Section 2.1: Relations

Ordered pairs and n-tuples

We have seen that a set is a collection of elements in no particular order. However, if we wish to describe relationships between elements, we will often write them as ordered pairs, (*x,y*), such as

(3, -2) or

(oranges, $1.29/lb)

Just as in graphing, the ordered pair (3, -2) is not the same as (-2, 3).

An ordered triple then has three elements, (*x,y,z*), while the more general **n-tuple** is an ordered collection of many elements, such as

(*a, b, c, d*) or

(computer, Dell, 3.0 GHz, 512 MB RAM, Windows XP).

Relations

We can now define the terms relation and function.

relation = a set of ordered pairs or (more generally) **n-tuples**

Relations can then be defined by listing all of the pairs or n-tuples using set notation. Another way is to list the relation in a table.

Example

Are the following two relations equal?

{(truck, blue), (truck, red), (car, red), (SUV, green)}

vehicle	**colour**
truck	blue
truck	red
car	red
SUV	green

Answer: Yes, because both relations have the same set of ordered pairs.

Relations can also be defined by listing the two sets and the relationship between them.

Example

Let $x \in \{2, 4, 6\}$ and $y \in \{1, 2, 3\}$. $(x,y) \in$ A if $x + y \leq 5$. Find A.

(Here, the definition of A looks a little tricky. It is read aloud as "an ordered pair (x,y) is a member of A if $x + y \leq 5$." So you are being asked to find all sets of ordered pairs that satisfy the equation.)

Answer: A = {(2,1), (2,2), (2,3), (4,1)} or

x	y
2	1
2	2
2	3
4	1

Functions

In general, functions are defined using n-tuples. However, for this course we will only consider functions of ordered pairs.

function = a relation which has for each value of the first component, exactly one value of the second component

Example

Is the relation {(truck, blue), (truck, red), (car, red), (SUV, green)} also a function?

Answer: No. When the first component is a truck, there are two possible second components, blue and red.

Example

Let $x \in \{2, 4, 6\}$ and $y \in \{1, 2, 3\}$. $(x,y) \in$ A if $x/y = 2$. Is A a function?

Yes. A = {(2,1), (4,2), (6,3)}. For each x value, there is only one possible y value, so A is a function.

Example

Let $x, y \in$ real numbers. $(x, y) \in$ A if $y = 2x - 3$. Is A a function?

Answer: Yes. For every x, plugging it into $y = 2x - 3$ gives a unique value of y. If you prefer, you can graph the equation (it's a straight line) and use the "vertical line" test to see that it's a function.

Note: to use the "vertical line test", draw a graph of the relation. If a vertical line cuts the graph at only one point anywhere along the graph, the relation is a function.

Cartesian Product

If X and Y are both sets, then the Cartesian product X × Y is the set of all ordered pairs (x,y), where x is an element of set X and y is an element of set Y.

Example

If A = {a, b} and B = {0, 1}, then find A × B and B × A.

Answer: A × B = { (a, 0), (a, 1), (b, 0), (b, 1) }, and

B × A = {(0, a), (0, b), (1, a), (1, b)}.

Because order matters, notice that A × B ≠ B × A unless A = B.

Example

If A = {a, b} and B = {0, 1}, then find A × A and B × A × A.

Answer: A × A = { (a, a), (a, b), (b, a), (b, b) }, and

B×A×A = {(0,a,a), (0,a,b), (0,b,a), (0,b,b), (1,a,a), (1,a,b), (1,b,a), (1,b,b)}.

Functions and relations between two sets are then **subsets** of the Cartesian product.

Example

Let A = {1, 2} and B = {1, 2, 3}. Is A × A ⊆ B × A? Is A × A a function?

Answer:

A × A = {(1,1), (1,2), (2,1), (2,2)}

B × A = {(1,1), (1,2), (2,1), (2,2), (3,1), (3,2)}

Yes, $A \times A \subseteq B \times A$, because every member of $A \times A$ is also in $B \times A$, and there is at least one member of $B \times A$ that isn't in $A \times A$. For example, (3,1) isn't in $A \times A$.

