DISCRETE MATHEMATICS
（离散数学）

云南大学数学系
李 建 平

Chapter 4 RELATIONS AND DIGRAPHS

4.1 PRODUCT SETS AND PARTITIONS

Product Sets

An **ordered pair** (a,b) is listing of the objects a and b in a prescribed order, with a appearing first and b appearing second. Ordered pairs (a_1,b_1) and (a_2,b_2) are equal if and only if $a_1=a_2$ and $b_1=b_2$.

If A and B are two nonempty sets, we define the **product set** or **Cartesian product** $A\times B$ as the set of all ordered pairs (a,b) with $a\in A$ and $b\in B$. Thus

$$A\times B=\{(a,b)|\ a\in A \text{ and } b\in B\}.$$

Theorem 1: For any two finite, nonempty sets A and B, $|A \times B|=|A||B|$.

Cartesian product $A_1 \times A_2 \times \ldots \times A_m$ of the nonempty sets $A_1, A_2, \ldots, A_m$ is the set of all ordered m-tuples $(a_1, a_2, \ldots, a_m)$, where $a_i \in A_i$, $i=1,2,\ldots,m$. Thus

$$A_1 \times A_2 \times \ldots \times A_m = \{(a_1, a_2, \ldots, a_m) | \ a_i \in A_i, \ i=1,2,\ldots,m\}.$$

- Partitions（划分）

A **partition** or **quotient set** of a nonempty set A is a collection P of nonempty subsets of A such that

1. Each element of A belongs to one of the sets in

P.

2. If A_1 and A_2 are distinct elements of P, then $A_1 \cap A_2 = \varnothing$.

The sets in P are called the blocks or cells of the partition. Figure 4.2 shows a partition $P=\{A_1,A_2,A_3,A_4,A_5,A_6,A_7\}$ into seven blocks.

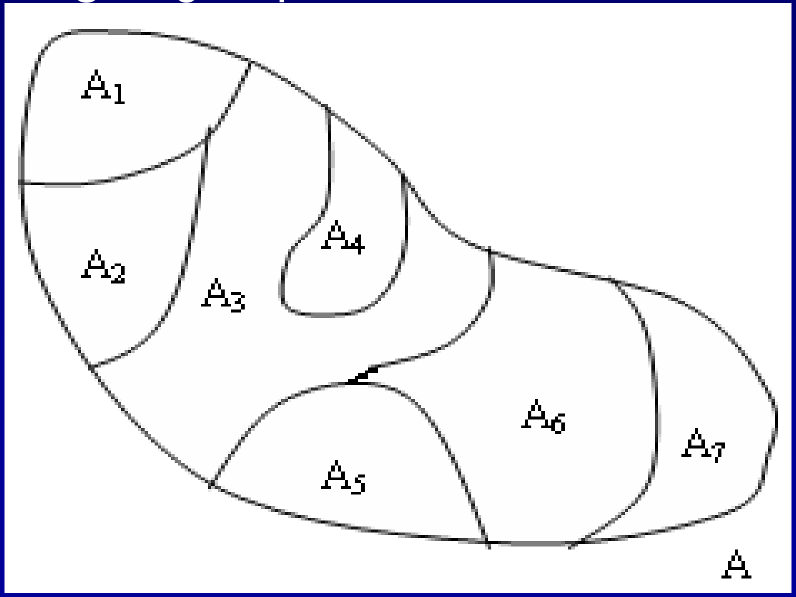


Figure 4.2

4.2 RELATIONS AND DIGRAPHS

If A is the set of all living humans males and B is the set of all living human females, then the relation F (father) can be defined between A and B. Thus, if $x \in A$ and $y \in B$, then x is related to y by the relation F if x is the father of y, and we write xFy.

Let A and B be nonempty sets. A **relation** R from A to B is a subset of $A \times B$. If $R \subseteq A \times B$ and $(a,b) \in R$, we say that **a is related to b by R**, and we also write aRb. If a is not related to b by R, and we write $a \not{R} b$. Frequently, A and B are equal. In this case, we often say that $R \subseteq A \times A$ is a relation on A, instead of a relation from A to A.

Sets Arising from Relations

Let $R \subseteq A \times B$ be a relation from A to B.

The **domain** of R, denoted by Dom (R), is the set of elements in A that are related to some element in B. We define the **range** of R, denoted by Ran (R), to be the set of elements in B that are second elements of pairs in R.

