

VCU, Department of Computer Science

CMSC 302

Relations

Vojislav Kecman

Sections 1, 2, 3 and 4 in
Ch 8 in 6th edition
Ch 9 in 7th edition

14/09/2015

1/53

We talk in terms of relations, we relate items, objects, ..., etc

- **Examples:**
- is bigger than,
- is more expensive than,
- is parallel to,
- is a subset of
- is
- All these expressions **relate some 'items'**, or they express the existence or nonexistence of a certain connection between a pair of objects taken **in a definite order**

14/09/2015

2/53

Relations

• If we want to describe a relationship between elements of two sets A and B, we can use **ordered pairs** with their first element taken from A and their second element taken from B.

• Since this is a relation between **two sets**, it is called a **binary relation**.

• **Definition:** Let A and B be sets. A binary relation from A to B is a **subset** of $A \times B$. (Hence, A **BINARY RELATION IS A SET of PAIRS**).

• In other words, for a binary relation R we have $R \subseteq A \times B$. We use the notation aRb to denote that $(a, b) \in R$ and $a \nabla R b$ to denote that $(a, b) \notin R$.

14/09/2015

3/53

Relations

- When (a, b) belongs to R, a is said to be **related** to b by R.
- **Example:** Let P be a set of people, C be a set of cars, and D be the relation describing which person drives which car(s).
- $P = \{\text{Carl, Suzanne, Peter, Carla}\}$,
- $C = \{\text{Mercedes, BMW, tricycle}\}$
- $D = \{(\text{Carl, Mercedes}), (\text{Suzanne, Mercedes}), (\text{Suzanne, BMW}), (\text{Peter, tricycle})\}$
- This means that Carl drives a Mercedes, Suzanne drives a Mercedes and a BMW, Peter drives a tricycle, and Carla does not drive any of these vehicles.

14/09/2015

4/53

And, what would be a Cartesian product $P \times C$?

- $P \times C = \{(Carl, Mercedes)^*, (Carl, BMW), (Carl, tricycle), (Suzanne, Mercedes)^*, (Suzanne, BMW)^*, (Suzanne, tricycle), (Peter, Mercedes), (Peter, BMW), (Peter, tricycle)^*, (Carla, Mercedes), (Carla, BMW), (Carla, tricycle)\}$

OBVIOUSLY!!!

$$pRc \subseteq P \times C$$

14/09/2015

5/53

Functions as Relations

- You might remember that a **function** f from a set A to a set B assigns a **unique** element of B to each element of A .
- The **graph** of f is the set of ordered pairs (a, b) such that $b = f(a)$.
- Since the graph of f is a subset of $A \times B$, it is a **relation** from A to B .
- Moreover, for each element a of A , there is exactly one ordered pair in the graph that has a as its first element.

14/09/2015

6/53

Functions as Relations

- Conversely, if R is a relation from A to B such that every element in A is the first element of exactly one ordered pair of R , then a function can be defined with R as its graph.
- This is done by assigning to an element $a \in A$ the unique element $b \in B$ such that $(a, b) \in R$.

14/09/2015

7/53

Thus,

- **Relations are a generalization of functions and they can be used to express a much wider class of relationships between sets**

14/09/2015

8/53

Relations on a Set

•**Definition:** A relation on the set A is a relation from A to A .

•In other words, a relation on the set A is a subset of $A \times A$.

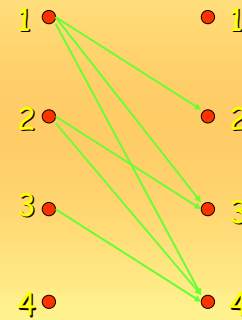
•**Example:** Let $A = \{1, 2, 3, 4\}$. Which ordered pairs are in the relation $R = \{(a, b) \mid a < b\}$?