No, $A \times A$ is not a function because for $x = 1$, there are two possible y's, 1 and 2.

Example

Let $X = \{1,2\}$ and $Y = \{1,3\}$. $(x,y) \in R$ if $x + y \geq 4$. Is R a function? Is $X \times X \subseteq X \times Y$?

Answer: $R = \{(1,3), (2,3)\}$. Yes, R is a function, because for each x there is only one y.

No, $X \times X$ is not a subset of $X \times Y$ because, for example, (1,2) is in $X \times X$ but not in $X \times Y$.

Section 2.1: Relations

Exercises

Write the relation as a set of ordered pairs.

1.

x	*y*
1	4
2	3
3	3
4	4

2.

fruit	*colour*
apple	red
orange	orange
banana	yellow
mango	green
peach	blue
apricot	purple

Write the relation as a table.

3. {(red,green},(green,red),(red,red),(green,green)}

4. {(Ford, 4 doors),(Honda, 4 doors), (Honda, 2 doors), (Saturn, 4 doors)}

Which of the following relations are functions?

5. The table in question 1.

6. The table in question 2.

7. The set in question 3.

8. The set in question 4.

Find the following relations and state whether each relation is also a function.

9. Let $x \in \{1,2\}$ and $y \in \{1, 2, 3\}$. $(x,y) \in$ A if x/y is an integer.

10. Let $x \in \{1,2\}$ and $y \in \{1, 2, 3\}$. $(x,y) \in$ A if $x < y$.

11. Let $x \in \{1, 2, 3\}$ and $y \in \{1, 2, 3, 4\}$. $(x,y) \in$ A if $x + y = 5$.

12. Let $x \in \{1, 2, 3, 4\}$ and $y \in \{1, 2, 3, 4\}$. $(x,y) \in$ A if $xy = 4$.

13. Let $x \in \{2, 4, 6\}$ and $y \in \{1, 2, 3\}$. $(x,y) \in$ A if $x - y = 3$.

14. Let $x \in \{2, 4, 6\}$ and $y \in \{1, 2, 3\}$. $(x,y) \in$ A if $x < y$.

Let A = {0,1}, B = {a,b}, and C = {α}. List the elements of the following Cartesian products.

15. C×C×C

16. A×C

17. B×C

18. A×A

19. B×B×A

20. A×B×C

Let A = {0,1}, B = {1,2}, and C = {0,1,2}. Are the following statements true or false? (Be sure to show your work.)

21. $A \times A \subset A \times C$

22. $A \times A \subseteq C \times A$

23. $B \times A \subseteq A \times B$

24. $C \times A \subset B \times C$

Section 2.1 – Answers

1. {(1,4),(2,3),(3,3),(4,4)}
2. {(apple,red),(orange,orange),(banana,yellow),(mango,green),(peach,blue),(apricot,purple)}
3.

red	green
green	red
red	red
green	green

4.

Ford	4 doors
Honda	4 doors
Honda	2 doors
Saturn	4 doors

5. Yes, it's a function.
6. Yes.
7. No, it's not a function. For example, "red" in the first coordinate has two possible second coordinates, "green" from the first row and "red" from the third.
8. No, Honda is repeated with two different "door" values.
9. A = {(1,1),{2,1),(2,2)}. It's not a function because when x is 2, there are two possible values for y.
10. The relation A is shown in the table below. No, it's not a function because when $x = 1$, there are two values for y.

x	y
1	2
1	3
2	3

11. A = {(1,4),(2,3),(3,2)}. Yes, it's a function.
12. The relation A is shown in the table below. Yes, it's a function because for each x-value, there is only one y-value.

x	y
1	4
2	2
4	1

13. A = {(4,1),(6,3)}. Yes, it's a function.

14. The relation A is shown in the table below. Yes, it's a function (if there's only one coordinate point, it's hard for it not to be!).