If R is a relation from A to B and $x \in A$, we define R(x), the **R-relative set of x**, to be the set of all y in B with the property that x is R-related to y, i.e.,

$$R(x) = \{ y \in B \mid x \, R \, y \}.$$

If $A_1 \subseteq A$, then $R(A_1)$, the **R-relative set of A_1**, is the set of all y in B with the property that x is R-related to y for some x in A_1, i.e.,

$$R(A_1)=\{ y \in B \mid x \; R \; y \text{ for some } x \text{ in } A_1\}.$$

See that $R(A_1)$ is the union of the sets $R(x)$, where $x \in A_1$.

Theorem 1: Let R be a relation from A to B, and let A_1 and A_2 be subsets of A. Then

(a) If $A_1 \subseteq A_2$, then $R(A_1) \subseteq R(A_2)$.

(b) $R(A_1 \cup A_2)=R(A_1) \cup R(A_2)$.

(c) $R(A_1 \cap A_2) \subseteq R(A_1) \cap R(A_2)$.

Theorem 2: Let R and S be relations from A to B. If $R(a)=S(a)$ for a in A, then $R=S$.

- The Matrix of a Relation

If $A=\{a_1,a_2,\ldots,a_m\}$ and $B=\{b_1,b_2,\ldots,b_n\}$ are finite sets containing *m* and *n* elements, respectively, and R is a relation from A to B, we represent R by the $m \times n$ matrix $M_R=[m_{ij}]$, which is defined by

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

The matrix M_R is called the matrix of R.

- The Digraph of a Relation

If A is a finite set and R is a relation on A, we can also represent R pictorially as follows. Draw a small circle for each element of A and label the circle with the corresponding element of A. These circles are called **vertices**. Draw an arrow, called an **arc,** from vertex a_i to vertex a_j if and only if a_i R a_j. The resulting pictorial representation of R is called a directed graph or **digraph** of R.

If R is a relation on A, the arcs in the digraph of R correspond exactly to the pairs in R, and the vertices correspond exactly to the elements of the set A.

If R is a relation on a set A and $x \in A$, then the **in-degree** of x (relative to the relation R) is the number of $y \in A$ such that $(y,x) \in R$.

The **out-degree** of x is the number of $y \in A$ such that $(x,y) \in R$.