14/09/2015

9/53

Relations on a Set

•**Solution:** $R = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$

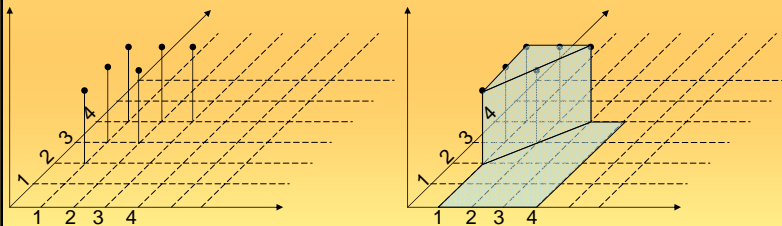


R	1	2	3	4
1		X	X	X
2			X	X
3				X
4				

14/09/2015

10/53

We can show this relation graphically over 2-dim space



14/09/2015

11/53

We can also state

If R is a relation from a set A to itself, that is, if R is a subset of $A^2 = A \times A$, then we say that R is a relation on A .

The *domain* of a relation R is the set of all first elements of the ordered pairs which belong to R , and the *range* of R is the set of second elements.

14/09/2015

12/53

Examples:

(a) Let $A = \{1, 2, 3\}$ and $B = \{x, y, z\}$, and let $R = \{(1, y), (1, z), (3, y)\}$. Then R is a relation from A to B since R is a subset of $A \times B$. With respect to this relation,

$$1Ry, 1Rz, 3Ry, \quad \text{but} \quad 1Rx, 2Rx, 2Ry, 2Rz, 3Rx, 3Rz$$

The domain of R is $\{1, 3\}$ and the range is $\{y, z\}$.

(b) Let $A = \{\text{eggs, milk, corn}\}$ and $B = \{\text{cows, goats, hens}\}$. We can define a relation R from A to B by $(a, b) \in R$ if a is produced by b . In other words,

$$R = \{(\text{eggs, hens}), (\text{milk, cows}), (\text{milk, goats})\}$$

With respect to this relation,

$$\text{eggs } R \text{ hens, milk } R \text{ cows, etc.}$$

(c) Suppose we say that two countries are *adjacent* if they have some part of their boundaries in common. Then "is adjacent to" is a relation R on the countries of the earth. Thus

$$(\text{Italy, Switzerland}) \in R \quad \text{but} \quad (\text{Canada, Mexico}) \notin R$$

14/09/2015

13/53

Properties of Relations

•We will now look at some useful ways to classify relations.

•**Definition:** A relation R on a set A is called **reflexive** if $(a, a) \in R$ for **every** element $a \in A$.

•Are the following relations on $\{1, 2, 3, 4\}$ reflexive?

• $R = \{(1, 1), (1, 2), (2, 3), (3, 3), (4, 4)\}$

•No.

• $R = \{(1, 1), (2, 2), (2, 3), (3, 3), (4, 4)\}$

•Yes.

• $R = \{(1, 1), (2, 2), (3, 3)\}$

•No.

•**Definition:** A relation on a set A is called **irreflexive** if $(a, a) \notin R$ for **every** element $a \in A$.

14/09/2015

14/53

Examples:

Consider the following five relations on the set $A = \{1, 2, 3, 4\}$:

$$R_1 = \{(1, 1), (1, 2), (2, 3), (1, 3), (4, 4)\}$$

$$R_2 = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}$$

$$R_3 = \{(1, 3), (2, 1)\}$$

$$R_4 = \emptyset, \text{ the empty relation}$$

$$R_5 = A \times A, \text{ the universal relation}$$

Determine which of the relations are reflexive.

Solutions:

Since A contain the four elements 1, 2, 3, and 4, a relation R on A is reflexive if it contains the four pairs $(1, 1)$, $(2, 2)$, $(3, 3)$, and $(4, 4)$. Thus only R_2 and the universal relation $R_5 = A \times A$ are reflexive. Note that R_1 , R_3 , and R_4 are not reflexive since, for example, $(2, 2)$ does not belong to any of them.

14/09/2015

15/53

More examples Consider the following five relations:

(1) Relation \leq (less than or equal) on the set \mathbf{Z} of integers

(2) Set inclusion \subseteq on a collection \mathbf{C} of sets

(3) Relation \perp (perpendicular) on the set \mathbf{L} of lines in the plane.

(4) Relation \parallel (parallel) on the set \mathbf{L} of lines in the plane.

(5) Relation $|$ of divisibility on the set \mathbf{N} of positive integers. (Recall $x|y$ if there exists z such that $xz = y$.)

Determine which of the relations are reflexive.

