Discrete Mathematics, KOM1062
Lecture #5

Instructor: Dr. Yavuz Eren

Lecture Book: **"Discrete Mathematics, Seventh Edt., Kenneth H. Rosen, 2007, McGraw Books Discrete Mathematics and Applications, Susanna S. Epp, Brooks, 4th Edt., 2011".**

Spring 2024

Relations

Relationships between elements of sets occur in many contexts. Every day we deal with relationships such as those between a business and its telephone number, an employee and his or her salary, a person and a relative, and so on.

In mathematics we study relationships such as those between a positive integer and one that it divides, an integer and one that it is congruent to modulo 5, a real number and one that is larger than it, a real number *x and the* value *f(x) where f is a function, and so on. Relationships such as that between a program and* a variable it uses, and that between a computer language and a valid statement in this language often arise in computer science.

Relationships between elements of sets are represented using the structure called a relation, which is just a subset of the Cartesian product of the sets. Relations can be used to solve problems such as determining which pairs of cities are linked by airline flights in a network, finding a viable order for the different phases of a complicated project, or producing a useful way to store information in computer databases.

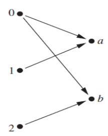
The most direct way to express a relationship between elements of two sets is to use ordered pairs made up of two related elements. For this reason, sets of ordered pairs are called binary relations.

Definition: Let *A and B be sets. A binary relation from A to B is a subset of A × B.*

A binary relation from *A to B is a set R of ordered pairs where the first element* of each ordered pair comes from *A and the second element comes from B. We use the notation a R b to denote that (a, b) ∈ R and "a R̸ b" to denote that (a,b) ∉ R. Moreover, when (a, b)* belongs to *R, a is said to be* related to b by R.

Ex. 1: Let *A be the set of students in your school, and let B be the set of courses. Let R be* the relation that consists of those pairs *(a,b), where a is a student enrolled in course b.* For instance, if Jason Goodfriend and Deborah Sherman are enrolled in CS518, the pairs (Jason Goodfriend, CS518) and (Deborah Sherman, CS518) belong to *R. If Jason Goodfriend* is also enrolled in CS510, then the pair (Jason Goodfriend, CS510) is also in *R. However,* if Deborah Sherman is not enrolled in CS510, then the pair (Deborah Sherman, CS510) is not in *R.*

Ex. 3 : Let *A = {0, 1, 2} and B = {a, b}. Then {(0, a), (0, b), (1, a), (2, b)} is a relation from A to B.* This means, for instance, that 0 *R a, but that 1 R̸ b. Relations can be represented graphically,* as shown in figure below, using arrows to represent ordered pairs. Another way to represent this relation is to use a table, which is also shown as



R	a	b
0	×	×
1	×	
2		×

Definition: A *relation on a set A is a relation from A to A(*Namely, a relation on a set *A is a subset of A × A).*

Ex. 5: Consider these relations on the set of integers:

R1 = {(a, b) | a ≤ b},
R2 = {(a, b) | a > b},
R3 = {(a, b) | a = b or a = −b},
R4 = {(a, b) | a = b},
R5 = {(a, b) | a = b + 1},
R6 = {(a, b) | a + b ≤ 3}.

Which of these relations contain each of the pairs *(1, 1), (1, 2), (2, 1), (1,−1), and (2, 2)?*

Solution: The pair *(*1, 1*)* is in R_1, R_3, R_4, and R_6;
*(*1, 2*)* is in R_1 and R_6;
*(*2, 1*)* is in R_2, R_5 and R_6;
*(*1,−1*)* is in R_2, R_3, and R_6;
*(*2, 2*)* is in R_1, R_3, and R_4.

Ex. 6: How many relations are there on a set with *n elements?*

Solution: A relation on a set *A* is a subset of *A* × *A*. Because *A* × *A* has n^2 elements when *A* has *n* elements, and a set with *m* elements has 2^m subsets, there are $2^{(n^2)}$ subsets of *A* × *A*. Thus, there are $2^{(n^2)}$ relations on a set with *n* elements. For example, there are $2^{(3^2)} = 2^9 = 512$ relations on the set {*a, b, c*}.

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

Ex. 7: Consider the following relations on {1, 2, 3, 4}:
R1 = {(1, 1), (1, 2), (2, 1), (2, 2), (3, 4), (4, 1), (4, 4)},
R2 = {(1, 1), (1, 2), (2, 1)},
R3 = {(1, 1), (1, 2), (1, 4), (2, 1), (2, 2), (3, 3), (4, 1), (4, 4)},
R4 = {(2, 1), (3, 1), (3, 2), (4, 1), (4, 2), (4, 3)},
R5 = {(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (3, 3), (3, 4), (4, 4)},
R6 = {(3, 4)}.
Which of these relations are reflexive?

Solution:
The relations R_3 and R_5 are reflexive because they both contain all pairs of the form *(a, a)*, namely, *(1, 1)*, *(2, 2)*, *(3, 3)*, and *(4, 4)*.

The other relations are not reflexive because they do not contain all of these ordered pairs. In particular, R_1, R_2, R_4, and R_6 are not reflexive because *(3, 3)* is not in any of these relations.

Ex. 9: Is the "divides" relation on the set of positive integers reflexive?

Solution:
Because $a \mid a$ whenever *a* is a positive integer, the "divides" relation is reflexive. (Note that if we replace the set of positive integers with the set of all integers the relation is not reflexive because by definition 0 does not divide 0.)

Definition: 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 such that for all a, b* $\in A$*, if (a, b)* $\in R$ *and (b, a)* $\in R$*, then a = b* is called *antisymmetric.*

Remark: Using quantifiers, we see that the relation R on the set A is symmetric if

$$\forall a \forall b\, ((a, b) \in R \rightarrow (b, a) \in R).$$

Similarly, the relation R on the set A is antisymmetric if

$$\forall a \forall b(((a, b) \in R \land (b, a) \in R) \rightarrow (a = b)).$$

That is, a relation is symmetric if and only if a is related to b implies that b is related to a.