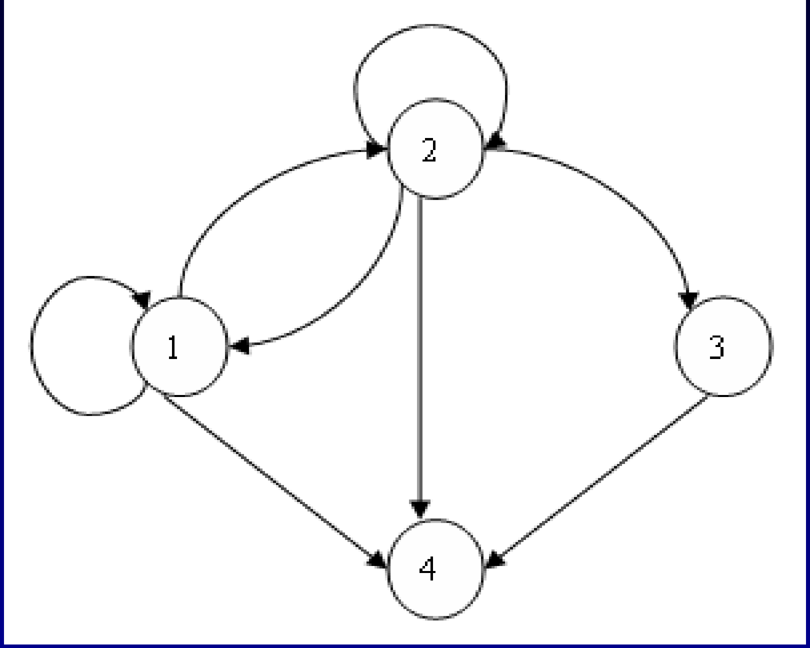


Figure 4.4

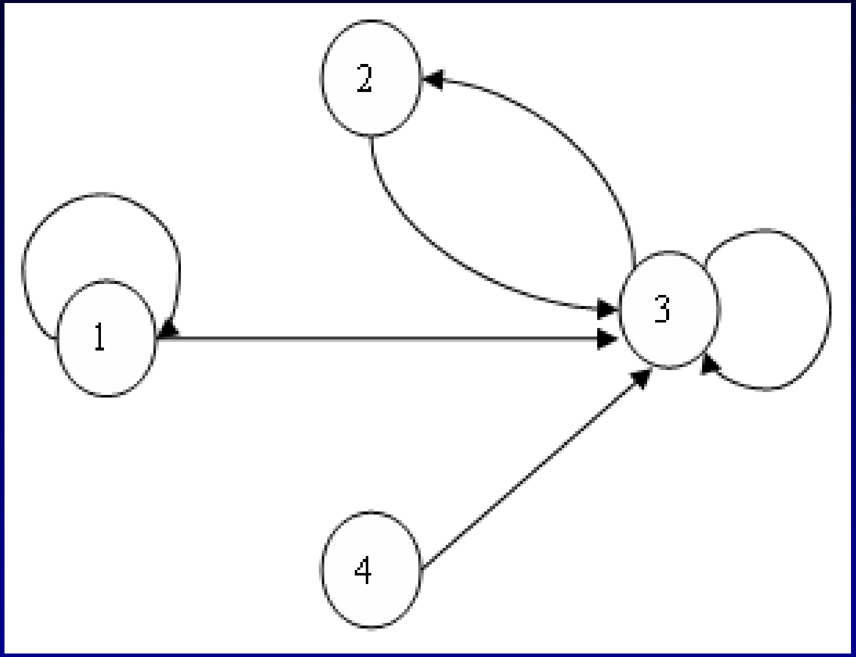


Figure 4.5

4.3 PATHS IN RELATIONS AND DIGRAPHS

Suppose that R is a relation on a set A. A **path** of length n in R from x_0 to x_n is a finite sequence Π: $x_0, x_1, x_2, \ldots, x_{n-1}, x_n$, beginning with the vertex x_0 and ending with the vertex x_n, such that

$$x_0 \; R \; x_1,\; x_1 \; R \; x_2,\; \ldots,\; x_{n-1} \; R \; x_n.$$

This path involves n+1 elements (not necessarily distinct) of A, and n is called the **length** of path Π.

A path that begins and ends at the same vertex is called a **cycle**. Paths in a relation R can be used to defined new relations that are quite useful. If n is a fixed positive integer, we define a new relation

R^n on A as follows: x R^n y means that there is a path of length n from x to y in R. We may also define a relation R^∞ on A, by letting x R^∞ y mean that there is some path from x to y. The relation R^∞ is sometimes called the **connectivity relation** for R.

Note that $R^n(x)$ consists of all vertices that can be reached from x by means of a path in R of length n. The set $R^\infty(x)$ consists of all vertices all vertices that can be reached from x by some path in R.

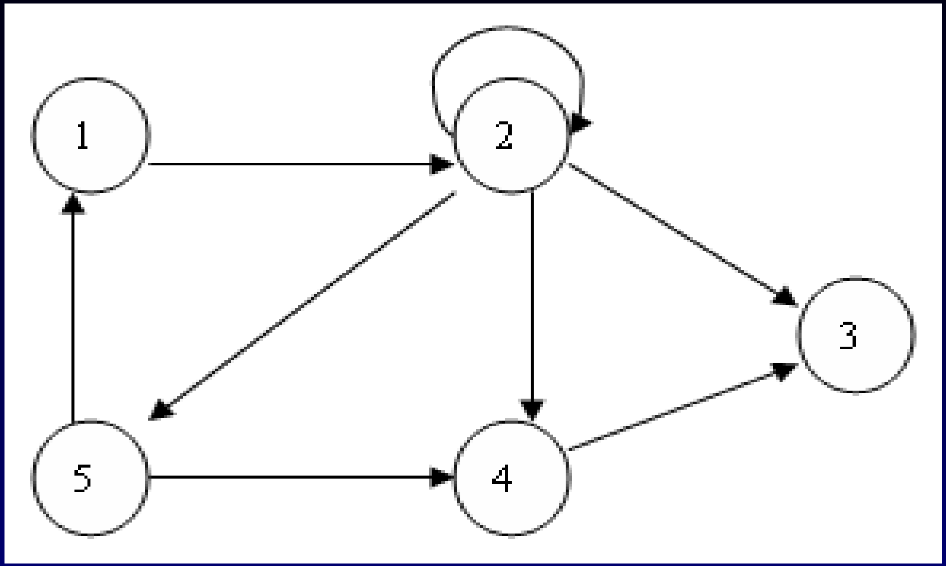


Figure 4.11

Let R be a relation on a finite set $A=\{a_1,a_2,\ldots,a_n\}$, and let M_R be the $n\times n$ matrix representing R. We will show how the matrix $M_R{}^2$, of R^2, can be computed form M_R.

