555-5555
2	666-6666	3	777-7777
2	666-6666	4	888-888
•••	•••	•••	•••

•First we have to disambiguate

- •p(P1, Phones)
- •p(P2, Phones)

 $P1 \times P2$

P1.CardNum	P1.Phone	P2.CardNum	P2.Phone
1	555-5555	1	555-5555
1	555-5555	2	666-6666
1	555-5555	2	555-5555
1	555-5555	3	777-7777
1	555-5555	4	888-888
•••	•••	•••	•••

•Then filter by matching phone number

•p(P1, Phones)

•p(P2, Phones)

$$P1 \times P2$$

P1.CardNum	P1.Phone	P2.CardNum	P2.Phone
1	555-5555	1	555-5555
1	555-5555	2	666-6666
1	555-5555	2	555-5555
1	555-5555	3	777-7777
1	555-5555	4	888-888
•••	•••	•••	•••

 $\bullet \sigma_{P1.Phone=P2.Phone}(P1 \times P2)$

•Then filter by *not* matching card number

•p(P1, Phones)

•p(P2, Phones)

$$P1 \times P2$$

P1.CardNum	P1.Phone	P2.CardNum	P2.Phone
1	555-5555	1	555-5555
1	555-5555	2	666-6666
1	555-5555	2	555-5555
1	555-5555	3	777-7777
1	555-5555	4	888-888
•••	•••	•••	•••

• □ P1.Phone=P2.Phone && P1.CardNum!= P2.CardNum(P1 × P2)

SQL Self-Join

```
select p1.CardNum, p2.CardNum
from Phones p1 join Phones p2
where p1.Phone = p2.Phone
and p1.CardNum != p2.CardNum;
```

• Renaming (aka "range variables") required for self-join

Join

•Default join is called an inner join

•x JOIN y WHERE ...

• Gives rows where condition is true

•Equivalent:

• x INNER JOIN y

• x, y

Outer Join

- •Two types of outer join: LEFT and RIGHT
- •ON clause required!
- •x LEFT JOIN y ON condition
 - Gives all rows where condition is true
 - And gives all rows from x

Outer Join

- •Two types of outer join: LEFT and RIGHT
- •ON clause required!
- •x LEFT JOIN y ON condition
 - Gives all rows where condition is true
 - And gives all rows from x
- •x RIGHT JOIN y ON condition
 - Gives all rows where condition is true
 - And gives all rows from y

Left Join

Patrons

Name	CardNum
Joe	1
Ann	2
Ben	3
Dan	4

CheckedOut

CardNum	Serial
1	1001
1	1004
4	1005

Left Join

Patrons

Name	CardNum
Joe	1
Ann	2
Ben	3
Dan	4

CheckedOut

CardNum	Serial
1	1001
1	1004
4	1005

Patrons p LEFT JOIN CheckedOut c ON p.CardNum = c.CardNum;

Name	CardNum	CardNum	Serial
Joe	1	1	1001
Joe	1	1	1004
Ann	2	NULL	NULL
Ben	3	NULL	NULL
Dan	4	4	1005

Left Join

\mathbf{D}		tr	_		_
Р	\mathbf{a}	r	1)	n	C.
1	ч	u	v	11	. •

C1		<u> </u>	74
L.r	ieck	eat	Jut

Name	CardNum
Joe	1
Ann	2
Ben	3
Dan	4

CardNum	Serial
1	1001
1	1004
4	1005

Patrons p LEFT JOIN CheckedOut c ON p.CardNum = c.CardNum;

Name	CardNum	CardNum	Serial
Joe	1	1	1001
Joe	1	1	1004
Ann	2	NULL	NULL
Ben	3	NULL	NULL
Dan	4	4	1005

•Unmatched tuples get NULL in right-side columns

Shortcut

Patrons NATURAL LEFT JOIN CheckedOut;

Name	CardNum	Serial
Joe	1	1001
Joe	1	1004
Ann	2	NULL
Ben	3	NULL
Dan	4	1005

Shortcut

Patrons NATURAL LEFT JOIN CheckedOut;