Solutions:

The relation (3) is not reflexive since no line is perpendicular to itself. Also (4) is not reflexive since no line is parallel to itself. The other relations are reflexive; that is, $x \leq x$ for every integer x in \mathbf{Z} , $A \subseteq A$ for any set A in \mathbf{C} , and $n|n$ for every positive integer n in \mathbf{N} .

14/09/2015

16/53

Properties of Relations

•Definitions:

- A relation R on a set A is called **symmetric** if $(b, a) \in R$ whenever $(a, b) \in R$ **for all** $a, b \in A$.
- A relation R on a set A is called **antisymmetric** if whenever $(a, b) \in R$, $(b, a) \notin R$.
A relation R on a set A such that for all $a, b \in A$, if $(a, b) \in R$ and $(b, a) \in R$, then $a = b$ is **antisymmetric**.
- A relation R on a set A is called **asymmetric** if $(a, b) \in R$ implies that $(b, a) \notin R$ **for some** $a, b \in A$.

R can be both symmetric and antisymmetric, but it can't be symmetric and asymmetric

14/09/2015

17/53

Properties of Relations

- Are the following relations on $\{1, 2, 3, 4\}$ symmetric, antisymmetric, or asymmetric?

- $R = \{(1, 1), (1, 2), (2, 1), (3, 3), (4, 4)\}$ •symmetric
- $R = \{(1, 1)\}$ •sym. and antisym.
- $R = \{(1, 3), (3, 2), (2, 1)\}$ •antisym. and asym.
- $R = \{(1, 3), (3, 2), (2, 1), (3, 1)\}$ •asym.
- $R = \{(4, 4), (3, 3), (1, 4)\}$ •antisym.

14/09/2015

18/53

Examples: Which relations are symmetric

Consider the following five relations on the set $A = \{1, 2, 3, 4\}$:

- $R_1 = \{(1, 1), (1, 2), (2, 3), (1, 3), (4, 4)\}$
- $R_2 = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}$
- $R_3 = \{(1, 3), (2, 1)\}$
- $R_4 = \emptyset$, the empty relation
- $R_5 = A \times A$, the universal relation

Solutions:

R_1 is not symmetric since $(1, 2) \in R_1$ but $(2, 1) \notin R_1$. R_3 is not symmetric since $(1, 3) \in R_3$ but $(3, 1) \notin R_3$. The other relations are symmetric.

14/09/2015

19/53

Properties of Relations

•**Definition:** A relation R on a set A is called **transitive** if whenever $(a, b) \in R$ and $(b, c) \in R$, then $(a, c) \in R$ for $a, b, c \in A$.

- Are the following relations on $\{1, 2, 3, 4\}$ transitive?

- $R = \{(1, 1), (1, 2), (2, 2), (2, 1), (3, 3)\}$ •Yes.
- $R = \{(1, 3), (3, 2), (2, 1)\}$ •No.
- $R = \{(1, 3), (3, 2), (1, 2)\}$ •Yes.
- $R = \{(2, 4), (4, 3), (2, 3), (4, 1)\}$ •No.

14/09/2015

20/53

Some relations' summary

	=	<	>	≤	≥
Reflexive	X			X	X
Irreflexive		X	X		
Symmetric	X				
Asymmetric		X	X		
Antisymmetric	X			X	X
Transitive	X	X	X	X	X

14/09/2015

21/53

Basics of Counting Permutations and Combinations

- Permutations - Order matters
- Combinations – Order doesn't matter

Combination	Permutations
<i>abc</i>	<i>abc, acb, bac, bca, cab, cba</i>
<i>abd</i>	<i>abd, adb, bad, bda, dab, dba</i>
<i>acd</i>	<i>acd, adc, cad, cda, dac, dca</i>
<i>bcd</i>	<i>bcd, bdc, cbd, cdb, dbc, dcb</i>

14/09/2015

22/53

- **Example 1:** How many words we can make from letters – H, O, W

HOW, HWO, OHW, OWH, WHO, WOH

- In languages – order matters

- **Example 2:** How many words we can make from letters – H, O, W by using 2 letters only