A relation is antisymmetric if and only if there are no pairs of distinct elements a and b with a related to b and b related to a. That is, the only way to have a related to b and b related to a is for a and b to be the same element.

The terms symmetric and antisymmetric are not opposites, because a relation can have both of these properties or may lack both of them. A relation cannot be both symmetric and antisymmetric if it contains some pair of the form (a, b), where a ≠ b.

Ex. 9: Consider the following relations on {1, 2, 3, 4}:
R1 = {(1, 1), (1, 2), (2, 1), (2, 2), (3, 4), (4, 1), (4, 4)},
R2 = {(1, 1), (1, 2), (2, 1)},
R3 = {(1, 1), (1, 2), (1, 4), (2, 1), (2, 2), (3, 3), (4, 1), (4, 4)},
R4 = {(2, 1), (3, 1), (3, 2), (4, 1), (4, 2), (4, 3)},
R5 = {(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (3, 3), (3, 4), (4, 4)},
R6 = {(3, 4)}.
Which of these relations are symmetric and which are antisymmetric?

Solution:
The relations R_2 and R_3 are symmetric, because in each case *(b, a)* belongs to the relation whenever *(a, b)* does.

For R_2, the only thing to check is that both *(2, 1)* and *(1, 2)* are in the relation. For *R*3, it is necessary to check that both *(1, 2)* and *(2, 1)* belong to the relation, and *(1, 4)* and *(4, 1)* belong to the relation.

The reader should verify that none of the other relations is symmetric. This is done by finding a pair *(a, b)* such that it is in the relation but *(b, a)* is not.

*R*4, *R*5, and *R*6 are all antisymmetric. For each of these relations there is no pair of elements *a* and *b* with $a = b$ such that both *(a, b)* and *(b, a)* belong to the relation. The reader should verify that none of the other relations is antisymmetric. This is done by finding a pair *(a, b)* with $a = b$ such that *(a, b)* and *(b, a)* are both in the relation.

Note that, R1 is not symmetric nor antisymmetric.

Ex. 10 : Consider these relations on the set of integers:
$R1 = \{(a, b) \mid a \leq b\}$,
$R2 = \{(a, b) \mid a > b\}$,
$R3 = \{(a, b) \mid a = b \text{ or } a = -b\}$,
$R4 = \{(a, b) \mid a = b\}$,
$R5 = \{(a, b) \mid a = b + 1\}$,
$R6 = \{(a, b) \mid a + b \leq 3\}$.
Which of theses relations are symmetric and which are antisymmetric?

Solution:
The relations R_3, R_4, and R_6 are symmetric.
R_3 is symmetric, for if $a = b$ or $a = -b$, then $b = a$ or $b = -a$.
R_4 is symmetric because $a = b$ implies that $b = a$.
R_6 is symmetric because $a + b \leq 3$ implies that $b + a \leq 3$.
The reader should verify that none of the other relations is symmetric.

The relations R_1, R_2, R_4, and R_5 are antisymmetric.
R_1 is antisymmetric because the inequalities $a \leq b$ and $b \leq a$ imply that $a = b$.
R_2 is antisymmetric because it is impossible that $a > b$ and $b > a$.
R_4 is antisymmetric, because two elements are related with respect to R_4 if and only if they are equal.
R_5 is antisymmetric because it is impossible that $a = b + 1$ and $b = a + 1$.
The reader should verify that none of the other relations is antisymmetric.

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 all* $a, b, c \in A$.

Remark: Using quantifiers we see that the relation R on a set A is transitive if we have $\forall a \forall b \forall c(((a, b) \in R \wedge (b, c) \in R) \rightarrow (a, c) \in R)$.

Ex. 11:
$R1 = \{(a, b) \mid a \leq b\}$,
$R2 = \{(a, b) \mid a > b\}$,
$R3 = \{(a, b) \mid a = b \text{ or } a = -b\}$,
$R4 = \{(a, b) \mid a = b\}$,
$R5 = \{(a, b) \mid a = b + 1\}$,
$R6 = \{(a, b) \mid a + b \leq 3\}$.
Which of the relations above are transitive?

Solution:
The relations R_1, R_2, R_3, and R_4 are transitive.
$R1$ is transitive because $a \leq b$ and $b \leq c$ imply that $a \leq c$.
$R2$ is transitive because $a > b$ and $b > c$ imply that $a > c$.
$R3$ is transitive because $a = \pm b$ and $b = \pm c$ imply that $a = \pm c$.
$R4$ is clearly transitive, as the reader should verify.
$R5$ is not transitive because $(2, 1)$ and $(1, 0)$ belong to $R5$, but $(2, 0)$ does not.
$R6$ is not transitive because $(2, 1)$ and $(1, 2)$ belong to $R6$, but $(2, 2)$ does not.

Ex. 12: Is the "divides" relation on the set of positive integers transitive?

Solution:
Suppose that a divides b and b divides c. Then there are positive integers k and l such that $b = ak$ and $c = bl$. Hence, $c = a(kl)$, so a divides c. It follows that this relation is transitive.

We can use counting techniques to determine the number of relations with specific properties. Finding the number of relations with a particular property provides information about how common this property is in the set of all relations on a set with n elements.

Ex. 16: How many reflexive relations are there on a set with n elements?
Solution: A relation R on a set A is a subset of $A \times A$. Consequently, a relation is determined by specifying whether each of the n^2 ordered pairs in $A \times A$ is in R. However, if R is reflexive, each of the n ordered pairs (a, a) for $a \in A$ must be in R. Each of the other $n(n - 1)$ ordered pairs of the form (a, b), where $a \neq b$, may or may not be in R. Hence, by the product rule for counting, there are $2^{n(n-1)}$ reflexive relations.

Combining Relations: Because relations from *A to B are subsets of A × B, two relations from A to B can be combined* in any way two sets can be combined.

Ex. 13 : Let *A = {1, 2, 3} and B = {1, 2, 3, 4}.*

The relations
R1 = {(1, 1), (2, 2), (3, 3)} and
R2 = {(1, 1), (1, 2), (1, 3), (1, 4)} can be combined to obtain
R1 ∪ R2 = {(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (3, 3)},
R1 ∩ R2 = {(1, 1)},
R1 − R2 = {(2, 2), (3, 3)},
R2 − R1 = {(1, 2), (1, 3), (1, 4)}.