Name	CardNum	Serial
Joe	Ī	1001
Joe	1	1004
Ann	2	NULL
Ben	3	NULL
Dan	4	1005

Only one copy of natural column(s)

- •NULL is a special value in SQL
 - It does not have the reflexive property
 - i.e. NULL != NULL

- •NULL is a special value in SQL
 - It does not have the reflexive property
 - i.e. NULL != NULL

```
•WHERE CardNum != NULL // wrong
```

•WHERE CardNum IS NOT NULL // right

•Boolean operators on NULL always return NULL

NULL == 5

NULL != 5

 $NULL == NULL \rightarrow NULL$

• NULL != NULL \rightarrow NULL

→ NULL

 \rightarrow NULL

•Boolean operators on NULL always return NULL

```
    NULL == 5
    NULL
    NULL != 5
    NULL
    NULL == NULL
    NULL != NULL
    NULL
```

•If NULL is used where a Boolean is expected, interpreted as FALSE

Quiz

•Find Names of Patrons who have not checked out a book. Use outer join.

CheckedOut

CardNum	Serial
1	1001
1	1004
4	1005
4	1006

Patrons

Name	CardNum
Joe	1
Ann	2
Ben	3
Dan	4

Solution

•Find Names of Patrons who have not checked out a book Select Name from Patrons **natural left join** CheckedOut where CheckedOut.Serial is null

Or

Select Name from Patrons left join CheckedOut on Patrons.CardNum=CheckedOut.CardNum where CheckedOut.Serial is null

CardNum	Serial
1	1001
1	1004
4	1005
4	1006

Name	CardNum
Joe	1
Ann	2
Ben	3
Dan	4

Nested Query as Condition

•Filter by nested query

```
SELECT x FROM y WHERE x IN (SELECT ...);

Condition
```

Nested Query as Condition

- •There are several of these operators
- •x is a value, A is a multi-set (e.g. from SELECT)

x IN A

EXISTS A

x OP y ANY A

x OP y ALL A

Nested Query as Condition

- •There are several of these operators
- •x is a value, A is a multi-set (e.g. from SELECT)

x IN A: True if x is in A.

EXISTS A: True if A is not empty.

x OP y ANY A: True if there exists a y in A such that x op y is True.

x OP y ALL A: True if for ALL y in A, x op y is True.

Find all students **not** enrolled in Databases

Students

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

Enrolled

sID	cID	Grd
1	3500	Α
1	3810	A-
1	5530	Α
2	3810	Α
2	5530	В
3	3500	С
3	3810	В
4	3500	С

Courses

cID	Name
3500	SW Practice
3810	Architecture
5530	Databases

•Find all students not enrolled in 'Databases'

```
select s.Name from Students s
where s.sID not in
(select e.sID from Enrolled e
natural join Courses c
where c.Name='Databases');
```

Find all students younger than everyone taking Databases

Students

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

Enrolled

sID	cID	Grd
1	3500	Α
1	3810	A-
1	5530	Α
2	3810	Α
2	5530	В
3	3500	С
3	3810	В
4	3500	С

Courses

cID	Name
3500	SW Practice
3810	Architecture
5530	Databases

•Find all students younger than everyone in 'Databases'

```
select s.sName from Students s
where s.DOB > all
(select DOB from
Students natural join Enrolled
natural join Courses c
where c.Name='Databases');
```

•Find all students younger than everyone in 'Databases'

```
select s.sName from Students s
where s.DOB > all

(select s2 DOB from Students s2
```

```
(select s2.DOB from Students s2 natural join Enrolled natural join Courses c where c.cName='Databases');
```

•Find all students younger than everyone in 'Databases'

```
select s.sName from Students s
where s.DOB > all

  (select s2.DOB from Students s2
   natural join Enrolled natural join
   Courses c where c.cName='Databases');
```

•Think of nested queries as nested for-loops

Nested Queries

•Think of nested queries as nested for-loops

```
select s.sName
from Students s
where s.DOB > all
```

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

```
(select s2.DOB
from Students s2
...);
```

DOB	
1980	
1979	

Nested Queries

•Think of nested queries as nested for-loops

```
select s.sName
from Students s
where s.DOB > all
```