HO, OH, HW, WH, OW, WO

- **Example 3:** How many SETS of 3 letters we can make from letters – H, O, W

HOW, => 1 SET only

- **Example 4:** How many SETS OF 2 LETTERS we can make from letters – H, O, W

HO, HW, OW

14/09/2015

23/53

Counting permutations and combinations

- **Permutations:** $P_r^n = \frac{n!}{(n-r)!}$, Note: $P_n^n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$

- **Combinations:** $C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{\overbrace{n(n-1)(n-2)\cdots}^{r \text{ terms}}}{1*2*3\cdots r}$

$$C_4^8 = \binom{8}{4} = \frac{8*7*6*5}{1*2*3*4} = 70$$

- **Examples:**

Note: $\binom{n}{r} = \binom{n}{n-r}$, eg. $\binom{8}{3} = \binom{8}{5}$,

Note also: $\binom{n}{0} = \binom{n}{n} = 1$, eg. $\binom{8}{0} = \binom{8}{8} = 1$

14/09/2015

24/53

Relations on a Set

•How many different relations can we define on a set A with n elements?

- A relation on a set A is a subset of $A \times A$.
- How many elements are in $A \times A$?

•There are n^2 elements in $A \times A$, so how many subsets (= relations on A) does $A \times A$ have?

•The number of subsets that we can form out of a set with m elements is 2^m . Therefore, 2^{n^2} subsets can be formed out of $A \times A$.

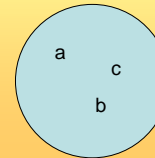
•Answer: We can define 2^{n^2} different relations on A.

14/09/2015

25/53

Let's check the claim about 2^m

Set $A = \{a, b, c\}$



How many subsets of A we can create?

1. $\{\emptyset\}$
2. $\{a\}$
3. $\{b\}$
4. $\{c\}$
5. $\{a,b\}$
6. $\{a,c\}$
7. $\{b,c\}$
8. $\{a,b,c\}$

How can we, in a clear and clean way, find out this number?

Well, it's basically the answer to the question – **how many combinations of 0th, 1st, 2nd, 3rd, ..., mth order** can we make from m elements of the set A?

$$\text{Number of subsets} = \binom{m}{0} + \binom{m}{1} + \binom{m}{2} + \dots + \binom{m}{m-1} + \binom{m}{m}$$

In a particular example above, $m = 3$ and

$$\text{Number of subsets} = \binom{3}{0} + \binom{3}{1} + \binom{3}{2} + \binom{3}{3} = 1 + 3 + 3 + 1 = 8$$

14/09/2015

26/53

Example of Relations Counting

• Show all the different relations on a set $A = \{a, b\}$?

•Solution: Relations on A are subsets of $A \times A$, which contains $2^2 = 4$ elements as follows $\{(a, a), (a, b), (b, a), (b, b)\}$.

•Therefore, different relations on A can be generated by choosing different subsets out of these 4 elements, so there are $2^4 = 16$ relations as follows:

1. $\{\emptyset\}$, 2. $\{(a, a)\}$, 3. $\{(a, b)\}$, 4. $\{(b, a)\}$, 5. $\{(b, b)\}$,
6. $\{(a, a), (a, b)\}$, 7. $\{(a, a), (b, a)\}$, 8. $\{(a, a), (b, b)\}$,
9. $\{(a, b), (b, a)\}$, 10. $\{(a, b), (b, b)\}$, 11. $\{(b, a), (b, b)\}$,
12. $\{(a, a), (a, b), (b, a)\}$, 13. $\{(a, a), (a, b), (b, b)\}$,
14. $\{(a, a), (b, a), (b, b)\}$, 15. $\{(a, b), (b, a), (b, b)\}$,
16. $\{(a, a), (a, b), (b, a), (b, b)\}$

14/09/2015

27/53

Example:

Find the number of relations from $A = \{a, b, c\}$ to $B = \{1, 2\}$.

14/09/2015

28/53

Counting Relations

•**Example:** How many different **reflexive** relations can be defined on a set A containing n elements?

•**Solution:** Relations on R are subsets of $A \times A$, which contains n^2 elements.

•Therefore, different relations on A can be generated by choosing different subsets out of these n^2 elements, so there are 2^{n^2} relations.

•A **reflexive** relation, however, **must** contain the n elements (a, a) for every $a \in A$.

