速成教育
SPEED UP EDUCATION

CSCB20

Introduction to Database & Web Applications

导师： **VC**

UTSC Week 02 Class | 2024/1/21

Disclaimer

This study handout is provided by Speed Up Education and its affiliated tutors. The study handout seeks to support your study process and should be used as a complement, not substitute to the university course material, lecture notes, problem, sets, past tests and other available resources.

The study handout is distributed for free to the students participating in Speed Up Education classes and is not for sale or other commercial uses whatsoever. We kindly ask you to refrain from copying or selling in part or in whole any information provided in the handout.

Thanks for choosing Speed Up Education. We wish you the best of luck in your exam.

What will we learn in this course?

Database 数据库

Important Terminology:

- **Database:** A Collection of interrelated data (Tables)
- **Relation**: Table
- **Tuples**: Rows
- **Attributes**: Columns
- **Instance** of database: A screenshot of the database at a particular time
- **Databae Schema:** The design of database
- **Relation Schema:** The design of a table (what attributes there are?)
- **Super Key:** A set of one or more attributes that taken together **uniquely identity** a tuple in the relation.
- **Candidate Key:** Minimal Superkey (cannot remove anymore without breaking uniqueness.)
- **Primary Key:** A candidate key chosen to distinguish between tuples
- **Foreign Keys**: A set of attributes in a relation that is a **primary key** in another relation.

Example:

Consider the instructor table given in lecture.

dept_name	*building*	*budget*
Biology	Watson	90000
Comp. Sci.	Taylor	100000
Elec. Eng.	Taylor	85000
Finance	Painter	120000
History	Painter	50000
Music	Packard	80000
Physics	Watson	70000

ID	*name*	*dept_name*	*salary*
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

1. What are the **relations**?

2. What are the **tuples**?

course_id	*title*	*dept_name*	*credits*
BIO-101	Intro. to Biology	Biology	4
BIO-301	Genetics	Biology	4
BIO-399	Computational Biology	Biology	3
CS-101	Intro. to Computer Science	Comp. Sci.	4
CS-190	Game Design	Comp. Sci.	4
CS-315	Robotics	Comp. Sci.	3
CS-319	Image Processing	Comp. Sci.	3
CS-347	Database System Concepts	Comp. Sci.	3
EE-181	Intro. to Digital Systems	Elec. Eng.	3
FIN-201	Investment Banking	Finance	3
HIS-351	World History	History	3
MU-199	Music Video Production	Music	3
PHY-101	Physical Principles	Physics	4

3. What are the **attributes** in Department Relation?

ID	*course_id*	*sec_id*	*semester*	*year*
10101	CS-101	1	Fall	2009
10101	CS-315	1	Spring	2010
10101	CS-347	1	Fall	2009
12121	FIN-201	1	Spring	2010
15151	MU-199	1	Spring	2010
22222	PHY-101	1	Fall	2009
32343	HIS-351	1	Spring	2010
45565	CS-101	1	Spring	2010
45565	CS-319	1	Spring	2010
76766	BIO-101	1	Summer	2009
76766	BIO-301	1	Summer	2010
83821	CS-190	1	Spring	2009
83821	CS-190	2	Spring	2009
83821	CS-319	2	Spring	2010
98345	EE-181	1	Spring	2009

4. What are the **Relation Schemas**?

5. Why some Schemas contains the same attributes?

dept_name	*building*	*budget*
Biology	Watson	90000
Comp. Sci.	Taylor	100000
Elec. Eng.	Taylor	85000
Finance	Painter	120000
History	Painter	50000
Music	Packard	80000
Physics	Watson	70000

6. What are the possible **superkeys** of Instructor Relation?

ID	*name*	*dept_name*	*salary*
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

7. Which are the **Candidate keys**?

course_id	*title*	*dept_name*	*credits*
BIO-101	Intro. to Biology	Biology	4
BIO-301	Genetics	Biology	4
BIO-399	Computational Biology	Biology	3
CS-101	Intro. to Computer Science	Comp. Sci.	4
CS-190	Game Design	Comp. Sci.	4
CS-315	Robotics	Comp. Sci.	3
CS-319	Image Processing	Comp. Sci.	3
CS-347	Database System Concepts	Comp. Sci.	3
EE-181	Intro. to Digital Systems	Elec. Eng.	3
FIN-201	Investment Banking	Finance	3
HIS-351	World History	History	3
MU-199	Music Video Production	Music	3
PHY-101	Physical Principles	Physics	4