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

(select s2.DOB
from Students s2
...);

DOB \$2 1980 1979

5

•Think of nested queries as nested for-loops

```
select s.sName
from Students s
where s.DOB > all
```

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

(select s2.DOB
from Students s2
...);

	DOB	
	1980	
s2	1979	

5

•Think of nested queries as nested for-loops

select s.sName
from Students s
where s.DOB > all

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

(select s2.DOB
from Students s2
...);

DOB \$2 1980 1979

S

•Think of nested queries as nested for-loops

select s.sName
from Students s
where s.DOB > all

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

(select s2.DOB
from Students s2
...);

	DOB	
	1980	
s2	1979	

S

```
(select s2.DOB
select s.sName
                          from Students s2
from Students s
                          . . . ) ;
where s.DOB > all
   foreach Student s in Students {
     foreach Student s2 in ... {
       if(s.DOB \le s2.DOB)
         don't select s
```

EXISTS

•Filter by complex nested query

```
select x from y where
EXISTS
(select ...);
```

•If any rows exist in nested query, x is selected

EXISTS

•Filter by complex nested query

```
select x from y where
NOT EXISTS
(select ...);
```

•If nested query empty (returns nothing), x is selected

•Find students taking all classes

Students s

sID	Name
1	Hermione
2	Harry

Courses c

cID	Name
3500	SW Practice
3810	Architecture

Enrolled e

sID	cID
1	3500
1	3810
2	3810

•Find students taking all classes:

•Find students taking all classes:

```
select s.sName

from Students s

where not exists(select c.cID

from Courses c

where not exists(select e.cID
```

foreach Student s foreach Course c foreach Enrollment e

```
(select e.cID
  from Enrolled e
  where e.cID = c.cID
  and e.sID = s.sID));
```

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exis
```

foreach Student s foreach Course c foreach Enrollment e

If the student s is enrolled in the course c, this is non-empty.

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
from Enrolle
```

If inner select is non-empty, then this is false

We will not select c if the student is enrolled in it.

foreach Student s foreach Course c foreach Enrollment e

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

•Find students taking all classes:

If inner select is non-empty, then this is false

If false for all {s, c}, then this select is empty foreach Student s foreach Course c foreach Enrollment e

```
where not exists(select e.cID

from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));

elect is non-empty
```

•Find students taking all classes:

```
from Students s

where not exists (select c.cID

from Courses c

where not exists (select e.cID

which makes

from Enrolle
```

which makes this true

foreach Student s foreach Course c foreach Enrollment e

```
If inner select is non-empty, then this is false
```

If false for all {s, c}, then this select is empty

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

•Find students taking all classes:

foreach Student s foreach Course c foreach Enrollment e

```
where not exists(select e.cID

from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));
```

Middle query is basically saying "what are all the courses the student is NOT enrolled in". If this is empty, the student is enrolled in every course.

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
```

What happens if the student is not enrolled in every course?

foreach Student s foreach Course c foreach Enrollment e

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

If there is **not** an enrollment for {s, c}
Then this select is empty for that {s,c}

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
from Enrolle
```

If inner select is empty, then this is true

foreach Student s foreach Course c foreach Enrollment e

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

•Find students taking all classes:

If inner select is empty, then this is true

If ever true, then this select is non-empty

foreach Student s foreach Course c foreach Enrollment e

```
where not exists (select e.cID

from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));
```

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
```

which makes this false

If inner select is empty, then this is true

If ever true, then this select is non-empty

foreach Student s foreach Course c foreach Enrollment e

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

foreach Student s foreach Course c foreach Enrollment e

```
where not exists(select e.cID

from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));
```

S

sID	Name
1	Hermione
2	Harry

C

cID	Name
3500	SW Practice
3810	Architecture

 ϵ

sID	cID
1	3500
1	3810
2	3810

foreach Student s foreach Course c foreach Enrollment e

S

sID	Name
1	Hermione
2	Harry

C

cID	Name
3500	SW Practice
3810	Architecture

 ϵ

s ID	cID
1	3500
1	3810
2	3810

foreach Student s foreach Course c foreach Enrollment e

S

sID	Name
1	Hermione
2	Harry

C

cID	Name
3500	SW Practice
3810	Architecture

e

sID	cID
1	3500
1	3810
2	3810

{3500}

foreach Students

select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));

s

foreach Course c
foreach Enrollment e

foreach Enrollment e

foreach Course c
foreach Enrollment e

from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));

sID	Name
1	Hermione
2	Harry

cID	Name
3500	SW Practice
3810	Architecture

sID	cID
1	3500
1	3810
2	3810

