CHAPTER 8

RELATIONS

Alessandro Artale – UniBZ - http://www.inf.unibz.it/~artale/

SECTION 8.1

Relations on Sets

Relations on Sets

A formal way to denote a **binary relation** between two sets is to define it as a subset of the Cartesian product of the two sets.

In general, we can define an ***n-ary* relation** to be a subset of a Cartesian product of n sets, where n is any integer greater than or equal to two.

Such a relation is the fundamental structure used in relational databases. However, because we focus on binary relations, when we use the term *relation* by itself, we will mean binary relation.

Example 2 – *The Congruence Modulo 2 Relation*

Define a relation E from **Z** to **Z** as follows: For all $(m, n) \in \mathbf{Z} \times \mathbf{Z}$,

$$m \; E \; n \quad \Leftrightarrow \quad m - n \text{ is even.}$$

a. Is 4 E 0? Is 2 E 6? Is 3 E (–3)? Is 5 E 2?

b. List five integers that are related by E to 1.

c. Prove that if n is any odd integer, then n E 1.

Solution:

a. Yes, 4 E 0 because 4 – 0 = 4 and 4 is even.

Yes, 2 E 6 because 2 – 6 = –4 and –4 is even.

Example 2 – *Solution*

cont'd

Yes, 3 E (–3) because 3 – (–3) = 6 and 6 is even.

No, 5 $\not E$ 2 because 5 – 2 = 3 and 3 is not even.

b. There are many such lists. One is

1 because 1 – 1 = 0 is even,

3 because 3 – 1 = 2 is even,

5 because 5 – 1 = 4 is even,

–1 because –1 – 1 = –2 is even,

–3 because –3 – 1 = –4 is even.

Example 1 – *Solution*

cont'd

c. Proof:

Suppose n is any odd integer.

Then $n = 2k + 1$ for some integer k. Now by definition of E, $n \; E \; 1$ if, and only if, $n - 1$ is even.

But by substitution,

$$n - 1 = (2k + 1) - 1 = 2k,$$

and since k is an integer, $2k$ is even.

Hence $n \; E \; 1$ *[as was to be shown]*.

Example 1 – *Solution*

cont'd

It can be shown that integers *m* and *n* are related by *E* if, and only if, *m mod* 2 = *n mod* 2 (that is, both are even or both are odd).

When this occurs *m* and *n* are said to be **congruent modulo 2**.

The Inverse of a Relation

The Inverse of a Relation

If R is a relation from A to B, then a relation R^{-1} from B to A can be defined by interchanging the elements of all the ordered pairs of R.

• Definition

Let R be a relation from A to B. Define the inverse relation R^{-1} from B to A as follows:

$$R^{-1} = \{(y, x) \in B \times A \mid (x, y) \in R\}.$$

This definition can be written operationally as follows:

For all $x \in A$ and $y \in B$, $\quad (y, x) \in R^{-1} \quad \Leftrightarrow \quad (x, y) \in R.$

Example 4 – *The Inverse of a Finite Relation*

Let $A = \{2, 3, 4\}$ and $B = \{2, 6, 8\}$ and let R be the "divides" relation from A to B: For all $(x, y) \in A \times B$,

$$x \; R \; y \quad \Leftrightarrow \quad x \mid y \qquad\qquad x \text{ divides } y.$$

a. State explicitly which ordered pairs are in R and R^{-1}, and draw arrow diagrams for R and R^{-1}.

b. Describe R^{-1} in words.

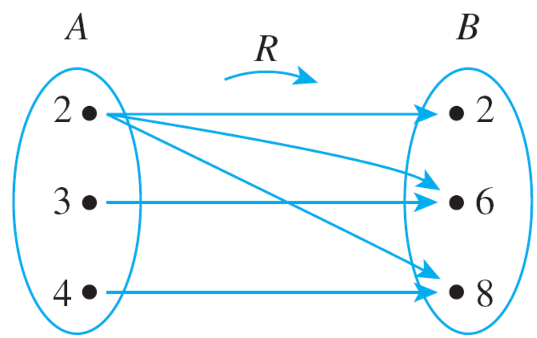
Solution:

a. $R = \{(2, 2), (2, 6), (2, 8), (3, 6), (4, 8)\}$

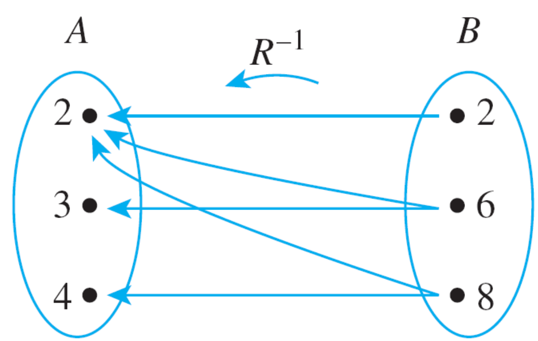
$R^{-1} = \{(2, 2), (6, 2), (8, 2), (6, 3), (8, 4)\}$