Ex. 14: Let A and B be the set of all students and the set of all courses at a school, respectively. Suppose that R_1 consists of all ordered pairs (a, b), where a is a student who has taken course b, and R_2 consists of all ordered pairs (a, b), where a is a student who requires course b to graduate.

What are the relations $R_1 \cup R_2$, $R_1 \cap R_2$, $R_1 \oplus R_2$, $R_1 - R_2$, and $R_2 - R_1$?

Solution:
The relation R1 ∪ R2 consists of all ordered pairs (a, b), where a is a student who either has taken course b or needs course b to graduate.

R1 ∩ R2 is the set of all ordered pairs (a, b), where a is a student who has taken course b and needs this course to graduate.

R1 ⊕ R2 consists of all ordered pairs (a, b), where student a has taken course b but does not need it to graduate or needs course b to graduate but has not taken it.

R1 − R2 is the set of ordered pairs (a, b), where a has taken course b but does not need it to graduate; that is, b is an elective course that a has taken.

R2 − R1 is the set of all ordered pairs (a, b), where b is a course that a needs to graduate but has not taken.

Ex. 15: Let R_1 *be the "less than" relation on the set of real numbers and let* R_2 *be the "greater than"* relation on the set of real numbers, that is, $R_1 = \{(x, y) \mid x < y\}$ *and* $R_2 = \{(x, y) \mid x > y\}$. What are $R_1 \cup R_2$, $R_1 \cap R_2$, $R_1 - R_2$, $R_2 - R_1$, *and* $R_1 \oplus R_2$?

Solution: We note that $(x, y) \in R_1 \cup R_2$ if and only if $(x, y) \in R_1$ or $(x, y) \in R_2$. Hence, $(x, y) \in R1 \cup R2$ if and only if $x < y$ or $x > y$. Because the condition $x < y$ or $x > y$ is the same as the condition $x = y$, it follows that $R_1 \cup R_2 = \{(x, y) \mid x = y\}$. In other words, the union of the "less than" relation and the "greater than" relation is the "not equals" relation.

Next, note that it is impossible for a pair (x, y) to belong to both R_1 and R_2 because it is impossible that $x < y$ and $x > y$. It follows that $R_1 \cap R_2 = \emptyset$.

We also see that $R_1 - R_2 = R_1$,
$R_2 - R_1 = R_2$, and
$R_1 \oplus R_2 = R_1 \cup R_2 - R_1 \cap R_2 = \{(x, y) \mid x = y\}$.

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* $S \circ R$.

Ex. 16: What is the composite of the relations *R and S, where*

R is the relation from {1, 2, 3} to {1, 2, 3, 4} with $R = \{(1, 1), (1, 4), (2, 3), (3, 1), (3, 4)\}$ *and*
S is the relation from {1, 2, 3, 4} to {0, 1, 2} with $S = \{(1, 0), (2, 0), (3, 1), (3, 2), (4, 1)\}$?

Solution: $S \circ R$ is constructed using all ordered pairs in *R* and ordered pairs in *S*, where the second element of the ordered pair in *R* agrees with the first element of the ordered pair in *S*.

For example, the ordered pairs $(2, 3)$ in *R* and $(3, 1)$ in *S* produce the ordered pair $(2, 1)$ in $S \circ R$.

Computing all the ordered pairs in the composite, we find $S \circ R = \{(1, 0), (1, 1), (2, 1), (2, 2), (3, 0), (3, 1)\}$.

Definition: Let *R be a relation on the set A. The powers* R^n, *n = 1, 2, 3, . . . , are defined recursively by* $R^1 = R$ *and* $R^{n+1} = R^n \circ R$ *(*The definition shows that $R^2 = R \circ R$, $R^3 = R^2 \circ R = (R \circ R) \circ R$, *and so on).*

Ex. 17: Let $R = \{(1, 1), (2, 1), (3, 2), (4, 3)\}$. *Find the powers* R^n, $n = 2, 3, 4, \ldots$.

Solution: Because $R^2 = R \circ R$, we find that $R^2 = \{(1, 1), (2, 1), (3, 1), (4, 2)\}$.

Furthermore, because $R^3 = R^2 \circ R$, $R^3 = \{(1, 1), (2, 1), (3, 1), (4, 1)\}$.

Additional computation shows that R^4 is the same as R^3, so $R^4 = \{(1, 1), (2, 1), (3, 1), (4, 1)\}$.

It also follows that $R^n = R^3$ for $n = 5, 6, 7, \ldots$. The reader should verify this.

Theorem: The relation *R on a set A is transitive if and only if* $R^n \subseteq R$ *for* $n = 1, 2, 3, \ldots$.

n-ary Relations and Their Applications

Relationships among elements of more than two sets often arise. For instance, there is a relationship involving the name of a student, the student's major, and the student's grade point average. Similarly, there is a relationship involving the airline, flight number, starting point, destination, departure time, and arrival time of a flight.

An example of such a relationship in mathematics involves three integers, where the first integer is larger than the second integer, which is larger than the third. Another example is the between ness relationship involving points on a line, such that three points are related when the second point is between the first and the third.

Definition: Let $A_1, A_2, \ldots, A_n$ *be sets. An n-ary relation on these sets is a subset of* $A_1 \times A_2 \times \cdots \times A_n$. The sets $A_1, A_2, \ldots, A_n$ *are called the domains of the relation, and n is called its degree.*

Ex.: Let *R be the relation on N × N × N consisting of triples (a, b, c), where a, b, and c are integers* with $a < b < c$. *Then* $(1, 2, 3) \in R$, *but* $(2, 4, 3)$ *is* $\not R$. *The degree of this relation is 3. Its domains* are all equal to the set of natural numbers.