x	y
2	3

15. C×C×C = {(α,α,α)}
16. A×C = {(0, α),(1, α)}
17. B×C = {(a,α),(b,α)}
18. A×A = {(0,0),(0,1),(1,0),(1,1)}
19. B×B×A = {(a,a,0), (a,a,1), (a,b,0), (a,b,1), (b,a,0), (b,a,1), (b,b,0), (b,b,1)}
20. A×B×C = {(0,a,α),(0,b,α),(1,a,α),(1,b,α)}
21. True. A×A = {(0,0),(0,1),(1,0), (1,1)} and A×C = {(0,0),(0,1),(0,2), (1,0),(1,1),(1,2)}. Since every ordered pair in the first set is also in the second, yes, it's true.
22. True. A×A = {(0,0),(0,1),(1,0), (1,1)} and C×A = {(0,0),(0,1),(1,0), (1,1),(2,0),(2,1)}. Since every ordered pair in the first set is also in the second, yes, it's true.
23. False. B×A = {(1,0),(1,1),(2,0), (2,1)} and A×B = {(0,1),(0,2),(1,1), (1,2)}. In fact, most members of the first set aren't in the second.
24. False. C×A = {(0,0),(0,1),(1,0), (1,1),(2,0),(2,1)} and B×C = {(1,0),(1,1),(1,2), (2,0),(2,1),(2,2)} and, for example, the ordered pair (0,0) in C×A is not in B×C.

Section 2.2: Relational Databases

Databases

A database is a set of records that can be manipulated by a computer. Database management systems allow users of the system to perform a variety of operations, such as retrieving records or adding, changing, or deleting records. Most commercial database products are written in the computing language SQL, which is based on the **relational database model**. In this section, we will look at **relational algebra**, which is the collection of fundamental operations used to manipulate databases.

The set of records in a database is essentially a relation. Up until now, we have been looking at binary relations – relations that have two columns when written as a table. However, for many practical applications, we are interested in relations with many more than two columns. For example,

STUDENT

number	name	course	grade
000123	John Smith	Math 161	C+
000124	Jane Doe	Math 161	B+
000125	Richard Jones	Comp 112	A-
000126	John Smith	Comp 112	D

This relation can also be expressed as a set of n-tuples: STUDENT = {(000123, John Smith, Math 161, C+),(000124, Jane Doe, Math 161, B+),(000125, Richard Jones, Comp 112, A-), (000126, John Smith, Comp 112, B+)}. However, the table is much easier to read.

The rows of a relation are called records. Each record is a unique combination of the elements. For example, each record of the STUDENT relation given above corresponds to one individual student at Camosun College. One database in a bank might have each banking transaction (withdrawal, deposit, transfer) as a record, and another database in which each bank account was a record.

The columns of a relation are called attributes. For example, the attributes of the STUDENT relation are the number, name, course, and grade.

A key is a single attribute or combination of attributes that uniquely identifies a record. For example, in the STUDENT relation above, it would not be a good idea to try to use the student name as a key, since two students have the same name. Rather, the college assigns each student a unique student number as a key. (In the table above, we could also use the grade as a key, since in this very short table, no two students have the same grade.)

Example

Is your drivers licence number a key for the records kept by the Motor Vehicle Branch?

Answer: Yes, because each individual is assigned a unique number.

Example

Is your postal code a key for the post office's database of all possible street addresses?

Answer: No, because there is more than one street address for each postal code.

The idea of a key ties in very nicely with the idea of a function that we discussed in previous sections. For a binary relation, the relation is a function if for each value of the first component, there is only one value of the second component. Therefore, the first component can serve as a key to identify the ordered pair.

Queries

A query is a request for information from a database. In order to answer a query, the database management system has to perform one or more fundamental operations on the database. We will be looking at the operations Select, Project, and Join.

Select

The Select operator selects certain records (rows of the table, or n-tuples) from a relation. Looking once again at the STUDENT relation, the operation

$$\sigma_{\text{course = "Math 161"}}(\text{STUDENT})$$

will select the n-tuples (000123, John Smith, Math 161, C+) and (000124, Jane Doe, Math 161, B+).

Example

What is the result of the operation $\sigma_{\text{grade = "A-"}}(\text{STUDENT})$?