Example 4 – *Solution*

cont'd



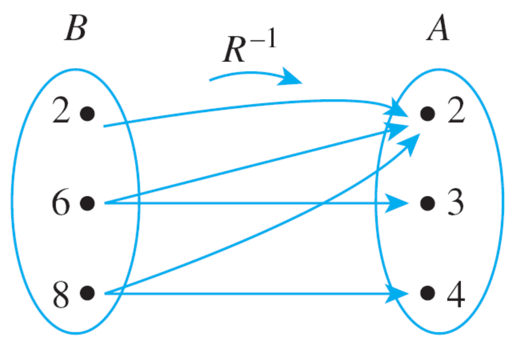
To draw the arrow diagram for R^{-1}, you can copy the arrow diagram for *R* but reverse the directions of the arrows.



Example 4 – *Solution*

cont'd

Or you can redraw the diagram so that *B* is on the left.



b. R^{-1} can be described in words as follows:
For all $(y, x) \in B \times A$,

$$y\, R^{-1}\, x \Leftrightarrow y \text{ is a multiple of } x.$$

Directed Graph of a Relation

Directed Graph of a Relation

• Definition

A **relation on a set** A is a relation from A to A.

When a relation R is defined *on* a set A, the arrow diagram of the relation can be modified so that it becomes a **directed graph**.

Instead of representing A as two separate sets of points, represent A only once, and draw an arrow from each point of A to each R-related point.

Directed Graph of a Relation

As with an ordinary arrow diagram,

> For all points x and y in A,
>
> there is an arrow from x to y $\Leftrightarrow$ $x\ R\ y$ $\Leftrightarrow$ $(x, y) \in R$.

If a point is related to itself, a loop is drawn that extends out from the point and goes back to it.

Example 6 – *Directed Graph of a Relation*

Let $A = \{3, 4, 5, 6, 7, 8\}$ and define a relation R on A as follows: For all $x, y \in A$,

$$x\ R\ y \Leftrightarrow 2 \mid (x - y).$$

Draw the directed graph of R.

Solution:

Note that $3\ R\ 3$ because $3 - 3 = 0$ and $2 \mid 0$ since $0 = 2 \cdot 0$. Thus there is a loop from 3 to itself.

Similarly, there is a loop from 4 to itself, from 5 to itself, and so forth, since the difference of each integer with itself is 0, and $2 \mid 0$.

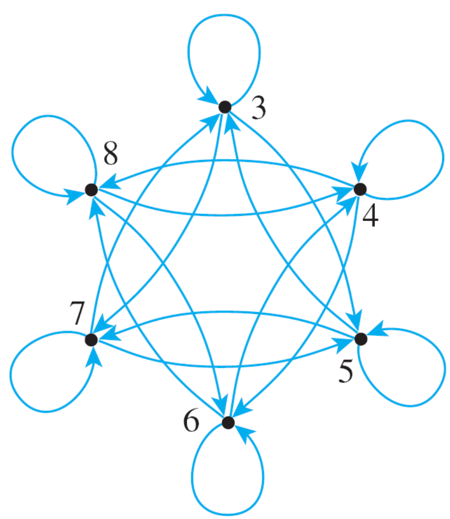
Example 6 – *Solution*

cont'd

Note also that 3 R 5 because $3 - 5 = -2 = 2 \cdot (-1)$. And 5 R 3 because $5 - 3 = 2 = 2 \cdot 1$.

Hence there is an arrow from 3 to 5 and also an arrow from 5 to 3.

The other arrows in the directed graph, as shown below, are obtained by similar reasoning.



N-ary Relations and Relational Databases

N-ary Relations and Relational Databases

N-ary relations form the mathematical foundation for relational database theory.

A binary relation is a subset of a Cartesian product of two sets, similarly, an *n-ary* relation is a subset of a Cartesian product of *n* sets.