```
select s.sName
from Students s
where not exists (select c.cID
                  from Courses c
                  where not exists (select e.cID
```

foreach Student s foreach Course c foreach Enrollment e

from Enrolled e where e.cID = c.cIDand e.sID = s.sID));

Name

Hermione

Harry

sID

S

 \mathbf{C}

cID	Name
3500	SW Practice
3810	Architecture

sID	cID
1	3500
1	3810
2	3810

select s.sName from Students s where not exists (select c.cID from Courses c where not exists (select e.cID

foreach Student s foreach Course c foreach Enrollment e

from Enrolled e where e.cID = c.cIDand e.sID = s.sID));

{3810}

Name

Hermione

S

Harry

sID

Name cID 3500 SW Practice 3810 Architecture

 \mathbf{C}

sID	cID
1	3500
1	3810
2	3810

select s.sName from Students s where not exists (select c.cID from Courses c where not exists (select e.cID

foreach Student s foreach Course c foreach Enrollment e

from Enrolled e where e.cID = c.cIDand e.sID = s.sID));

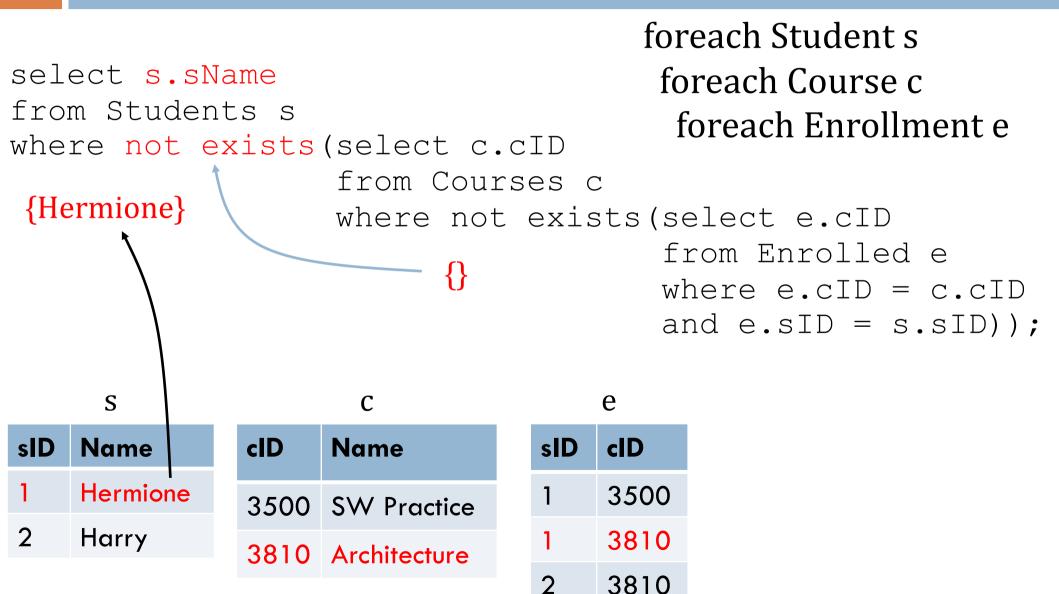
sID Name Hermione Harry

S

Name cID 3500 SW Practice 3810 Architecture

 \mathbf{C}

sID	cID
1	3500
1	3810
2	3810



foreach Student s foreach Course c foreach Enrollment e

S **Name**

sID

Hermione Harry cID Name3500 SW Practice3810 Architecture

 \mathbf{C}

sID	cID
1	3500
1	3810
2	3810