8. Which should be the **primary key**?

ID	*course_id*	*sec_id*	*semester*	*year*
10101	CS-101	1	Fall	2009
10101	CS-315	1	Spring	2010
10101	CS-347	1	Fall	2009
12121	FIN-201	1	Spring	2010
15151	MU-199	1	Spring	2010
22222	PHY-101	1	Fall	2009
32343	HIS-351	1	Spring	2010
45565	CS-101	1	Spring	2010
45565	CS-319	1	Spring	2010
76766	BIO-101	1	Summer	2009
76766	BIO-301	1	Summer	2010
83821	CS-190	1	Spring	2009
83821	CS-190	2	Spring	2009
83821	CS-319	2	Spring	2010
98345	EE-181	1	Spring	2009

9. What are the **Foreign Keys**?

Example Tables

Product

ProductID	ProductName	SupplierID	CategoryID	Unit	Price
1	Chais	1	1	10 boxes x 20 bags	18
2	Chang	1	1	24 - 12 oz bottles	19
3	Aniseed Syrup	1	2	12 - 550 ml bottles	10
4	Chef Anton's Cajun Seasoning	2	2	48 - 6 oz jars	22
5	Chef Anton's Gumbo Mix	2	2	36 boxes	21.35
6	Grandma's Boysenberry Spread	3	2	12 - 8 oz jars	25
7	Uncle Bob's Organic Dried Pears	3	7	12 - 1 lb pkgs.	30
8	Northwoods Cranberry Sauce	3	2	12 - 12 oz jars	40
9	Mishi Kobe Niku	4	6	18 - 500 g pkgs.	97
10	Ikura	4	8	12 - 200 ml jars	31

Categories

CategoryID	CategoryName	Description
1	Beverages	Soft drinks, coffees, teas, beers, and ales
2	Condiments	Sweet and savory sauces, relishes, spreads, and seasonings
3	Confections	Desserts, candies, and sweet breads
4	Dairy Products	Cheeses
5	Grains/Cereals	Breads, crackers, pasta, and cereal
6	Meat/Poultry	Prepared meats
7	Produce	Dried fruit and bean curd
8	Seafood	Seaweed and fish

Relational Algebra

A query language, but in a more Mathematical way ☺

Select tuples (rows) from a relation (table) statisfying a predicate (condition)

SELECT (σ)

The SELECT operation is used for selecting a subset of the tuples according to a given selection condition. Sigma(σ)Symbol denotes it. It is used as an expression to choose tuples which meet the selection condition. Select operation selects tuples that satisfy a given predicate.

$$\sigma_p(r)$$

σ is the predicate

r stands for relation which is the name of the table

p is prepositional logic (boolean formula) **Connectives**: and, or, not **Operators**: <, >, ≤, ≥, =, ≠

Examples:

$$\sigma_{price \geq 20}(Products)$$

ProductID	ProductName	SupplierID	CategoryID	Unit	Price
4	Chef Anton's Cajun Seasoning	2	2	48 - 6 oz jars	22
5	Chef Anton's Gumbo Mix	2	2	36 boxes	21.35
6	Grandma's Boysenberry Spread	3	2	12 - 8 oz jars	25
7	Uncle Bob's Organic Dried Pears	3	7	12 - 1 lb pkgs.	30
8	Northwoods Cranberry Sauce	3	2	12 - 12 oz jars	40

$$\sigma_{price \geq 20 \text{ and productName="Ikura"}}(Products)$$

ProductID	ProductName	SupplierID	CategoryID	Unit	Price
10	Ikura	4	8	12 - 200 ml jars	31

Projection(π)

The projection eliminates all attributes of the input relation but those mentioned in the projection list. The projection method defines a relation that contains a vertical subset of Relation.

This helps to extract the values of specified attributes to eliminates duplicate values. (pi) The symbol used to choose attributes from a relation. This operation helps you to keep specific columns from a relation and discards the other columns.

Example 1:

$$\Pi_{\text{productName, price}}(Products)$$

ProductName	Price
Chais	18
Chang	19
Aniseed Syrup	10
Chef Anton's Cajun Seasoning	22
Chef Anton's Gumbo Mix	21.35
Grandma's Boysenberry Spread	25
Uncle Bob's Organic Dried Pears	30

$$\Pi_{\text{productName, price}}(\sigma_{price \geq 40 \text{ or productID}=5}(Products))$$

ProductName	Price
Chef Anton's Gumbo Mix	21.35
Mishi Kobe Niku	97
Carnarvon Tigers	62.5
Sir Rodney's Marmalade	81
Schoggi Schokolade	43.9
Rössle Sauerkraut	45.6

Difference (-)

The result of A - B, is a relation which includes **all tuples that are in A but not in B.**

- The attribute name of A has to match with the attribute name in B.
- The two-operand relations A and B should be either compatible or Union compatible.
- It should be defined relation consisting of the tuples that are in relation A, but not in B.