Theorem 1: If R is a relation on $A=\{a_1,a_2,\ldots,a_n\}$, then $M_R^2=M_R\odot M_R$ (see Section1.5).

Proof: Let $M_R=[m_{ij}]$ and $M_R^2=[n_{ij}]$. By definition, the i, jth element of $M_R\odot M_R$ is equal to 1 if and only if row i of M_R and column j of M_R have a 1 in the same relative position, say position k. This means that $m_{ik}=1$ and $m_{kj}=1$ for some k, $1\leqslant k\leqslant n$. By definition of the matrix M_R, the preceding conditions mean that $a_i\ R\ a_k$ and $a_k\ R\ a_j$. Thus $a_i\ R^2\ a_j$, and so $n_{ij}=1$. We have therefore shown that position i, j of $M_R\odot M_R$ is equal to 1 if and only if $n_{ij}=1$. This means that $M_R\odot M_R=M_R^2$.

Theorem 2: For $n \geqslant 2$ and R a relation on a finite set A, we have $M_R^n = M_R \odot M_R \odot \cdots \odot M_R$ (n factors).

Proof: Let P(n) be the assertion that the statement holds for an integer $n \geqslant 2$.

Basis Step: P(2) is true by Theorem 1.

Induction Step: We use P(k) to show P(k+1). Consider the matrix M_R^{k+1}. Let $M_R^{k+1}=[x_{ij}]$, $M_R^k=[y_{ij}]$, and $M_R=[m_{ij}]$. If $x_{ij}=1$, we must have a path of length k+1 from a_i to a_j. If we let a_s be the vertex that this path reaches just length 1 from a_s to the last vertex a_j. Thus $y_{is}=1$ and $m_{sj}=1$, so $M_R^k \odot M_R$ has a 1 in position i, j, then $x_{ij}=1$. This means that $M_R^{k+1} = M_R^k \odot M_R$.

Using

$$P(k): M_R^k = M_R \odot \cdots \odot M_R \text{ (k factors)},$$

we have

$$M_R^{k+1} = M_R^k \odot M_R = (M_R \odot M_R \odot \cdots \odot M_R) \odot M_R$$

and hence

$$P(k+1): M_R^{k+1} = M_R \odot \cdots \odot M_R \odot M_R \text{ (k+1 factors)}$$

is true. Thus, by the principle of mathematical induction, $P(n)$ is true for all $n \geqslant 2$. This proves the theorem. As before, we write $M_R^k = M_R \odot \cdots \odot M_R$ (k factors) as $(M_R)^n{}_\odot$.

Suppose that R is a relation on a finite set A, and $x \in A$, $y \in A$. $x\ R^\infty\ y$ means that x and y are connected by a path in R of length n for some n.

x R^∞ y if and only if x R y or x R^2 y or x R^3 y or … .

$$R^\infty = R \cup R^2 \cup R^3 \cup \cdots = \bigcup_{n=1}^{\infty} R^n.$$

If R and S are relations on A, the relation $R \cup S$ is defined by x $(R \cup S)$ y if and only if x R y or x S y.

We can verify that $M_{R \cup S} = M_R \vee M_S$. Then

$$M_{R^\infty} = M_R \vee M_{R^2} \vee M_{R^3} \vee \cdots = M_R \vee (M_R)^2_\odot \vee (M_R)^3_\odot \vee \cdots.$$

The **reachability relation** R* of a relation R on a set A that has n elements is defined as follows:

x R* y means that x=y or x R^∞ y.

$M_{R^*} = M_{R^\infty} \vee I_n$, where I_n is the $n \times n$ identity matrix.

$$M_{R^*} = I_n \vee M_R \vee (M_R)^2_\odot (M_R)^3_\odot \vee \cdots .$$

Let Π_1: a, x_1, x_2, …, x_{n-1}, b be a path in a relation R of length n from a to b, and let Π_2: b, y_1, y_2,…,y_{m-1},c be a path in R of length m from b to c. Then the **composition（复合）of** Π_1 **and** Π_2 is the path a, x_1, x_2, …, x_{n-1}, b , y_1, y_2,…,y_{m-1},c of the length n+m, which is denoted by $\Pi_2 \circ \Pi_1$. This is a path from a to c.