- **Definition**

Given sets $A_1, A_2, \ldots, A_n$, an ***n*-ary relation** R on $A_1 \times A_2 \times \cdots \times A_n$ is a subset of $A_1 \times A_2 \times \cdots \times A_n$. The special cases of 2-ary, 3-ary, and 4-ary relations are called **binary, ternary,** and **quaternary relations,** respectively.

Example 7 – *A Simple Database*

The following is a radically simplified version of a database that might be used in a hospital.

Let A_1 be a set of positive integers, A_2 a set of alphabetic character strings, A_3 a set of numeric character strings, and A_4 a set of alphabetic character strings.

Define a quaternary relation R on $A_1 \times A_2 \times A_3 \times A_4$ as follows:

$(a_1, a_2, a_3, a_4) \in R \Leftrightarrow$ a patient with patient ID number a_1, named a_2, was admitted on date a_3, with primary diagnosis a_4.

Example 7 – *A Simple Database* cont'd

At a particular hospital, this relation might contain the following 4-tuples:

(011985, John Schmidt, 020710, asthma)

(574329, Tak Kurosawa, 0114910, pneumonia)

(466581, Mary Lazars, 0103910, appendicitis)

(008352, Joan Kaplan, 112409, gastritis)

(011985, John Schmidt, 021710, pneumonia)

(244388, Sarah Wu, 010310, broken leg)

(778400, Jamal Baskers, 122709, appendicitis)

Example 7 – *A Simple Database* cont’d

In discussions of relational databases, the tuples are normally thought of as being written in tables.

Each row of the table corresponds to one tuple, and the header for each column gives the descriptive attribute for the elements in the column.

Operations within a database allow the data to be manipulated in many different ways.

Example 7 – *A Simple Database* cont'd

For example, in the database language SQL, if the above database is denoted *S*, the result of the query

SELECT Patient_ID#, Name FROM *S* WHERE
Admission_Date = 010310

would be a list of the ID numbers and names of all patients admitted on 01-03-10:

466581 Mary Lazars,
244388 Sarah Wu.

Example 7 – *A Simple Database* cont'd

This is obtained by taking the intersection of the set $A_1 \times A_2 \times \{010310\} \times A_4$ with the database and then projecting onto the first two coordinates.

Similarly, SELECT can be used to obtain a list of all admission dates of a given patient.

For John Schmidt this list is

02-07-10   and

02-17-10

Example 7 – *A Simple Database* cont'd

Individual entries in a database can be added, deleted, or updated, and most databases can sort data entries in various ways.

In addition, entire databases can be merged, and the entries common to two databases can be moved to a new database.

SECTION 8.2

Reflexivity, Symmetry, and Transitivity

Reflexivity, Symmetry, and Transitivity

Let A = {2, 3, 4, 6, 7, 9} and define a relation R on A as follows: For all x, $y \in A$,

$$x\ R\ y \quad \Leftrightarrow \quad 3 \mid (x - y).$$

Then 2 R 2 because 2 – 2 = 0, and 3 | 0.

Similarly, 3 R 3, 4 R 4, 6 R 6, 7 R 7, and 9 R 9.

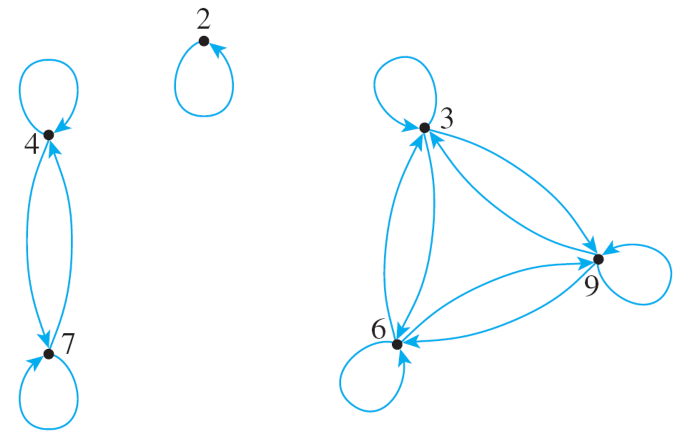
Also 6 R 3 because 6 – 3 = 3, and 3 | 3.

And 3 R 6 because 3 – 6 = –(6 – 3) = –3, and 3 | (–3).

Similarly, 3 R 9, 9 R 3, 6 R 9, 9 R 6, 4 R 7, and 7 R 4.