Ex. 2: Let *R be the relation consisting of 5-tuples (A,N,S,D,T) representing airplane flights,* where *A is the airline, N is the flight number, S is the starting point, D is the destination, and T is* the departure time. For instance, if Nadir Express Airlines has flight 963 from Newark to Bangor at 15:00, then (Nadir, 963, Newark, Bangor, 15:00) belongs to *R. The degree of this relation* is 5, and its domains are the set of all airlines, the set of flight numbers, the set of cities, the set of cities, and the set of times.

Databases and Relations

A database consists of records, which are *n-tuples, made up of fields. The fields are the* entries of the *n-tuples. For instance, a database of student records may be made up of fields* containing the name, student number, major, and grade point average of the student. The relational data model represents a database of records as an *n-ary relation.* Thus, student records are represented as 4-tuples of the form (*Student_name, ID_number, Major, GPA). A sample* database of two such records is

(Ackermann, 231455, Computer Science, 3.88)
(Adams, 888323, Physics, 3.45).

Student_name	*ID_number*	*Major*	*GPA*
Ackermann	231455	Computer Science	3.88
Adams	888323	Physics	3.45
Chou	102147	Computer Science	3.49
Goodfriend	453876	Mathematics	3.45
Rao	678543	Mathematics	3.90
Stevens	786576	Psychology	2.99

A domain of an *n-ary relation is called a primary key when the value of the n-tuple from* this domain determines the *n-tuple. That is, a domain is a primary key when no two n-tuples in* the relation have the same value from this domain.

Records are often added to or deleted from databases. Because of this, the property that a domain is a primary key is time-dependent. Consequently, a primary key should be chosen that remains one whenever the database is changed. The current collection of *n-tuples in a relation* is called the extension of the relation. The more permanent part of a database, including the name and attributes of the database, is called its intension. When selecting a primary key, the goal should be to select a key that can serve as a primary key for all possible extensions of the database. To do this, it is necessary to examine the intension of the database to understand the set of possible *n-tuples that can occur in an extension.*

Ex. 3: Which domains are primary keys for the *n-ary relation displayed the table below, assuming that no n-tuples will be added in the future?*
Solution: *Because there is only one 4-tuple in this table for each student name, the domain* of student names is a primary key. Similarly, the ID numbers in this table are unique, so the domain of ID numbers is also a primary key. However, the domain of major fields of study is not a primary key, because more than one 4-tuple contains the same major field of study. The domain of grade point averages is also not a primary key, because there are two 4-tuples containing the same GPA.

Student_name	*ID_number*	*Major*	*GPA*
Ackermann	231455	Computer Science	3.88
Adams	888323	Physics	3.45
Chou	102147	Computer Science	3.49
Goodfriend	453876	Mathematics	3.45
Rao	678543	Mathematics	3.90
Stevens	786576	Psychology	2.99

Operations on *n-ary Relations*

There are a variety of operations on *n-ary relations that can be used to form new n-ary relations.* Applied together, these operations can answer queries on databases that ask for all *n-tuples that* satisfy certain conditions.

Definition: Let *R be an n-ary relation and C a condition that elements in R may satisfy. Then the selection operator* s_C *maps the n-ary relation R to the n-ary relation of all n-tuples from R that satisfy* the condition *C.*

Ex. 3: To find the records of computer science majors in the *n-ary relation R shown in Table, we use* the operator s_{C1} *, where* c_1 *is the condition Major="Computer Science." The result is the two* 4-tuples (Ackermann, 231455, Computer Science, 3.88) and (Chou, 102147, Computer Science, 3.49).

Similarly, to find the records of students who have a grade point average above 3.5 in this database, we use the operator s_{C2} *, where* c_2 *is the condition GPA > 3.5. The result is the* two 4-tuples (Ackermann, 231455, Computer Science, 3.88) and (Rao, 678543, Mathematics, 3.90).

Finally, to find the records of computer science majors who have a GPA above 3.5, we use the operator s_{C3} *, where* C_3 *is the condition (Major="Computer Science" ∧ GPA > 3.5).* The result consists of the single 4-tuple (Ackermann, 231455, Computer Science, 3.88).

Operations on *n-ary Relations*

Definition: The *projection* $P_{i_1,i_2,\ldots,i_m}$ *where* $i_1 < i_2 < \cdot \cdot \cdot < i_m$*, maps the n-tuple* $(a_1, a_2, \ldots, a_n)$ *to the m-tuple* $(a_{i_1}, a_{i_2}, \ldots, a_{i_m})$*, where* $m \leq n$.

In other words, the projection $P_{i1,i2,\ldots,im}$ *deletes n − m of the components of an n-tuple, leaving the* i_1*th,* i_2*th, . . . , and* i_m*th components.*

Ex. 4: What results when the projection $P_{1,3}$ *is applied to the 4-tuples (2, 3, 0, 4),* (Jane Doe, 234111001, Geography, 3.14), and (a_1, a_2, a_3, a_4)?
Solution: The projection $P_{1,3}$ sends these 4-tuples to (2, 0), (Jane Doe, Geography), and (a_1, a_3), respectively.

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

Ex. 5: What relation results when the join operator J_2 *is used to combine the relation displayed in the following* tables?

Professor	Department	Course_number
Cruz	Zoology	335
Cruz	Zoology	412
Farber	Psychology	501
Farber	Psychology	617
Grammer	Physics	544
Grammer	Physics	551
Rosen	Computer Science	518
Rosen	Mathematics	575

Department	Course_number	Room	Time
Computer Science	518	N521	2:00 P.M.
Mathematics	575	N502	3:00 P.M.
Mathematics	611	N521	4:00 P.M.
Physics	544	B505	4:00 P.M.
Psychology	501	A100	3:00 P.M.
Psychology	617	A110	11:00 A.M.
Zoology	335	A100	9:00 A.M.
Zoology	412	A100	8:00 A.M.

Solution:

Professor	Department	Course_number	Room	Time
Cruz	Zoology	335	A100	9:00 A.M.
Cruz	Zoology	412	A100	8:00 A.M.
Farber	Psychology	501	A100	3:00 P.M.
Farber	Psychology	617	A110	11:00 A.M.
Grammer	Physics	544	B505	4:00 P.M.
Rosen	Computer Science	518	N521	2:00 P.M.
Rosen	Mathematics	575	N502	3:00 P.M.