```
select s.sName
from Students s
where not exists (select c.cID
```

foreach Student s foreach Course c foreach Enrollment e

{Hermoine}

from Courses c

where not exists (select e.cID from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));

S

sID Name Hermione Harry

 \mathbf{C}

cID	Name
3500	SW Practice
3810	Architecture

sID	cID
1	3500
1	3810
2	3810

```
select s.sName
from Students s
where not exists (select c.cID
```

foreach Student s foreach Course c foreach Enrollment e

{Hermoine}

from Courses c

where not exists (select e.cID from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));

S

sID	Name
1	Hermione
2	Harry

cID	Name
3500	SW Practice
3810	Architecture

sID	cID
1	3500
1	3810
2	3810

```
select s.sName
from Students s
where not exists (select c.cID
```

foreach Student s foreach Course c foreach Enrollment e

{Hermoine}

from Courses c

where not exists (select e.cID

from Enrolled e where e.cID = c.cID

and e.sID = s.sID));

S

sID Name Hermione Harry

 \mathbf{C}

cID	Name
3500	SW Practice
3810	Architecture

sID	cID
1	3500
1	3810
2	3810

select s.sName from Students s where not exists (select c.cID from Courses c {Hermoine} where not exists (select e.cID {3500} S Name sID Name cID Hermione 3500 SW Practice Harry 3810 Architecture

foreach Student s foreach Course c foreach Enrollment e

from Enrolled e where e.cID = c.cIDand e.sID = s.sID));

sID cID 3500 3810 3810

```
foreach Students

select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

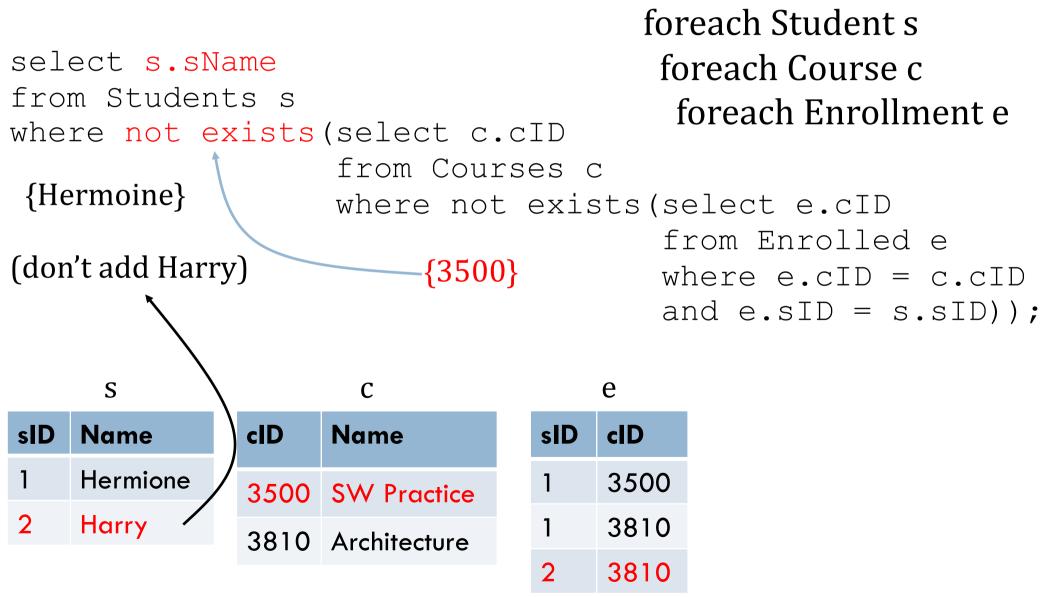
3	
sID	Name
1	Hermione
2	Harry

S

cID	Name
3500	SW Practice
3810	Architecture

 \mathbf{C}

sID	cID
1	3500
1	3810
2	3810



 foreach Student s foreach Course c foreach Enrollment e

where not exists (select e.cID

from Enrolled e
where e.cID = c.cID

and e.sID = s.sID));

Only student taking all classes

S

{Hermoine}

sID	Name
1	Hermione
2	Harry

C

cID	Name
3500	SW Practice
3810	Architecture

 ϵ

sID	cID
1	3500
1	3810
2	3810