Answer: (000125, Richard Jones, Comp 112, A-)

Project

The Project operator selects attributes, or columns, from a relation. For example, the projection $\pi_{\text{name}}(\text{STUDENT})$ gives (John Smith), (Jane Doe), (Richard Jones), (John Smith). More than one column may be projected, as in the following example.

Example

What is the result of the operation $\pi_{name,grade}$(STUDENT)?

Answer: (John Smith, C+), (Jane Doe, B+), (Richard Jones, A-), (John Smith, D).

When you use the project operator in order to choose certain rows of a relation, you may end up with identical n-tuples. For example, suppose you had the relation

INVENTORY

ID	Item	Brand Name	Colour	In Stock
1	hockey stick	Cooper	black	3
2	football	Trooper	pink	5
3	football	Trooper	yellow	0
4	football	Cooper	yellow	2
5	snowboard	Nike	white	0

and you used the command $\pi_{Item,Colour}$ (INVENTORY). You should then get the relation

Item	Colour
hockey stick	black
football	pink
football	yellow
football	yellow
snowboard	white

Note that there are two instances of the ordered pair (football, yellow). According to set theory, the duplicate should be removed and the table collapsed to only four rows.

Item	Colour
hockey stick	black
football	pink
football	yellow
snowboard	white

This is how the Project operator would be handled according to relational algebra. As a practical matter, what real databases often do is give you either option – either preserving all of the original rows intact, or collapsing duplicate rows. SQL, for example, uses the keyword `distinct` to collapse duplicate rows.

Join

The Join operator (also called Natural Join) joins two relations. Suppose we had a second relation, EMAIL

number	email
000123	johnny@yahoo.com
000124	jane@telus.net
000125	rick@badboy.ca
000126	c000126@camosun.bc.ca

We wish to match the "number" column in the STUDENT relation with the "number" column in the EMAIL relation, and so the operation

$$\text{STUDENT} \bowtie \text{EMAIL}$$

will give the resulting relation

number	name	course	grade	email
000123	John Smith	Math 161	C+	johnny@yahoo.com
000124	Jane Doe	Math 161	B+	jane@telus.net
000125	Richard Jones	Comp 112	A-	rick@badboy.ca
000126	John Smith	Comp 112	D	c000126@camosun.bc.ca

Note that the two relations joined by the $\bowtie$ operator must share one common attribute (in this example, each relation has the column "number").

Sequences of Operations

Suppose the query which we wish to make requires more than one fundamental operations. Our approach then is to make a series of operations that will yield the desired result. The first operation will result in a new relation, which we then perform another operation on. We could either give that temporary relation a name to be specified in the next operation, or the more compact way is to nest two operations.

For example, consider the following two relations.

INVENTORY

ID	Item	Brand Name	Colour	In Stock
1	hockey stick	Cooper	black	3
2	football	Trooper	pink	5
3	football	Trooper	yellow	0
4	football	Cooper	yellow	2
5	snowboard	Nike	white	0

CONTACT INFO

Brand Name	Phone
Cooper	555-1212
Trooper	370-4542
Nike	370-3001

Suppose we wished to list the ID and phone number for all items. We will first have to join the two databases using INVENTORY $\bowtie$ CONTACT INFO, giving

ID	Item	Brand Name	Colour	In Stock	Phone
1	hockey stick	Cooper	black	3	555-1212
2	football	Trooper	pink	5	370-4542
3	football	Trooper	yellow	0	370-4542
4	football	Cooper	yellow	2	555-1212
5	snowboard	Nike	white	0	370-3001

And then we'll want to use the Project operator to choose the ID and Phone columns.

The sequence of operations would then look like

$\pi_{\text{ID, Phone}}$ (INVENTORY $\bowtie$ CONTACT INFO)

giving the relation

ID	Phone
1	555-1212
2	370-4542
3	370-4542
4	555-1212
5	370-3001

which is our desired result.

Please note that for this example, if we tried to do the Project operator first, to get rid of the unnecessary columns, that we would have to leave the Brand Name column in for the Join operation. Then we'd have to strip it out using *another* Project operator. So, it's simpler to do the Join first.

Example

For the relation INVENTORY, write an operation or sequence of operations that will list the ID number for all out-of-stock items.