4.4 PROPERTIES OF RELATIONS

Reflexive（自反）and Irreflexive（反自反）Relations

A relation R on a set A is **reflexive** if $(a,a) \in R$ for all $a \in A$, that is, if a R a for all $a \in A$. A relation R on a set A is **irrefexive** if $a \not R a$ for every $a \in A$.

Example (a) Let $\triangle=\{(a,a)|a \in A\}$, so that $\triangle$ is the relation of equality on the set A. Then $\triangle$ is reflexive, since $(a,a) \in \triangle$ for all $a \in A$.

(b) Let A be a nonempty set. Let $R=\varnothing \subseteq A \times A$, the **empty relation**. Then R is not reflexive, since $(a,a) \notin R$ for all $a \in A$ (the empty set has no elements). However, R is irreflexive.

The matrix of a reflexive relation must have all 1's on its main diagonal, while the matrix of an irreflexive relation must have all 0's on its main diagonal.

A reflexive relation has a cycle of length 1 at every vertex, while an irreflexive relation has no cycles of length 1. R is reflexive if and only if $\triangle \subseteq R$, and R is irreflexive if and only if $\triangle \cap R = \varnothing$.

Symmetric, Asymmetric and Antisymmetric Relations

A relation R on a set A is **symmetric**（对称）if whenever a R b, then b R a. A relation R on a set A is **asymmetric**（斜对称）if whenever a R b, then b $\not R$ a. A relation R on a set A is **antisymmetric**（反对称）if whenever a R b and b R a, then a=b.

The matrix $M_R=[m_{ij}]$ of a symmetric relation satisfies the property that

if $m_{ij}=1$, then $m_{ji}=1$.

Moreover, if $m_{ji}=0$, then $m_{ij}=0$, i.e., $M_R= M_R^T$, so that M_R is a symmetric matrix.

The matrix $M_R=[m_{ij}]$ of an asymmetric relation R satisfies the property that

if $m_{ij}=1$, then $m_{ji}=0$.

If R is asymmetric, it follows that $m_{ii}=0$ for all i.

The matrix $M_R=[m_{ij}]$ of an antisymmetric relation R satisfies the property that if $i\neq j$, then $m_{ij}=0$ or $m_{ji}=0$.

We now consider the digraphs of these three types of relations. If R is an asymmetric relation, then the digraph of R cannot simultaneously have an arc from vertex i to vertex j and an arc from vertex j to vertex i. Thus there can be no cycles of length 1, and all arcs are “one-way streets.”

If R is an antisymmetric relation, then for different vertices i and j there cannot be an arc from vertex i to vertex j and an arc from vertex j to vertex i. Thus there may be cycles of length 1, but again all arcs are “one way.”

The digraph of a symmetric relation R has the property that if there is an arc from vertex i to vertex j, then there is an arc from vertex j to vertex i. If two vertices a and b are connected by arcs in each direction, we replace these two arcs with one undirected edge, or a “two-way street.” The resulting diagram will be called the **graph** of the symmetric relation.