•Consequently, we can only choose among $n^2 - n = n(n - 1)$ elements to generate reflexive relations, so there are $2^{n(n - 1)}$ of them.

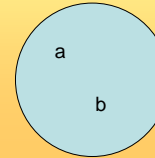
14/09/2015

29/53

Let's check the claim about $2^{n(n - 1)}$

Set $A = \{a, b\}$

How many reflexive relations of A we can create?



$$2^{2(2-1)} = 2^2 = 4$$

1. $\{0\}$, 2. $\{(a, a)\}$, 3. $\{(a, b)\}$, 4. $\{(b, a)\}$, 5. $\{(b, b)\}$,

6. $\{(a, a), (a, b)\}$, 7. $\{(a, a), (b, a)\}$, 8. $\{(a, a), (b, b)\}$,

9. $\{(a, b), (b, a)\}$, 10. $\{(a, b), (b, b)\}$, 11. $\{(b, a), (b, b)\}$,

12. $\{(a, a), (a, b), (b, a)\}$, 13. $\{(a, a), (a, b), (b, b)\}$,

14. $\{(a, a), (b, a), (b, b)\}$, 15. $\{(a, b), (b, a), (b, b)\}$,

16. $\{(a, a), (a, b), (b, a), (b, b)\}$

14/09/2015

30/53

Combining Relations

•Relations are sets, and therefore, we can apply the usual **set operations** to them.

•If we have two relations R_1 and R_2 , and both of them are from a set A to a set B, then we can combine them to $R_1 \cup R_2$, $R_1 \cap R_2$, or $R_1 - R_2$.

•In each case, the result will be **another relation from A to B**.

14/09/2015

31/53

Combining Relations

•... and there is another important way to combine relations.

•**Definition:** Let R be a relation from a set A to a set B and S a relation from B to a set C. The **composite** of R and S is the relation consisting of ordered pairs (a, c), where $a \in A$, $c \in C$, and for which there exists an element $b \in B$ such that $(a, b) \in R$ and $(b, c) \in S$. We denote the composite of R and S by **$S \circ R$** .

•In other words, if relation R contains a pair (a, b) and relation S contains a pair (b, c), then $S \circ R$ contains a pair (a, c).

14/09/2015

32/53

Combining Relations

•**Example:** Let D and S be relations on $A = \{1, 2, 3, 4\}$.

• $D = \{(a, b) \mid b = 5 - a\}$ “ b equals $(5 - a)$ ”

• $S = \{(b, c) \mid b < c\}$ “ b is smaller than c ”

• $D = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$

• $S = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$

• $S \circ D = \{$

$(2, 4), (3, 3), (3, 4), (4, 2), (4, 3), (4, 4)\}$

• D maps an element a to the element $(5 - a)$, and afterwards S maps $(5 - a)$ to all elements larger than $(5 - a)$, resulting in

$S \circ D = \{(a, c) \mid c > 5 - a\}$.

14/09/2015

33/53

Combining Relations

•We already know that **functions** are just **special cases** of **relations** (namely those that map each element in the domain onto exactly one element in the codomain).

•If we formally convert two functions into relations, that is, write them down as sets of ordered pairs, the composite of these relations will be exactly the same as the composite of the functions (as defined earlier).

14/09/2015

34/53

Combining Relations

•**Definition:** Let R be a relation on the set A . The powers R^n , $n = 1, 2, 3, \dots$, are defined inductively by

• $R^1 = R$

• $R^{n+1} = R^n \circ R$

•In other words:

• $R^n = R \circ R \circ \dots \circ R$ (n times the letter R)

14/09/2015

35/53

Combining Relations

•**Theorem:** The relation R on a set A is **transitive** if and only if $R^n \subseteq R$ for all positive integers n .

•Remember the definition of transitivity: 

•**Definition:** A relation R on a set A is called **transitive** if whenever $(a, b) \in R$ and $(b, c) \in R$, then $(a, c) \in R$ for $a, b, c \in A$.

•The composite of R with itself contains exactly these pairs (a, c) .
 •Therefore, for a transitive relation R , $R \circ R$ does not contain any pairs that are not in R , so $R \circ R \subseteq R$.
 •Since $R \circ R$ does not introduce any pairs that are not already in R , it must also be true that $(R \circ R) \circ R \subseteq R$, and so on, so that $R^n \subseteq R$.