Reflexivity, Symmetry, and Transitivity

Thus the directed graph for *R* has the appearance shown at the right.



This graph has three important properties:

1. Each point of the graph has an arrow looping around from it back to itself.

2. In each case where there is an arrow going from one point to a second, there is an arrow going from the second point back to the first.

Reflexivity, Symmetry, and Transitivity

3. In each case where there is an arrow going from one point to a second and from the second point to a third, there is an arrow going from the first point to the third. That is, there are no "incomplete directed triangles" in the graph.

Properties (1), (2), and (3) correspond to properties of general relations called *reflexivity*, *symmetry*, and *transitivity*.

• Definition

Let R be a relation on a set A.

1. R is **reflexive** if, and only if, for all $x \in A$, $x \, R \, x$.
2. R is **symmetric** if, and only if, for all $x, y \in A$, ***if*** $x \, R \, y$ then $y \, R \, x$.
3. R is **transitive** if, and only if, for all $x, y, z \in A$, ***if*** $x \, R \, y$ and $y \, R \, z$ then $x \, R \, z$.

Reflexivity, Symmetry, and Transitivity

Because of the equivalence of the expressions $x \, R \, y$ and $(x, y) \in R$ for all x and y in A, the reflexive, symmetric, and transitive properties can also be written as follows:

1. R is reflexive $\Leftrightarrow$ for all x in A, $(x, x) \in R$.
2. R is symmetric $\Leftrightarrow$ for all x and y in A, ***if*** $(x, y) \in R$ then $(y, x) \in R$.
3. R is transitive $\Leftrightarrow$ for all x, y and z in A, ***if*** $(x, y) \in R$ and $(y, z) \in R$ then $(x, z) \in R$.

Example 1 – *Properties of Relations on Finite Sets*

Let $A = \{0, 1, 2, 3\}$ and define relations R, S, and T on A as follows:

$R = \{(0, 0), (0, 1), (0, 3), (1, 0), (1, 1), (2, 2), (3, 0), (3, 3)\},$

$S = \{(0, 0), (0, 2), (0, 3), (2, 3)\},$

$T = \{(0, 1), (2, 3)\}.$

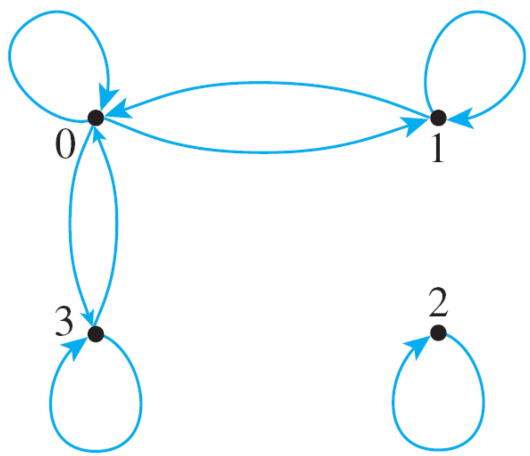
a. Is R reflexive? symmetric? transitive?

b. Is S reflexive? symmetric? transitive?

c. Is T reflexive? symmetric? transitive?

Example 1(a) – *Solution*

The directed graph of *R* has the appearance shown below.



***R* is reflexive:** There is a loop at each point of the directed graph. This means that each element of *A* is related to itself, so *R* is reflexive.

Example 1(a) – *Solution*

cont'd

***R* is symmetric:** In each case where there is an arrow going from one point of the graph to a second, there is an arrow going from the second point back to the first.

This means that whenever one element of A is related by R to a second, then the second is related to the first. Hence R is symmetric.

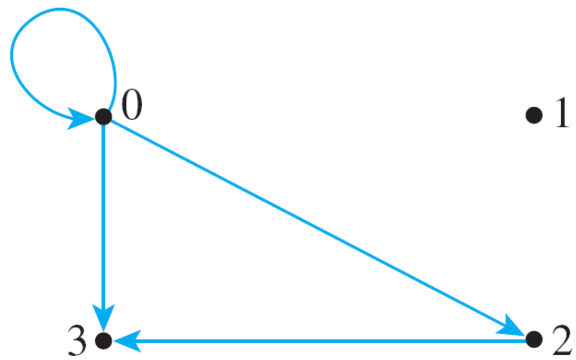
***R* is not transitive:** There is an arrow going from 1 to 0 and an arrow going from 0 to 3, but there is no arrow going from 1 to 3.

