CZ2007 Introduction to Database Systems (Week 6)

Topic 5: Relational Algebra (3)

Copyright Notice

Last Lecture

- Assignment: $T_1 := \sigma_{A > 100} R_1$
- Rename: $\rho_{test(A', B', C')} R_1$
- Duplicate Elimination δ
- Extended Projection Π
- Grouping and Aggregation γ

This Lecture

- Division: $\div$
- Left Outerjoin: $\overset{\circ}{\bowtie}_{L}$ condition
- Right Outerjoin: $\overset{\circ}{\bowtie}_{R}$ condition
- Full Outerjoin: $\overset{\circ}{\bowtie}$

Division ÷

Owns

Name	Product
Alice	iPad
Alice	iPhone
Bob	iPhone
Cathy	iPad

AppleP

Product
iPhone
iPad

Results

Name
Alice

- Query: "Find each person that owns all Apple products"
- Owns ÷ AppleP
- In general, $R_1(A, B) \div R_2(B)$ returns a table that contains only A
- The table contains each A value in R_1 that is associated with every B value in R_2
- Intuitive interpretation: "Find the A that R_1 all the B in R_2"
- Example: "Find the Name that Owns all the Product in AppleP"

Division ÷

Joins

Name	Club
Alice	ABC
Bob	DEF
Bob	ABC
Cathy	DEF

Clubs

Club
ABC
DEF

Results

Name
Bob

- Query: “Find each person that has joined all clubs”
- Joins ÷ Clubs

Division ÷

Joins

Name	Club
Alice	ABC
Bob	DEF
Bob	ABC
Cathy	DEF

Clubs

Name
ABC
DEF

Results

Club

- Joins ÷ Clubs ?
- No result.
- In general, $R_1(A, B) \div R_2(B)$ returns a table that contains only A

Division ÷

Owns

Name	Product
Alice	iPad
Alice	iPhone
Bob	iPhone
Cathy	iPad

AppleP

Product	Price
iPhone	999
iPad	699

Results

Name
Alice

- Query: "Find each person that owns all Apple products"
- Owns ÷ AppleP ?
- Wrong, since "Price" does not appear in "Owns"
- In general, $R_1(A, B) \div R_2(B)$ returns a table that contains only A
- Correct answer: Owns ÷ $(\Pi_{Product}$AppleP)

Division ÷

Owns

Name	Since	Product
Alice	2013	iPhone
Alice	2013	iPad
Bob	2013	iPhone
Bob	2010	iPad

AppleP

Product
iPhone
iPad

Results

Name	Since
Alice	2013

- Owns ÷ AppleP
- The result is a table with attributes in Owns but not in AppleP, i.e., Name and Since
- The table contains every {Name, Since} combination that is associated with all Product in AppleP

Exercise

Grades

Name	Course	Grade
Alice	DB	A
Alice	DM	C
Bob	DB	B
Bob	NN	B
Cathy	SP	B
Cathy	NN	A

CrsSch

Course	School
DB	SCSE
DM	SCSE
NN	EEE
SP	EEE

- Query: "Find the students who have taken all courses from SCSE"
- R1 := $\sigma_{\text{School = 'SCSE'}}$ CrsSch
- R2 := Π_{Course} R1
- R3 := $\Pi_{\text{Name, Course}}$ (Grades)
- R4 := R3 ÷ R2

Exercise

Grades

Name	Course	Grade
Alice	DB	A
Alice	DM	C
Bob	DB	B
Bob	NN	B
Cathy	SP	B
Cathy	NN	A

CrsSch

Course	School
DB	SCSE
DM	SCSE
NN	EEE
SP	EEE

- Query: "Find the students who have taken all courses from SCSE but not all courses from EEE"
- R1 := $\Pi_{\text{Course}}(\sigma_{\text{School = 'SCSE'}}\ \text{CrsSch})$
- R2 := $\Pi_{\text{Name, Course}}(\text{Grades})$
- R3 := R2 ÷ R1
- R4 := $\Pi_{\text{Course}}(\sigma_{\text{School = 'EEE'}}\ \text{CrsSch})$
- R5 := $\Pi_{\text{Name, Course}}(\text{Grades})$
- R6 := R3 – (R5 ÷ R4)

Left Outerjoin $\overset{\circ}{\bowtie}_L$ condition

Students

Name	School
Alice	SCSE
Bob	EEE
Cathy	CEE
David	SCSE

Donations

Name	Amount
Cathy	100
David	200
Eddie	300
Fred	400

Results

Name	School	Amount
Alice	SCSE	NULL
Bob	EEE	NULL
Cathy	CEE	100
David	SCSE	200

- Query: "For each student, find the amount of his/her donation"
- Students $\overset{\circ}{\bowtie}_L$ Donations
- All tuples in Students are retained in the results
- For each student who has not made a donation, a "NULL" value is given as his/her Amount

Left Outerjoin $\overset{\circ}{\bowtie}_L$ condition

Students

SName	School
Alice	SCSE
Bob	EEE
Cathy	CEE
David	SCSE

Donations

Name	Amount
Cathy	100
David	200
Eddie	300
Fred	400

Results

SName	Name	School	Amount
Alice	NULL	SCSE	NULL
Bob	NULL	EEE	NULL
Cathy	Cathy	CEE	100
David	David	SCSE	200