14/09/2015

36/53

Examples: Which relations are transitive?

Consider the following five relations on the set $A = \{1, 2, 3, 4\}$:

$$R_1 = \{(1, 1), (1, 2), (2, 3), (1, 3), (4, 4)\}$$

$$R_2 = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}$$

$$R_3 = \{(1, 3), (2, 1)\}$$

$$R_4 = \emptyset, \text{ the empty relation}$$

$$R_5 = A \times A, \text{ the universal relation}$$

Solutions:

The relation R_3 is not transitive since $(2, 1), (1, 3) \in R_3$ but $(2, 3) \notin R_3$. All the other relations are transitive.

14/09/2015

37/53

Examples: Which relations are transitive?

- (1) Relation \leq (less than or equal) on the set \mathbf{Z} of integers
- (2) Set inclusion \subseteq on a collection \mathcal{C} of sets
- (3) Relation \perp (perpendicular) on the set L of lines in the plane.
- (4) Relation \parallel (parallel) on the set L of lines in the plane.
- (5) Relation $|$ of divisibility on the set \mathbf{N} of positive integers. (Recall $x|y$ if there exists z such that $xz = y$.)

Solutions:

The relations \leq , \subseteq , and $|$ are transitive. That is: (i) If $a \leq b$ and $b \leq c$, then $a \leq c$. (ii) If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$. (iii) If $a|b$ and $b|c$, then $a|c$.

On the other hand the relation \perp is not transitive. If $a \perp b$ and $b \perp c$, then it is not true that $a \perp c$. Since no line is parallel to itself, we can have $a \parallel b$ and $b \parallel a$, but $a \not\parallel a$. Thus \parallel is not transitive. (We note that the relation "is parallel or equal to" is a transitive relation on the set L of lines in the plane.)

14/09/2015

38/53

n-ary Relations

•In order to study an interesting application of relations, namely **databases**, we first need to generalize the concept of binary relations to **n-ary relations**.

•**Definition:** Let A_1, A_2, \dots, A_n be sets. An **n-ary relation** on these sets is a **subset of** $A_1 \times A_2 \times \dots \times A_n$.

•The sets A_1, A_2, \dots, A_n are called the **domains** of the relation, and n is called its **degree**.

14/09/2015

39/53

n-ary Relations

•Example:

•Let $R = \{(a, b, c) \mid a = 2b \wedge b = 2c \text{ with } a, b, c \in \mathbf{Z}\}$

•What is the degree of R ?

•The degree of R is 3, so its elements are triples.

•What are its domains?

•Its domains are all equal to the set of integers.

•Is $(2, 4, 8)$ in R ?

•No.

•Is $(4, 2, 1)$ in R ?

•Yes.

14/09/2015

40/53

Databases and Relations

- Let us take a look at a type of database representation that is based on relations, namely the **relational data model**.
- A database consists of n-tuples called **records**, which are made up of **fields**.
- These fields are the **entries** of the n-tuples.
- The relational data model represents a database as an n-ary relation, that is, a set of records.

14/09/2015

41/53

Databases and Relations

Alternative Terminology

- Although **not all tables are relations**, the terms table and relation are normally used interchangeably
- The following sets of terms are equivalent:

Table	Column	Row
Relation	Attribute	Tuple
File	Field	Record



Domain

14/09/2015

42/53

Databases and Relations

• **Example:** Consider a database of students, whose records are represented as 4-tuples with the fields **Student Name**, **ID Number**, **Major**, and **GPA**:

• $R = \{(Ackermann, 231455, CS, 3.88),$
(Adams, 888323, Physics, 3.45),
(Chou, 102147, CS, 3.79),
(Goodfriend, 453876, Math, 3.45),
(Rao, 678543, Math, 3.90),
(Stevens, 786576, Psych, 2.99)\}

• Relations that represent databases are also called **tables**, since they are often displayed as tables.

14/09/2015

43/53

Databases and Relations

- A **domain** of an n-ary relation is called a **primary key** if the n-tuples are **uniquely determined** by their values from this domain.
- This means that no two records have the same value from the same primary key.
- In **our** example, which of the fields **Student Name**, **ID Number**, **Major**, and **GPA** are primary keys?
- **Student Name** and **ID Number** are primary keys, because no two students have identical values in these fields.
- In a real student database, only **ID Number** would be a primary key.