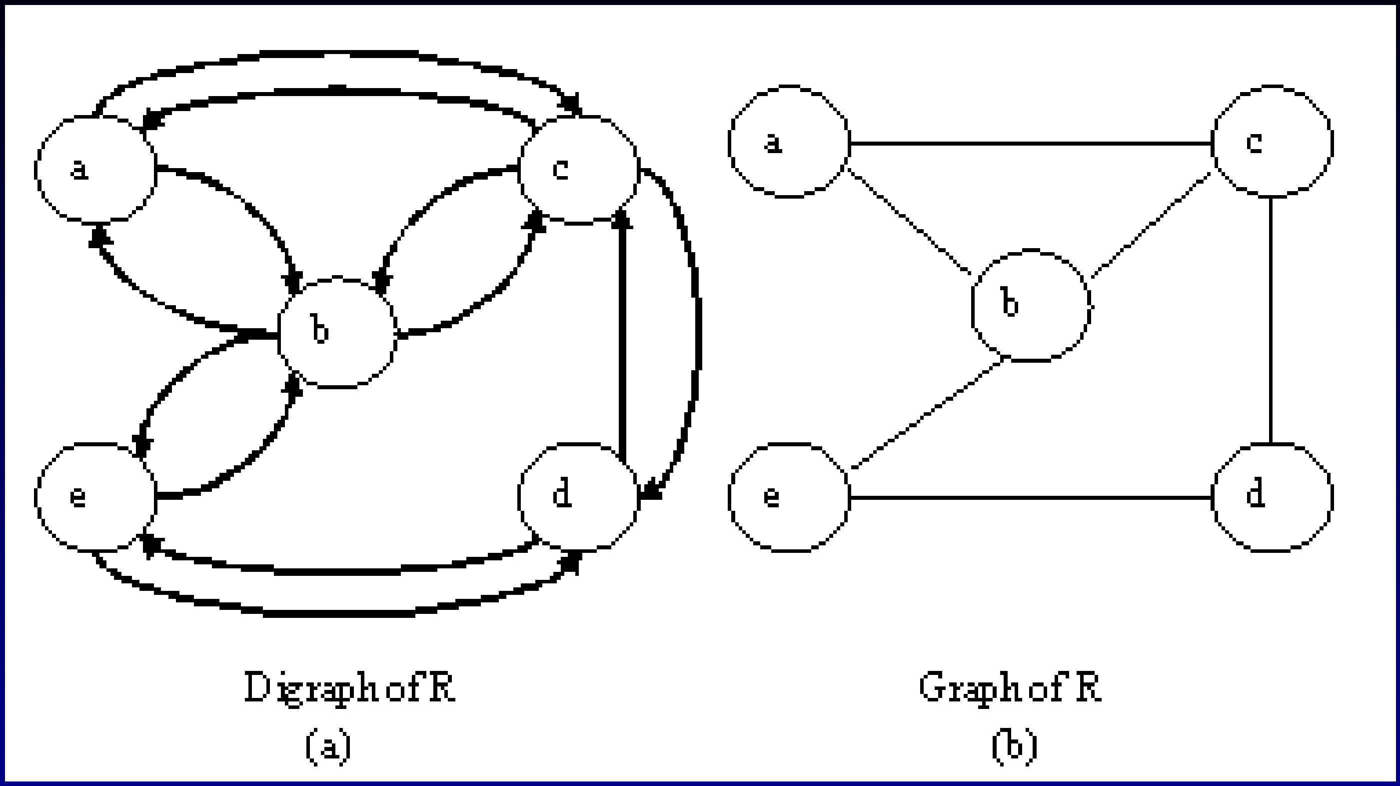


Figure 4.20

A symmetric relation R on a set A is called connected if there is a path from any element of A to any other element of A. In Figure 4.21 we show the graphs of two symmetric relations. The graph in Figure 4.21(a) is connected, whereas that in Figure 4.21(b) is not connected.