Representing Relations

Generally, matrices are appropriate for the representation of relations in computer programs. On the other hand, people often find the representation of relations using directed graphs useful for understanding the properties of these relations.

Representing Relations Using Matrices

A relation between finite sets can be represented using a zero–one matrix. Suppose that *R is a* relation from $A = \{a_1, a_2, \ldots, a_m\}$ to $B = \{b_1, b_2, \ldots, b_n\}$. *(Here the elements of the sets A* and *B have been listed in a particular, but arbitrary, order. Furthermore, when A = B we use* the same ordering for *A and B.) The relation R can be represented by the matrix* $\mathbf{M}_R = [\mathbf{m}_{ij}]$, where

$$m_{ij} = \begin{cases} 1 \text{ if } (a_i, b_j) \in R, \\ 0 \text{ if } (a_i, b_j) \notin R. \end{cases}$$

Representing Relations

Ex. 1: Suppose that A = {1, 2, 3} and B = {1, 2}. Let R be the relation from A to B containing (a, b) if a ∈ A, b ∈ B, and a > b.

What is the matrix representing R if $a_1 = 1$, $a_2 = 2$, and $a_3 = 3$, and $b_1 = 1$ and $b_2 = 2$?

Solution: Because R = {(2, 1), (3, 1), (3, 2)}, the matrix for R is

$$\mathbf{M}_R = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

The 1s in M_R show that the pairs (2, 1), (3, 1), and (3, 2) belong to R. The 0s show that no other pairs belong to R.

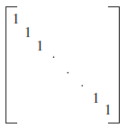
Ex. 2: Let *A = {a_1, a_2, a_3} and B = {b_1, b_2, b_3, b_4, b_5}. Which ordered pairs are in the relation R* represented by the matrix:

$$\mathbf{M}_R = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

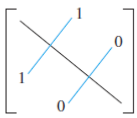
Solution:

$$R = \{(a_1, b_2), (a_2, b_1), (a_2, b_3), (a_2, b_4), (a_3, b_1), (a_3, b_3), (a_3, b_5)\}.$$

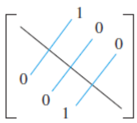
**R is reflexive if and only if (a_i, a_i) ∈ R for i = 1, 2, . . . , n. R is reflexive if all the elements on* the main diagonal of M_R *are equal to 1, as shown in the figure below. Note that the elements off the* main diagonal can be either 0 or 1.



*The relation *R is symmetric if (a, b) ∈ R implies that (b, a) ∈ R. Consequently, the* relation *R on the set A = {a_1, a_2, . . . , a_n} is symmetric if and only if (a_j, a_i) ∈ R whenever (a_i, a_j) ∈ R.*



*The matrix of an antisymmetric relation has the property that **if $m_{ij} = 1$ with i ≠ j , then $m_{ji} = 0$**. *Or, in other words, either $m_{ij} = 0$ or $m_{ji} = 0$ when i ≠ j .*



Ex. 3: Suppose that the relation *R on a set is represented by the matrix.* Is *R reflexive, symmetric, and/or antisymmetric?*

$$\mathbf{M}_R = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Solution: Because all the diagonal elements of this matrix are equal to 1, *R* is reflexive. Moreover, because M_R is symmetric, it follows that *R* is symmetric. It is also easy to see that *R* is not antisymmetric.

Suppose that R_1 and R_2 are relations on a set A represented by the matrices M_{R_1} and M_{R_2}, respectively. The matrix representing the union of these relations has a 1 in the positions where either M_{R_1} or M_{R_2} has a 1. The matrix representing the intersection of these relations has a 1 in the positions where both M_{R_1} and M_{R_2} have a 1. Thus, the matrices representing the union and intersection of these relations are as follows:

$$\mathbf{M}_{R_1 \cup R_2} = \mathbf{M}_{R_1} \vee \mathbf{M}_{R_2} \qquad \mathbf{M}_{R_1 \cap R_2} = \mathbf{M}_{R_1} \wedge \mathbf{M}_{R_2}$$

Ex. 4: Suppose that the relations R_1 and R_2 on a set A are represented by the matrices

$$\mathbf{M}_{R_1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{M}_{R_2} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

What are the matrices representing R1 ∪ R2 and R1 ∩ R2?

Solution:

$$\mathbf{M}_{R_1 \cup R_2} = \mathbf{M}_{R_1} \vee \mathbf{M}_{R_2} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad \mathbf{M}_{R_1 \cap R_2} = \mathbf{M}_{R_1} \wedge \mathbf{M}_{R_2} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

*(Composite of relations) Suppose that *R is a relation from A to B and S is a relation* from *B to C. Suppose also that A, B, and C have m, n, and p elements, respectively. Let the zero–* one matrices for *S∘R, R, and S be $M_{S\circ R}$ = $[t_{ij}]$, M_R = $[r_{ij}]$, and M_S = $[s_{ij}]$, respectively* (these matrices have sizes *m × p, m × n, and n × p, respectively). The ordered pair (a_i, c_j)* belongs to *S∘R if and only if there is an element b_k such that (a_i, b_k) belongs to R and (b_k, c_j)* belongs to *S. It follows that* $t_{ij} = 1$ *if and only if* $r_{ik} = s_{kj} = 1$ *for some k. From the definition* of the Boolean product, this means that

$$\mathbf{M}_{S \circ R} = \mathbf{M}_R \odot \mathbf{M}_S.$$

Ex. 5: Find the matrix representing the relations *S∘R, where the matrices representing R and S are*

$$\mathbf{M}_R = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{M}_S = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Solution:

$$\mathbf{M}_{S \circ R} = \mathbf{M}_R \odot \mathbf{M}_S = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Ex. 6: Find the matrix representing the relation R^2, *where the matrix representing R is*

$$\mathbf{M}_R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Solution:

$$\mathbf{M}_{R^2} = \mathbf{M}_R^{[2]} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{M}_{R^n} = \mathbf{M}_R^{[n]}$$

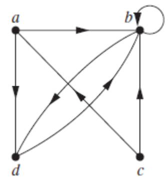
Representing Relations Using Digraphs

There is another important way of representing a relation using a pictorial representation. Each element of the set is represented by a point, and each ordered pair is represented using an arc with its direction indicated by an arrow. We use such pictorial representations when we think of relations on a finite set as directed graphs, or digraphs.