Example:

Table A		Table B	
column 1	**column 2**	**column 1**	**column 2**
a	1	a	2
a	2	b	3
b	1		

Table A - B	
column 1	**column 2**
a	1
b	1

Union operation (υ)

UNION is symbolized by ∪ symbol. It includes **all tuples that are in tables A or in B**. It also **eliminates duplicate tuples**. So, set A UNION set B would be expressed as: A ∪ B

For a union operation to be valid, the following conditions must hold -

- A and B must be the same number of attributes.
- Attribute domains need to be compatible.
- Duplicate tuples should be automatically removed.

Example:

Table A		Table B	
column 1	**column 2**	**column 1**	**column 2**
a	1	a	2
a	2	b	3
b	1		

Table A ∪ B	
column 1	**column 2**
a	1
a	2
b	1
b	3

Intersection operation (∩)

An intersection is defined by the symbol ∩.

A ∩ B

Defines a relation consisting of a set of all tuple that are in both A and B. However, A and B must be union-compatible.

Example:

Table A		Table B	
column 1	**column 2**	**column 1**	**column 2**
a	1	a	2
a	2	b	3
b	1		

Table A ∩ B	
column 1	**column 2**
a	2

Note: A ∩ B = A – (A – B)

Natural Join (⋈)

The tuples are joined if the attributes common to both relations are equal.

Natural join can only be performed if there is a common attribute (column) between the relations. The name and type of the attribute must be same.

Example:

Table A		Table B	
Num	**Square**	**Num**	**Cube**
2	4	2	8
3	9	3	27
4	16	5	125

Table A ⋈ B		
Num	**Square**	**Cube**
2	4	8
3	9	27

Theta Join ($A \bowtie_p B$)

Products ⋈ Categories

ProductID	ProductName	SupplierID	CategoryID	Unit	Price	CategoryName	Description
1	Chais	1	1	10 boxes x 20 bags	18	Beverages	Soft drinks, coffees, teas, beers, and ales
2	Chang	1	1	24 - 12 oz bottles	19	Beverages	Soft drinks, coffees, teas, beers, and ales
3	Aniseed Syrup	1	2	12 - 550 ml bottles	10	Condiments	Sweet and savory sauces, relishes, spreads, and seasonings
4	Chef Anton's Cajun Seasoning	2	2	48 - 6 oz jars	22	Condiments	Sweet and savory sauces, relishes, spreads, and seasonings
5	Chef Anton's Gumbo Mix	2	2	36 boxes	21.35	Condiments	Sweet and savory sauces, relishes, spreads, and seasonings
6	Grandma's Boysenberry Spread	3	2	12 - 8 oz jars	25	Condiments	Sweet and savory sauces, relishes, spreads, and seasonings
7	Uncle Bob's Organic Dried Pears	3	7	12 - 1 lb pkgs.	30	Produce	Dried fruit and bean curd
8	Northwoods Cranberry Sauce	3	2	12 - 12 oz jars	40	Condiments	Sweet and savory sauces, relishes, spreads, and seasonings
9	Mishi Kobe Niku	4	6	18 - 500 g pkgs.	97	Meat/Poultry	Prepared meats

Cartesian product(X)

Cartesian Product is an operation used to merge columns from two relations. Generally, a cartesian product is never a meaningful operation when it performs alone. However, it becomes meaningful when it is followed by other operations. It is also called Cross Product or Cross Join.

Table A		Table B	
column 1	**column 2**	**column 1**	**column 2**
a	1	a	2
a	2	b	3
b	1		

A x B

column 1	**column 2**	**column 1**	**column 2**
a	1	a	2
a	1	b	3
a	2	a	2
a	2	b	3
b	1	a	2
b	1	b	3

$$p_v(A) \times p_c(B)$$

v.column 1	**v.column 2**	**c.column 1**	**c.column 2**
a	1	a	2
a	1	b	3
a	2	a	2
a	2	b	3
b	1	a	2
b	1	b	3