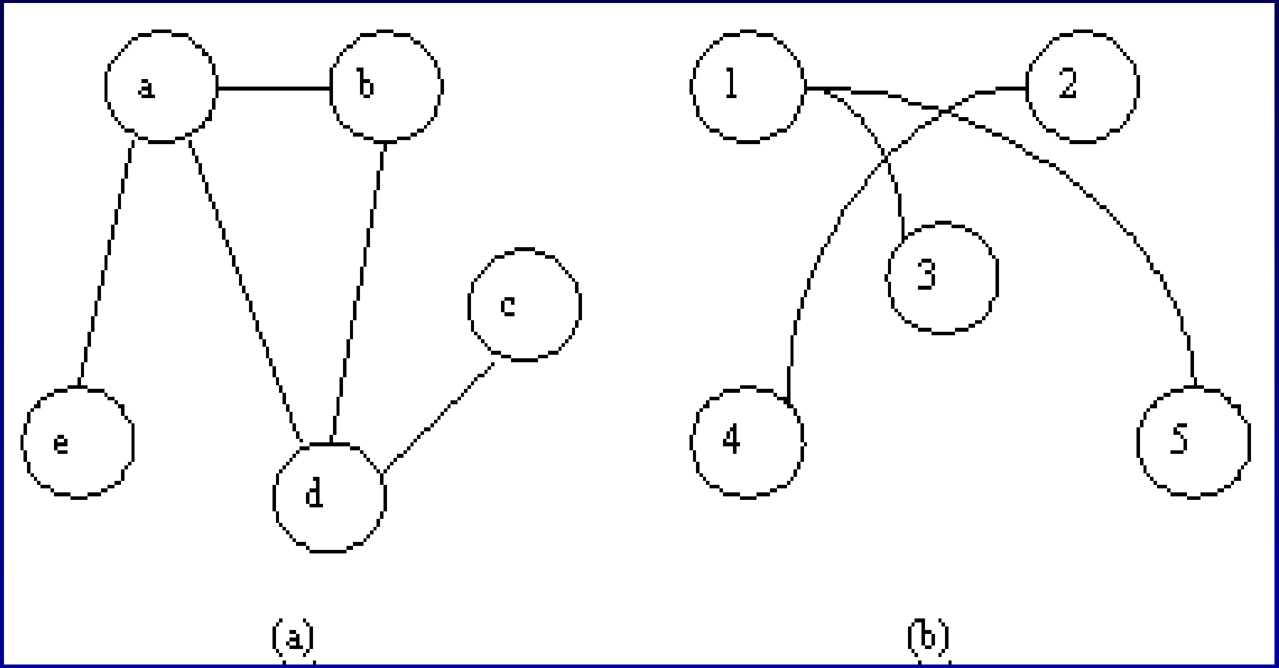


Figure 4.21

Transitive Relations（传递关系）

We say that a relation R on a set A is **transitive** if whenever a R b and b R c, then a R c. A relation R on A is not transitive if there exist a, b, and c in A so that a R b and b R c, but a $\not R$ c. If such a, b, and c do not exist, then R is transitive.

Example　Let $A=Z^+$ and let R be the division relation considered in Example 5. Is R transitive?

Solution: Suppose that a R b and b R c, so that a | b and b | c. It then does follow that a | c. [see Theorem 2(d) of Section 1.4.]. Thus R is transitive.

Example Let A={1,2,3,4} and let

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

Is R transitive?

Solution: Since there are no elements a, b, and c in such that a R b and b R c, but a $\not R$ c, we conclude that R is transitive.

A relation R is transitive if and only if its matrix $M_R=[m_{ij}]$] has the property

if $m_{ij}=1$ and $m_{jk}=1$, then $m_{ik}=1$.

The left-hand side of this statement simply means that $(M_R)^2_\odot$ has a 1 in position i, k. Thus the transitivity of R means that if $(M_R)^2_\odot$ has a 1 in any position, then M_R must have a 1 in the same position. Thus, in particular, if $(M_R)^2_{\odot} = M_R$, then R is transitive. The converse is not true.

If we consider particular vertices a and c, the conditions a R b and b R c mean that there is a path of length 2 in R from a to c. The definition of transitivity as follows: If a R^2 c, then a R c; that is, $R^2 \subseteq R$. In other words, if a and c are connected by a path of length 2 in R, then they must be connected by a path of length 1.

Theorem 1: A relation R is transitive if and only if it satisfies the following property: If there is a path of length greater than 1 from vertex x to vertex y, there is a path of length 1 from x to y (that is, x is related to y). Algebraically stated, R is transitive if and only if $R^n \subseteq R$ for all $n \geqslant 1$.

Theorem 2: Let R be a relation on a set A. Then

(a) Reflexivity of R means that $a \in R(a)$ for all a in A.

(b) Symmetry of R means that $a \in R(b)$ if and only if $b \in R(a)$.

(c) Transitivity of R means that if $b \in R(a)$ and $c \in R(b)$, then $c \in R(a)$.

4.5 EQUIVALENCE RELATIONS（等价关系）

A relation R on a set A is called an **equivalence relation** if it is reflexive, symmetric, and transitive.

Example Let A=Z, the set of integers, and let R be defined by a R b if and only if $a \leqslant b$. Is R an equivalence relation?

Solution: R is not an equivalence relation.

Example Let A=Z and $n \in Z^+$. Let

$$R=\{(a,b) \in A \times A \mid a \equiv b \pmod{n}\}.$$

i.e., $a \equiv b \pmod{n}$ if and only if a and b yield the same remainder when divided by n. We can show that congruence mod n is an equivalence relation.

We note that if $a \equiv b \pmod{n}$, then $a=qn+r$ and $b=tn+r$ and $a-b$ is a multiple of n. Thus, $a \equiv b \pmod{n}$ if $n|(a-b)$.

Equivalence Relations and Partitions

Theorem 1: Let P be a partition of a set A. Recall that the sets in P are called the blocks of P.

Define the relation R on A as follows:

a R b if and only if a and b are members of the same block.

Then R is an equivalence relation on A.

Proof:

(1) If $a \in A$, then clearly a is in the same block as itself;

(2) If a R b, then a and b are in the same block; so

b R a.

(3) If a R b and b R c, then a, b, and c must all lie in the same block of P. Thus a R c.

This relation R is called the **equivalence relation determined by P**.

All equivalence relations on A can be produced from partitions.

Lemma 1: Let R be an equivalence relation on a set A, and let $a \in A$ and $b \in A$. Then

a R b if and only if R(a)=R(b).