Answer:

$$\pi_{\text{ID}} (\sigma_{\text{In Stock} = 0} (\text{INVENTORY}))$$

Section 2.2 Exercises

1. For a database of Camosun College students, would the student name be a key?
2. For a database of Camosun College students, would the student number be a key?
3. For a database of Canadian citizens, would the social insurance number be a key?
4. For a database of Canadian residents, would the person's street address be a key?
5. For a bank's database of bank accounts, would the name of the account holder be a key?
6. For a bank's database of bank transactions, could you use the time and date of the transaction as a key?
7. For a database of students currently taking Math 161, under what conditions could you use the student name as a key?
8. For a database of students currently taking Math 161, under what conditions could you use the students' email addresses as a key?

Use the following tables to answer questions 9 through 22.

Account

Account Number	Customer Name	Branch	Balance	Type
123	Abby	Hillside	$4000	Chequing
124	Ben	Hillside	$5000	Savings
125	Charlie	Douglas	$50	Chequing
126	Danielle	Hillside	$-100	Chequing
127	Ed	Cadboro Bay	$200	Savings
128	Ed	Cadboro Bay	$4000	Chequing

Address

Branch	Street Address
Hillside	100 Hillside Avenue
Douglas	1050 Douglas Street
Cadboro Bay	2000 Cadboro Bay Road

Activity

Account Number	Status
125	Dormant
126	Overdrawn
127	Dormant

9. Write an operation to select the records of all accounts at the Cadboro Bay Branch.
10. Write an operation to list the names of all the customers and the type of account they have.
11. Write an operation on the relation Activity to select the records all of the accounts that are overdrawn.
12. Write an operation on the relation Address to select the current street addresses.
13. Write an operation on the relation Account to list the account numbers and the branch they belong to.
14. Write an operation on the relation Address to select the record of the Hillside branch.
15. Write an operation to add the branch address to the records in the top table.
16. Write an operation to add the activity to the records in the top table.
17. Write one or more operations to list the names of the customers with dormant accounts.
18. Write one or more operations to list the balance associated with the overdrawn accounts.
19. Write one or more operations to list the names of all customers with chequing accounts.
20. Write one or more operations to list the balance of all of Ed's accounts.
21. Write one or more operations to list the account number and the street address of the branch that the account is at.
22. Write one or more operations to list the account number and the street address of all dormant accounts.

Section 2.2 – Answers

1. No, because more than one Camosun College student can have the same name.
2. Yes. Each number is unique to each student.
3. Yes, because the social insurance number is unique to each individual.
4. No, because many people can live at the same street address.
5. No, because a person can have more than one bank account.
6. No, because most banks are large enough that two transactions could conceivably go through at the same time.
7. You could use the student name as a key if you don't have two students with the same name.
8. You could if each person's email address is unique (no two students share the same email.).
9. $\sigma_{\text{Branch= "Cadboro Bay"}}$ (Account)
10. $\pi_{\text{Customer Name, Type}}$ (Account)
11. $\sigma_{\text{Status= "Overdrawn"}}$ (Activity)
12. $\pi_{\text{Street Address}}$ (Address)
13. $\pi_{\text{Account Number, Branch}}$ (Account)
14. $\sigma_{\text{Branch= "Hillside"}}$ (Address)
15. Account $\bowtie$ Address
16. Account $\bowtie$ Activity
17. $\pi_{\text{Customer Name}}$ ($\sigma_{\text{Status = "Dormant"}}$ (Account $\bowtie$ Activity))
18. π_{Balance} ($\sigma_{\text{Status = "Overdrawn"}}$ (Account $\bowtie$ Activity))
19. $\pi_{\text{Customer Name}}$($\sigma_{\text{Type = "Chequing"}}$ (Account))
20. π_{Balance} ($\sigma_{\text{Customer Name = "Ed"}}$ (Account))
21. $\pi_{\text{Account Number, Street Address}}$ (Account $\bowtie$ Address)
22. $\pi_{\text{Account Number, Street Address}}$ ($\sigma_{\text{Status = "Dormant"}}$ ((Account $\bowtie$ Address) $\bowtie$ Activity))