- Query: "For each student, find the amount of his/her donation"
- Students $\overset{\circ}{\bowtie}_{L\ \text{Sname} = \text{Name}}$ Donations
- Similar to theta joins in that all attributes are retained

Right Outerjoin $\overset{\circ}{\bowtie}_R$ condition

Students

Name	School
Alice	SCSE
Bob	EEE
Cathy	CEE
David	SCSE

Donations

Name	Amount
Cathy	100
David	200
Eddie	300
Fred	400

Results

Name	School	Amount
Cathy	CEE	100
David	SCSE	200
Eddie	NULL	300
Fred	NULL	400

- Query: "For each donator, find the school he/she is in"
- Students $\overset{\circ}{\bowtie}_R$ Donations
- All tuples in Donations are retained in the results
- For each donator who is not students, a "NULL" value is given as his/her School

Right Outerjoin $\overset{\circ}{\bowtie}_R$ condition

Students

SName	School
Alice	SCSE
Bob	EEE
Cathy	CEE
David	SCSE

Donations

Name	Amount
Cathy	100
David	200
Eddie	300
Fred	400

Results

SName	Name	School	Amount
Cathy	Cathy	CEE	100
David	David	SCSE	200
NULL	Eddie	NULL	300
NULL	Fred	NULL	400

- Query: "For each donator, find the school he/she is in"
- Students $\overset{\circ}{\bowtie}_{R\ \text{Sname} = \text{Name}}$ Donations
- Similar to theta joins in that all attributes are retained

Full Outerjoin $\overset{\circ}{\bowtie}_{\text{condition}}$

Students

Name	School
Alice	SCSE
Bob	EEE
Cathy	CEE
David	SCSE

Donations

Name	Amount
Cathy	100
David	200
Eddie	300
Fred	400

Results

Name	School	Amount
Alice	SCSE	NULL
Bob	EEE	NULL
Cathy	CEE	100
David	SCSE	200
Eddie	NULL	300
Fred	NULL	400

- The combination of left and right outerjoins
- Students $\overset{\circ}{\bowtie}$ Donations

Full Outerjoin $\overset{\circ}{\bowtie}_{\text{condition}}$

Students

SName	School
Alice	SCSE
Bob	EEE
Cathy	CEE
David	SCSE

Donations

Name	Amount
Cathy	100
David	200
Eddie	300
Fred	400

Results

SName	Name	School	Amount
Alice	NULL	SCSE	NULL
Bob	NULL	EEE	NULL
Cathy	Cathy	CEE	100
David	David	SCSE	200
NULL	Eddie	NULL	300
NULL	Fred	NULL	400

- Students $\overset{\circ}{\bowtie}_{\text{Sname = Name}}$ Donations

Example

CastIn

Name	Movie	Year
John	Batman	2012
Steve	Batman	2012
Meg	The Women	2008

Stars

Name	Gender	Birth
John	Male	1980
Meg	Female	1981
Steve	Male	1990

- For each movie, count the number of male movie stars that were cast in the movie
- R1 := $\sigma_{\text{Gender = 'Male'}}$ Stars
- R2 := CastIn ⋈ R1
- Incomplete!
- "All female cast" movies not included

Name	Movie	Year	Gender	Birth
John	Batman	2012	Male	1980
Steve	Batman	2012	Male	1990

Example

CastIn

Name	Movie	Year
John	Batman	2012
Steve	Batman	2012
Meg	The Women	2008

Stars

Name	Gender	Birth
John	Male	1980
Meg	Female	1981
Steve	Male	1990

- For each movie, count the number of male movie stars that were cast in the movie
- R1 := $\sigma_{\text{Gender = 'Male'}}$ Stars
- R2 := CastIn $\overset{\circ}{\bowtie}_{L}$ R1
- R3 := $\gamma_{\text{Movie, COUNT(Gender)} \rightarrow \text{MaleStars}}$ (R2)

Example

CastIn

Name	Movie	Year
John	Batman	2012
Steve	Batman	2012
Meg	The Women	2008

Stars

Name	Gender	Birth
John	Male	1980
Meg	Female	1981
Steve	Male	1990

R2

Name	Movie	Year	Gender	Birth
John	Batman	2012	Male	1980
Steve	Batman	2012	Male	1990
Meg	The Women	2008	NULL	NULL

R3

Movie	MaleStars
Batman	2
The Women	1

- R2 := CastIn $\overset{\circ}{\bowtie}_L$ R1
- R3 := $\gamma_{\text{Movie, COUNT(Gender)} \rightarrow \text{MaleStars}}$ (R2)

Example

- For each movie, count the number of male movie stars that were cast in the movie
- R1 := $\sigma_{\text{Gender = 'Male'}}$ Stars
- R2 := $\gamma_{\text{Name, COUNT(Gender)} \rightarrow \text{Male}}$ (R1)
- R3 := CastIn $\overset{\circ}{\bowtie}_{L}$ R2
- R4 := $\gamma_{\text{Movie, SUM(Male)} \rightarrow \text{MaleStars}}$ (R3)

R2

Name	Male
John	1
Steve	1

R3

Name	Movie	Year	Male
John	Batman	2012	1
Steve	Batman	2012	1
Meg	The Women	2008	NULL

R4

Movie	MaleStars
Batman	2
The Women	0

The End