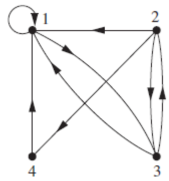
Definition: A *directed graph, or digraph, consists of a set V of vertices (or nodes) together with a set E of ordered pairs of elements of V called edges (or arcs). The vertex a is called the initial vertex of the edge (a, b), and the vertex b is called the terminal vertex of this edge.*

*An edge of the form *(a, a) is represented using an arc from the vertex a back to itself. Such* an edge is called a **loop.**

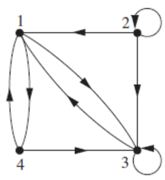
Ex. 6: The directed graph with vertices *a, b, c, and d, and edges (a, b), (a, d), (b, b), (b, d), (c, a), (c, b), and (d, b) is displayed in the Figure below.*



Ex. 7: The directed graph of the relation *R = {(1, 1), (1, 3), (2, 1), (2, 3), (2, 4), (3, 1), (3, 2), (4, 1)}* on the set {1*, 2, 3, 4} is shown in the figure below?.*

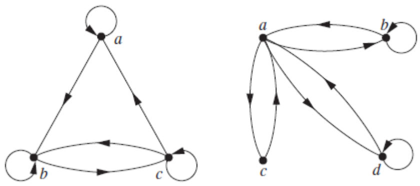


Ex. 8: What are the ordered pairs in the relation *R represented by the directed graph shown in the figure below?*



Solution: *R = {(1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (3, 1), (3, 3), (4, 1), (4, 3)}.*

Ex. 9: Determine whether the relations for the directed graphs shown in the figure below are reflexive, symmetric, antisymmetric, and/or transitive(directed graphs of R and S, respectively.).



Solution:

Because there are loops at every vertex of the directed graph of *R*, it is reflexive. *R* is neither symmetric nor antisymmetric because there is an edge from *a* to *b* but not one from *b* to *a*, but there are edges in both directions connecting *b* and *c*. Finally, *R* is not transitive because there is an edge from *a* to *b* and an edge from *b* to *c*, but no edge from *a* to *c*.

Because loops are not present at all the vertices of the directed graph of *S*, this relation is not reflexive. It is symmetric and not antisymmetric, because every edge between distinct vertices is accompanied by an edge in the opposite direction. It is also not hard to see from the directed graph that *S* is not transitive, because *(c, a)* and *(a, b)* belong to *S*, but *(c, b)* does not belong to *S*.

Closures of Relations

Assume that, a computer network has data centers in **Boston**, **Chicago**, **Denver**, **Detroit**, **New York**, and **San Diego**. There are direct, one-way telephone lines from Boston to Chicago, from Boston to Detroit, from Chicago to Detroit, from Detroit to Denver, and from New York to San Diego. Let *R be the relation containing (a, b) if there is a telephone line from the data center in a to* that in *b.*

How can we determine if there is some (possibly indirect) link composed of one or more telephone lines from one center to another? Because not all links are direct, such as the link from Boston to Denver that goes through Detroit, *R cannot be used directly to answer this.* In the language of relations, *R is not transitive, so it does not contain all the pairs that can be* linked.

As we will show in this lecture, we can find all pairs of data centers that have a link by constructing a transitive relation *S containing R such that S is a subset of every transitive* relation containing *R. Here, S is the smallest transitive relation that contains R. This relation is* called the transitive closure of *R.*

In general, let *R be a relation on a set A. R may or may not have some property P, such as* reflexivity, symmetry, or transitivity. If there is a relation *S with property P containing R such* that *S is a subset of every relation with property P containing R, then S is called the closure* of *R with respect to P.*

Closures of Relations

The relation R = {(1, 1), (1, 2), (2, 1), (3, 2)} on the set A = {1, 2, 3} is not reflexive. How can we produce a reflexive relation containing R that is as small as possible? This can be done by adding (2, 2) and (3, 3) to R, because these are the only pairs of the form (a, a) that are not in R. Clearly, this new relation contains R.

Furthermore, any reflexive relation that contains R must also contain (2, 2) and (3, 3). Because this relation contains R, is reflexive, and is contained within every reflexive relation that contains R, it is called the **reflexive closure of R.** As this example illustrates, given a relation R on a set A, the reflexive closure of R can be formed by adding to R all pairs of the form (a, a) with a ∈ A, not already in R. The addition of these pairs produces a new relation that is reflexive, contains R, and is contained within any reflexive relation containing R. We see that the reflexive closure of R equals R ∪ Δ, where Δ= {(a, a) | a ∈ A} is the **diagonal relation on A.**

Ex. 1: What is the reflexive closure of the relation R = {(a, b) | a < b} on the set of integers?
Solution: The reflexive closure of R is

$$R \cup \Delta = \{(a, b) \mid a < b\} \cup \{(a, a) \mid a \in Z\} = \{(a, b) \mid a \leq b\}.$$

The relation *{(1, 1), (1, 2), (2, 2), (2, 3), (3, 1), (3, 2)} on {1, 2, 3} is not symmetric. How* can we produce a symmetric relation that is as small as possible and contains *R?*

To do this, we need only add *(2, 1) and (1, 3), because these are the only pairs of the form (b, a) with (a, b) ∈ R that are not in R.*

The symmetric closure of a relation can be constructed by taking the union of a relation with its inverse that is, $R \cup R^{-1}$ is the symmetric closure of *R, where* $R^{-1} = \{(b, a) \mid (a, b) \in R\}$.

Ex. 2: What is the symmetric closure of the relation *R = {(a, b) | a < b} on the set of integers?*
Solution: The symmetric closure of *R is the relation*

$$R \cup R^{-1} = \{(a, b) \mid a > b\} \cup \{(b, a) \mid a > b\} = \{(a, b) \mid a \neq b\}.$$

This last equality follows because *R contains all ordered pairs of positive integers where the* first element is greater than the second element and R^{-1} contains all ordered pairs of positive integers where the first element is less than the second.