This means that there are elements of A—0, 1, and 3—such that $1\ R\ 0$ and $0\ R\ 3$ but $1 \not R\ 3$. Hence R is not transitive.

Example 1(b) – *Solution*

cont'd

The directed graph of S has the appearance shown below.



S is not reflexive: There is no loop at 1, for example. Thus $(1, 1) \notin S$, and so S is not reflexive.

S is not symmetric: There is an arrow from 0 to 2 but not from 2 to 0. Hence $(0, 2) \in S$ but $(2, 0) \notin S$, and so S is not symmetric.

Example 1(b) – *Solution*

cont'd

***S* is transitive:** There are three cases for which there is an arrow going from one point of the graph to a second and from the second point to a third:

Namely, there are arrows going from 0 to 2 and from 2 to 3; there are arrows going from 0 to 0 and from 0 to 2; and there are arrows going from 0 to 0 and from 0 to 3.

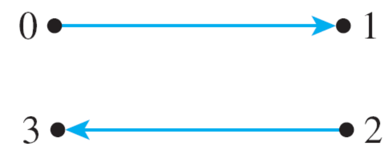
In each case there is an arrow going from the first point to the third. (Note again that the "first," "second," and "third" points need not be distinct.)

This means that whenever $(x, y) \in S$ and $(y, z) \in S$, then $(x, z) \in S$, for all $x, y, z \in \{0, 1, 2, 3\}$, and so S is transitive.

Example 1(c) – *Solution*

cont'd

The directed graph of T has the appearance shown at right.



***T* is not reflexive:** There is no loop at 0, for example. Thus $(0, 0) \notin T$, so T is not reflexive.

***T* is not symmetric:** There is an arrow from 0 to 1 but not from 1 to 0. Thus $(0, 1) \in T$ but $(1, 0) \notin T$, and so T is not symmetric.

***T* is transitive:** The transitivity condition is vacuously true for T. To see this, observe that the transitivity condition says that

For all $x, y, z \in A$, if $(x, y) \in T$ and $(y, z) \in T$ then $(x, z) \in T$.

Example 1(c) – *Solution*

cont'd

The only way for this to be false would be for there to exist elements of A that make the hypothesis true and the conclusion false.

That is, there would have to be elements x, y, and z in A such that $(x, y) \in T$ and $(y, z) \in T$ and $(x, z) \notin T$.

In other words, there would have to be two ordered pairs in T that have the potential to "link up" by having the *second* element of one pair be the *first* element of the other pair.

But the only elements in T are (0, 1) and (2, 3), and these do not have the potential to link up. Hence the hypothesis is never true. It follows that it is impossible for T *not* to be transitive, and thus T is transitive.

Properties of Relations on Infinite Sets

Properties of Relations on Infinite Sets

Suppose a relation R is defined on an infinite set A. To prove the relation is reflexive, symmetric, or transitive, first write down what is to be proved.

For instance, for symmetry you need to prove that

$$\forall\, x, y \in A, \text{ if } x \mathrel{R} y \text{ then } y \mathrel{R} x.$$

Then use the definitions of A and R to rewrite the statement for the particular case in question. For instance, for the "equality" relation on the set of real numbers, the rewritten statement is

$$\forall\, x, y \in \mathbf{R}, \text{ if } x = y \text{ then } y = x.$$

Properties of Relations on Infinite Sets

Sometimes the truth of the rewritten statement will be immediately obvious (as it is here).

At other times you will need to prove it using the method of generalizing from the generic particular.

We begin with the relation of equality, one of the simplest and yet most important relations.

Example 2 – *Properties of Equality*

Define a relation *R* on **R** (the set of all real numbers) as follows: For all real numbers *x* and *y*.

$$\boxed{x \; R \; y \quad \Leftrightarrow \quad x = y.}$$

a. Is *R* reflexive?

b. Is *R* symmetric?

c. Is *R* transitive?

Example 2(a) – *Solution*

***R* is reflexive:** R is reflexive if, and only if, the following statement is true:

$$\text{For all } x \in \mathbf{R},\ x\, R\, x.$$

Since $x\,R\,x$ just means that $x = x$, this is the same as saying

$$\text{For all } x \in \mathbf{R},\ x = x.$$

But this statement is certainly true; every real number is equal to itself.