14/09/2015

44/53

Databases and Relations

- In a database, a **primary key should remain same** even if new records are added.
- **Combinations of domains** can also uniquely identify n-tuples in an n-ary relation.
- When the values of a **set of domains** determine an n-tuple in a relation, the **Cartesian product** of these domains is called a **composite key**.

14/09/2015

45/53

Databases and Relations

Keys

- A **key** is a combination of one or more columns that is used to identify rows in a relation
- A **composite key** is a key that consists of two or more columns

14/09/2015

46/53

Candidate and Primary Keys

- A **candidate key** is a key that **determines all of the other columns** in a relation
- A **primary key** is a candidate key selected as the **primary means of identifying rows** in a relation:
 - There is one and only one primary key per relation
 - The primary key may be a composite key
 - The ideal primary key is short, numeric and never changes

14/09/2015

47/53

Databases and Relations

- We can apply a variety of **operations** on n-ary relations to form new relations.
- **Definition:** The **projection** P_{i_1, i_2, \dots, i_m} maps the n-tuple (a_1, a_2, \dots, a_n) to the m-tuple $(a_{i_1}, a_{i_2}, \dots, a_{i_m})$, where $m \leq n$.
- In other words, a projection P_{i_1, i_2, \dots, i_m} keeps the m components $a_{i_1}, a_{i_2}, \dots, a_{i_m}$ of an n-tuple and deletes its $(n - m)$ other components.
- **Example:** What is the result when we apply the projection $P_{2,4}$ to the student record (Stevens, 786576, Psych, 2.99) ?
- **Solution:** It is the pair (786576, 2.99).

14/09/2015

48/53

Databases and Relations

- In some cases, applying a projection to an entire table may not only result in fewer columns, but also in **fewer rows**.

- Why is that?

- Some records may only have differed in those fields that were deleted, so they become **identical**, and there is no need to list identical records more than once.

14/09/2015

49/53

Databases and Relations

- We can use the **join** operation to combine two tables into one **if they share some identical fields**.

- **Definition:** Let R be a relation of degree m and S a relation of degree n. The **join** $J_p(R, S)$, where $p \leq m$ and $p \leq n$, is a relation of degree $m + n - p$ that consists of all $(m + n - p)$ -tuples $(a_1, a_2, \dots, a_{m-p}, c_1, c_2, \dots, c_p, b_1, b_2, \dots, b_{n-p})$, where the m-tuple $(a_1, a_2, \dots, a_{m-p}, c_1, c_2, \dots, c_p)$ belongs to R and the n-tuple $(c_1, c_2, \dots, c_p, b_1, b_2, \dots, b_{n-p})$ belongs to S.

14/09/2015

50/53

Databases and Relations

- In other words, to generate $J_p(R, S)$, we have to find all the elements in R whose p last components match the p first components of an element in S.

- The new relation contains exactly these matches, which are combined to tuples that contain each matching field only once.

14/09/2015

51/53

Databases and Relations

- **Example:** What is $J_1(Y, R)$, where Y contains the fields **Student Name** and **Year of Birth**,

- $Y = \{(1978, \text{Ackermann}), (1972, \text{Adams}), (1917, \text{Chou}), (1984, \text{Goodfriend}), (1982, \text{Rao}), (1970, \text{Stevens})\}$,

- and R contains the student records as defined before ?

14/09/2015

52/53

Databases and Relations

•**Solution:** The resulting relation is:

- $\{(1978, \text{Ackermann}, 231455, \text{CS}, 3.88),$
 $(1972, \text{Adams}, 888323, \text{Physics}, 3.45),$
 $(1917, \text{Chou}, 102147, \text{CS}, 3.79),$
 $(1984, \text{Goodfriend}, 453876, \text{Math}, 3.45),$
 $(1982, \text{Rao}, 678543, \text{Math}, 3.90),$
 $(1970, \text{Stevens}, 786576, \text{Psych}, 2.99)\}$

•Since Y has two fields and R has four, the relation $J_1(Y, R)$ has $2 + 4 - 1 = 5$ fields.