Suppose that a relation *R is not transitive. How can we produce a transitive relation that* contains *R such that this new relation is contained within any transitive relation that contains R?*
Can the transitive closure of a relation R be produced by adding all the pairs of the form *(a, c), where (a, b) and (b, c) are already in the relation?*

Consider the relation R = {(1, 3), (1, 4), (2, 1), (3, 2)} on the set {1, 2, 3, 4}. This relation is not transitive because it does not contain all pairs of the form *(a, c) where (a, b) and (b, c) are in R.*

The pairs of this form not in *R are (1, 2), (2, 3), (2, 4), and (3, 1). Adding these pairs does not produce a* transitive relation, because the resulting relation contains *(3, 1) and (1, 4) but does not contain (3, 4).*

This shows that constructing the transitive closure of a relation is more complicated than constructing either the reflexive or symmetric closure.

Paths in Directed Graphs

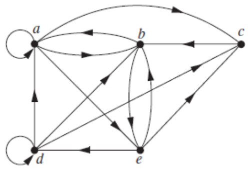
A path in a directed graph is obtained by traversing along edges and this helps in the construction of transitive closures

Definition: A *path from a to b in the directed graph G is a sequence of edges* $(x_0, x_1), (x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)$ *in G, where n is a nonnegative integer, and* $x_0 = a$ *and* $x_n = b$, that is, a sequence of edges where the terminal vertex of an edge is the same as the initial vertex in the next edge in the path. This path is denoted by $x_0, x_1, x_2, \ldots, x_{n-1}, x_n$ *and has length n. We view the empty set of edges as a path of length zero from a to a. A path of* length $n \geq 1$ *that begins and ends at the same vertex is called a circuit or cycle.*

A path in a directed graph can pass through a vertex more than once. Moreover, an edge in a directed graph can occur more than once in a path.

Ex. 3: Which of the following are paths in the directed graph shown in the figure below: *a, b, e, d; a, e, c, d, b; b, a, c, b, a, a, b; d, c; c, b, a; e, b, a, b, a, b, e? What are the lengths of those that are paths?* Which of the paths in this list are circuits?

Solution:



The term *path also applies to relations. Carrying over the definition from directed graphs to* relations, there is a path from *a to b in R if there is a sequence of elements* $a, x_1, x_2, \ldots, x_{n-1}, b$ with $(a, x_1) \in R$, $(x_1, x_2) \in R, \ldots,$ *and* $(x_{n-1}, b) \in R$. *Following theorem can be obtained from the* definition of a path in a relation.

Theorem: Let *R be a relation on a set A. There is a path of length n, where n is a positive integer, from a to b if and only if* $(a, b) \in R^n$.

Transitive Closures

Finding the transitive closure of a relation is equivalent to determining which pairs of vertices in the associated directed graph are connected by a path.

Definition: Let *R be a relation on a set A. The connectivity relation* R^* consists of the pairs *(a, b) such* that there is a path of length at least one from *a to b in R.*

Because R^n *consists of the pairs (a, b) such that there is a path of length n from a to b, it follows* that R^* is the union of all the sets R^n.

$$R^* = \bigcup_{n=1}^{\infty} R^n.$$

Ex. 4: Let R be the relation on the set of all subway stops in New York City that contains (a, b) if it is possible to travel from stop 'a' to stop 'b' without changing trains. What is R^n when n is a positive integer? What is R^*?

Solution: The relation R^n contains (a, b) if it is possible to travel from stop a to stop b by making at most n − 1 changes of trains. The relation R^* consists of the ordered pairs (a, b) where it is possible to travel from stop a to stop b making as many changes of trains as necessary.

Ex. 5: Let *R be the relation on the set of all states in the United States that contains (a, b) if state a* and state *b have a common border. What is R^n, where n is a positive integer? What is R^**?

Solution:
The relation R^n consists of the pairs *(a, b)*, where it is possible to go from state *a* to state *b* by crossing exactly *n* state borders. R^* consists of the ordered pairs *(a, b)*, where it is possible to go from state *a* to state *b* crossing as many borders as necessary.

Theorem: The transitive closure of a relation *R equals the connectivity relation R^**.

Now, we know that the transitive closure equals the connectivity relation, we turn our attention to the problem of computing this relation. We do not need to examine arbitrarily long paths to determine whether there is a path between two vertices in a finite directed graph. As the following Lemma shows, it is sufficient to examine paths containing no more than *n edges, where n is* the number of elements in the set.

Lemma: Let *A be a set with n elements, and let R be a relation on A. If there is a path of length at* least one in *R from a to b, then there is such a path with length not exceeding n. Moreover,* when *a ≠ b, if there is a path of length at least one in R from a to b, then there is such a path* with length not exceeding *n − 1.*

From Lemma, we see that the transitive closure of R is the union of R, R^2, R^3, . . . ,and R^n. This follows because there is a path in R^* between two vertices if and only if there is a path between these vertices in R_i , for some positive integer i with i ≤ n. Because

$$R^* = R \cup R^2 \cup R^3 \cup \cdots \cup R^n$$

and the zero–one matrix representing a union of relations is the join of the zero–one matrices of these relations, the zero–one matrix for the transitive closure is the join of the zero–one matrices of the first *n powers of the zero–one matrix of R.*

Theorem: Let M_R *be the zero–one matrix of the relation R on a set with n elements. Then the zero–one* matrix of the transitive closure R^* is $\mathbf{M}_{R^*} = \mathbf{M}_R \vee \mathbf{M}_R^{[2]} \vee \mathbf{M}_R^{[3]} \vee \cdot\cdot\cdot \vee \mathbf{M}_R^{[n]}$.