Example 2(b) – *Solution*

cont'd

***R* is symmetric:** R is symmetric if, and only if, the following statement is true:

For all $x, y \in \mathbf{R}$, if $x\ R\ y$ then $y\ R\ x$.

By definition of R, $x\ R\ y$ means that $x = y$ and $y\ R\ x$ means that $y = x$. Hence R is symmetric if, and only if,

For all $x, y \in \mathbf{R}$, if $x = y$ then $y = x$.

But this statement is certainly true; if one number is equal to a second, then the second is equal to the first.

Example 2(c) – *Solution*

cont'd

***R* is transitive:** R is transitive if, and only if, the following statement is true:

For all $x, y, z \in \mathbf{R}$, if $x\ R\ y$ and $y\ R\ z$ then $x\ R\ z$.

By definition of R, $x\ R\ y$ means that $x = y$, $y\ R\ z$ means that $y = z$, and $x\ R\ z$ means that $x = z$. Hence R is transitive if, and only if, the following statement is true:

For all $x, y, z \in \mathbf{R}$, if $x = y$ and $y = z$ then $x = z$.

But this statement is certainly true: If one real number equals a second and the second equals a third, then the first equals the third.

Example 4 – *Properties of Congruence Modulo 3*

Define a relation *T* on **Z** (the set of all integers) as follows: For all integers *m* and *n*,

$$\boxed{m \; T \; n \quad \Leftrightarrow \quad 3 \mid (m - n).}$$

This relation is called **congruence modulo 3.**

a. Is *T* reflexive?

b. Is *T* symmetric?

c. Is *T* transitive?

Example 4(a) – *Solution*

***T* is reflexive:** To show that T is reflexive, it is necessary to show that

$$\text{For all } m \in \mathbf{Z}, \qquad m\ T\ m.$$

By definition of T, this means that

$$\text{For all } m \in \mathbf{Z}, \qquad 3 \mid (m - m).$$

Or, since $m - m = 0$, $\qquad$ For all $m \in \mathbf{Z}$, $\quad 3 \mid 0$.

But this is true: $3 \mid 0$ since $0 = 3 \cdot 0$. Hence T is reflexive. This reasoning is formalized in the following proof.

Proof of Reflexivity: Suppose m is a particular but arbitrarily chosen integer. *[We must show that $m\ T\ m$.]* Now $m - m = 0$. But $3 \mid 0$ since $0 = 3 \cdot 0$. Hence $3 \mid (m - m)$. Thus, by definition of T, $m\ T\ m$ *[as was to be shown]*.

Example 4(b) – *Solution*

cont'd

***T* is symmetric:** To show that T is symmetric, it is necessary to show that

$$\text{For all } m, n \in \mathbf{Z}, \text{ if } m \; T \; n \text{ then } n \; T \; m.$$

By definition of T this means that

$$\text{For all } m, n \in \mathbf{Z}, \text{ if } 3 \mid (m - n) \text{ then } 3 \mid (n - m).$$

Is this true? Suppose m and n are particular but arbitrarily chosen integers such that $3 \mid (m - n)$.

Must it follow that $3 \mid (n - m)$? *[In other words, can we find an integer so that $n - m = 3 \cdot$ (that integer)?]*

Example 4(b) – *Solution*

cont'd

By definition of "divides," since

$$3 \mid (m - n),$$

then $\quad m - n = 3k \quad$ for some integer k.

The crucial observation is that $n - m = -(m - n)$. Hence, you can multiply both sides of this equation by –1 to obtain

$$-(m - n) = -3k,$$

which is equivalent to $\quad n - m = 3(-k)$.

[Thus we have found an integer, namely $-k$*, so that* $n - m = 3 \cdot$ *(that integer).]*

Example 4(b) – *Solution*

cont'd

Since $-k$ is an integer, this equation shows that

$$3 \mid (n - m).$$

It follows that T is symmetric.

The reasoning is formalized in the following proof.

Proof of Symmetry: Suppose m and n are particular but arbitrarily chosen integers that satisfy the condition $m\ T\ n$. *[We must show that $n\ T\ m$.]* By definition of T, since $m\ T\ n$ then $3 \mid (m - n)$. By definition of "divides," this means that $m - n = 3k$, for some integer k. Multiplying both sides by -1 gives $n - m = 3(-k)$. Since $-k$ is an integer, this equation shows that $3 \mid (n - m)$. Hence, by definition of T, $n\ T\ m$ *[as was to be shown]*.