Ex. 6: Find the zero–one matrix of the transitive closure of the relation *R where*

$$\mathbf{M}_R = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

Solution:

$$\mathbf{M}_{R^*} = \mathbf{M}_R \vee \mathbf{M}_R^{[2]} \vee \mathbf{M}_R^{[3]}.$$

$$\mathbf{M}_R^{[2]} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{M}_R^{[3]} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$

$$\mathbf{M}_{R^*} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \vee \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \vee \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Warshall's Algorithm

Warshall's algorithm is based on the construction of a sequence of zero–one matrices.

These matrices are $W_0, W_1, \ldots, W_n$, *where* $W_0 = M_R$ *is the zero–one matrix of this relation, and* $W_k = [w^{(k)}_{ij}]$, where $w^{(k)}_{ij} = 1$ if there is a path from v_i *to* v_j *such that all the interior vertices of* this path are in the set $\{v1, v2, \ldots, vk\}$ *(the first k vertices in the list) and is 0 otherwise.* (The first and last vertices in the path may be outside the set of the first *k vertices in the list.)*

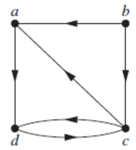
Note that $\mathbf{W}_n = \mathbf{M}_{R^*}$, because the *(i, j)th entry of* $\mathbf{M}_{R^*}$ is 1 if and only if there is a path from v_i to v_j, with all interior vertices in the set $\{v_1, v_2, \ldots, v_n\}$ *(but these are the only vertices in the* directed graph).

Suppose that *R is a relation on a set with n elements. Let* $v_1, v_2, \ldots, v_n$ *be an arbitrary* listing of these *n elements. The concept of the interior vertices of a path is used in Warshall's* algorithm.

If $a, x_1, x_2, \ldots, x_{m-1}, b$ *is a path, its interior vertices are* $x_1, x_2, \ldots, x_{m-1}$, *that* is, all the vertices of the path that occur somewhere other than as the first and last vertices in the path. For instance, the interior vertices of a path *a, c, d, f, g, h, b, j in a directed graph* are *c, d, f, g, h, and b. The interior vertices of a, c, d, a, f, b are c, d, a, and f* .

Note that the first vertex in the path is not an interior vertex unless it is visited again by the path, except as the last vertex. Similarly, the last vertex in the path is not an interior vertex unless it was visited previously by the path, except as the first vertex.

Ex. 7: Let R be the relation with directed graph shown in figure below. Let a, b, c, d be a listing of the elements of the set. Find the matrices W_0, W_1, W_2, W_3, and W_4. The matrix W_4 is the transitive closure of R.



Solution: Let v_1 = a, v_2 = b, v_3 = c, and v_4 = d. $\mathbf{W}_0$ is the matrix of the relation. Hence,

$$\mathbf{W}_0 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

$\mathbf{W}_1$ has 1 as its (i, j)th entry if there is a path from v_i to v_j that has only "v_1 = a" as an interior vertex. Note that all paths of length one can still be used because they have no interior vertices. Also, there is now an allowable path from b to d, namely, b, a, d. Hence,

$$\mathbf{W}_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$\mathbf{W}_2$ has 1 as its (i, j)th entry if there is a path from v_i to v_j that has only v_1 = a and/or v_2 = b as its interior vertices, if any. Because there are no edges that have b as a terminal vertex, no new paths are obtained when we permit b to be an interior vertex. Hence, $\mathbf{W}_2 = \mathbf{W}_1$.

$\mathbf{W}_3$ has 1 as its *(i, j)*th entry if there is a path from v_i to v_j that has only $v_1 = a$, $v_2 = b$, and/or $v_3 = c$ as its interior vertices, if any. We now have paths from *d* to *a*, namely, *d, c, a*, and from *d* to *d*, namely, *d, c, d*. Hence,

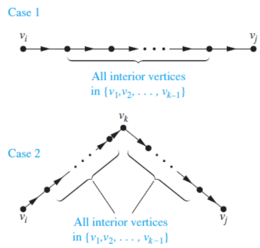
$$\mathbf{W}_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}.$$

Finally, $\mathbf{W}_4$ has 1 as its *(i, j)*th entry if there is a path from v_i to v_j that has $v_1 = a$, $v_2 = b$, $v_3 = c$, and/or $v_4 = d$ as interior vertices, if any. Because these are all the vertices of the graph, this entry is 1 if and only if there is a path from v_i to v_j . Hence,

$$\mathbf{W}_4 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}.$$

This last matrix, $\mathbf{W}_4$, is the matrix of the transitive closure.

Warshall's algorithm computes $\mathbf{M}_{R*}$ by efficiently computing $W_0 = M_R$, W_1, W_2, . . . , $W_n = M_{R*}$. This observation shows that we can compute W_k *directly from* W_{k-1}*: There is a* path from *vi to vj with no vertices other than* $v_1, v_2, \ldots, v_k$ *as interior vertices if and only if* either there is a path from v_i *to vj with its interior vertices among the first* $k-1$ *vertices in the* list, or there are paths from v_i *to* v_k *and from* v_k *to vj that have interior vertices only among the* first $k-1$ *vertices in the list. That is, either a path from* v_i *to* v_j *already existed before* v_k *was* permitted as an interior vertex, or allowing v_k *as an interior vertex produces a path that goes* from v_i *to* v_k *and then from* v_k *to* v_j *. These two cases are shown in the following Figure.*



The first type of path exists if and only if $w_{ij}^{[k-1]} = 1$, and the second type of path exists if and only if both $w_{ik}^{[k-1]}$ *and* $w_{kj}^{[k-1]}$ *are 1. Hence,* $w_{ij}^{[k]}$ *is 1 if and only if either* $w_{ij}^{[k-1]}$ is 1 *or both* $w_{ik}^{[k-1]}$ *and* $w_{kj}^{[k-1]}$ *are 1.*

This gives us following Lemma .

Lemma: Let $W_k = [w_{ij}^{[k]}]$ be the zero–one matrix that has a 1 in its *(i, j)th position if and only if there* is a path from *vi to vj with interior vertices from the set {v1, v2, . . . , vk}. Then* $w_{ij}^{[k]} = w_{ij}^{[k-1]} \vee (w_{ik}^{[k-1]} \wedge w_{kj}^{[k-1]})$*, whenever i, j , and k are positive integers not exceeding n.*