Example 4(c) – *Solution*

cont' d

***T* is transitive:** To show that T is transitive, it is necessary to show that

For all $m, n, p \in \mathbf{Z}$, if $m\ T\ n$ and $n\ T\ p$ then $m\ T\ p$.

By definition of T this means that

For all $m, n \in \mathbf{Z}$,
if $3 \mid (m - n)$ and $3 \mid (n - p)$ then $3 \mid (m - p)$.

Is this true? Suppose m, n, and p are particular but arbitrarily chosen integers such that $3 \mid (m - n)$ and $3 \mid (n - p)$.

Must it follow that $3 \mid (m - p)$? *[In other words, can we find an integer so that* $m - p = 3 \cdot$ (*that integer*)*?]*

Example 4(c) – *Solution*

cont' d

By definition of "divides," since

$$3 \mid (m - n) \text{ and } 3 \mid (n - p),$$

then $\qquad m - n = 3r \qquad$ for some integer r,

and $\qquad n - p = 3s \qquad$ for some integer s.

The crucial observation is that $(m - n) + (n - p) = m - p$.

Add these two equations together to obtain

$$(m - n) + (n - p) = 3r + 3s,$$

which is equivalent to $m - p = 3(r + s)$.

[Thus we have found an integer so that $m - p = 3 \cdot ($*that integer*$)$*.]*

Example 4(c) – *Solution*

cont'd

Since r and s are integers, $r + s$ is an integer. So this equation shows that

$$3 \mid (m - p).$$

It follows that T is transitive.

The reasoning is formalized in the following proof.

Proof of Transitivity: Suppose m, n, and p are particular but arbitrarily chosen integers that satisfy the condition $m\ T\ n$ and $n\ T\ p$. *[We must show that $m\ T\ p$.]* By definition of T, since $m\ T\ n$ and $n\ T\ p$, then $3 \mid (m - n)$ and $3 \mid (n - p)$. By definition of "divides," this means that $m - n = 3r$ and $n - p = 3s$, for some integers r and s. Adding the two equations gives $(m - n) + (n - p) = 3r + 3s$, and simplifying gives that $m - p = 3(r + s)$. Since $r + s$ is an integer, this equation shows that $3 \mid (m - p)$. Hence, by definition of T, $m\ T\ p$ *[as was to be shown]*.

The Transitive Closure of a Relation

The Transitive Closure of a Relation

Generally speaking, a relation fails to be transitive because it fails to contain certain ordered pairs.

For example, if (1, 3) and (3, 4) are in a relation R, then the pair (1, 4) *must* be in R if R is to be transitive.

To obtain a transitive relation from one that is not transitive, it is necessary to add ordered pairs.

Roughly speaking, the relation obtained by adding the least number of ordered pairs to ensure transitivity is called the ***transitive closure*** of the relation.

The Transitive Closure of a Relation

In a sense made precise by the formal definition, the transitive closure of a relation is the smallest transitive relation that contains the relation.

• Definition

Let A be a set and R a relation on A. The **transitive closure** of R is the relation R^t on A that satisfies the following three properties:

1. R^t is transitive.
2. $R \subseteq R^t$.
3. If S is any other transitive relation that contains R, then $R^t \subseteq S$.

Example 5 – *Transitive Closure of a Relation*

Let $A = \{0, 1, 2, 3\}$ and consider the relation R defined on A as follows:

$$R = \{(0, 1), (1, 2), (2, 3)\}.$$

Find the transitive closure of R.

Solution:

Every ordered pair in R is in R^t, so

$$\{(0, 1), (1, 2), (2, 3)\} \subseteq R^t.$$

Thus the directed graph of R^t contains the arrows shown at the right.

Example 5 – *Solution*

cont' d

Since there are arrows going from 0 to 1 and from 1 to 2, R^t must have an arrow going from 0 to 2.

Hence $(0, 2) \in R^t$. Then $(0, 2) \in R^t$ and $(2, 3) \in R^t$, so since R^t is transitive, $(0, 3) \in R^t$.

Also, since $(1, 2) \in R^t$ and $(2, 3) \in R^t$, then $(1, 3) \in R^t$.

Thus R^t contains at least the following ordered pairs:

$$\{(0, 1), (0, 2), (0, 3), (1, 2), (1, 3), (2, 3)\}.$$

But this relation is transitive; hence it equals R^t. Note that the directed graph of R^t is as shown at the right.