Proof. Suppose that R(a)=R(b). Since R is reflexive, then $b \in R(b)$; therefore, $b \in R(a)$, and then a R b.

Conversely, suppose that a R b. We must show that R(a)=R(b). First, choose an element $x \in R(b)$, then b R x; since R is also transitive, we have aRx, and then $x \in R(a)$. Thus $R(b) \subseteq R(a)$.

Similarly, $R(a) \subseteq R(b)$. Thus R(a) = R(b).

Theorem 2. Let R be an equivalence relation on A, and let P be the collection of all distinct relative sets R(a) for a in A. Then P is a partition of A, and R is the equivalence relation determined by P .

Proof. According to the definition of a partition, we must show the following two properties:

(a) Every element of A belongs to some relative set.

(b) If R(a) and R(b) are not identical, then

$R(a) \cap R(b) = \varnothing$.

Now the property (a) is true, since $a \in R(a)$ by reflexivity of R. To show property (b) we prove the following equivalent statement:

If $R(a) \cap R(b) \neq \varnothing$, then $R(a)=R(b)$.

To prove this, we assume that $c \in R(a) \cap R(b)$. Then a R c and b R c.

Since R is symmetric, we have c R b. Then by transitivity of R, we have a R b. Lemma 1 then tells us that $R(a)=R(b)$.

We have now proved that P is a partition. By Lemma 1 we see that a R b if and only if a and b belong to the same block of P . Thus P determines R, and the theorem is proved.

If R is an equivalence relation on A, then the sets R(a) are traditionally called **equivalence classes**（等价类）of R. Denote the class R(a) by [a] or $[a]_R$. The partition P consists of all equivalence classes of R, and this partition will be denoted by A/R.

Partitions of A are also called quotient sets（商集合）of A.

A general procedure for determining partitions A/R for A finite or countable.

The procedure is as follows:

Step 1: Choose any element of A and compute the equivalence class R(a).

Step 2: If R(a)$\neq$A, choose an element b, not included in R(a), and compute the equivalence class R(b).

Step 3: If A is not the union of previously computed equivalence classes, then choose an element x of A that is not in any of those equivalence classes and compute R(x).

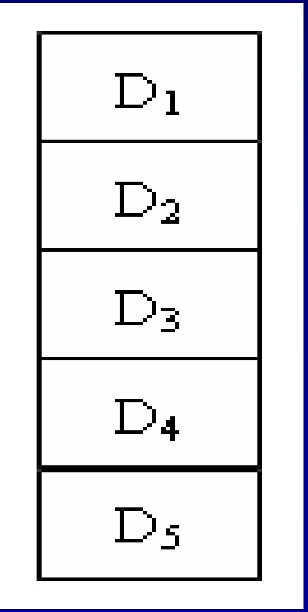
Step 4: Repeat step 3 until all elements of A are included in the computed equivalence classes. If A is countable, this process could continue indefinitely. In that case, continue until a pattern emerges that allows you to describe or give a formula for all equivalence classes.

4.6 COMPUTER REPRESENTATION OF RELATIONS AND DIGRAPHS

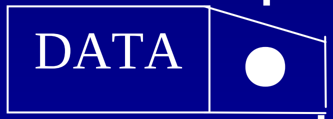
The most straightforward method of storing data items is to place them in a linear list or array. This generally corresponds to putting consecutive data items in consecutively numbered storage locations in a computer memory. Figure 4.28 illustrates this method for five data items $D_1,\ldots,D_5$. The linear array might be A and the data would be in locations A[1], A[2], A[3], A[4], A[5], and we would have access to any data item D_i by simply supplying its index i.

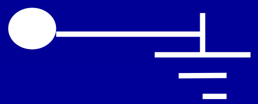
The main problem with this storage method is that we cannot insert new data between existing data without moving a possibly large number of items. To add another item E between D_2 and D_3, would have to move D_3 to A[4], D_4 to A[5], and D_5 to A[6], if room exists, and then assign E to A[3].

Figure 4.28



An alternative method of representing this sequence is by a **linked list**, shown in schematic fashion in Figure 4.29. The basic unit of information storage is the **storage cell**. Such cells to have room for two information items. The first can be data (numbers or symbols), and the second item is a **pointer**, that is, a number that tells us (points to) the location of the next cell to be considered.

We represent the storage cell as a partitioned box 

, with a dot in the right-hand side representing a pointer. A line is drawn from each such dot to the cell that the corresponding pointer designates as next. The symbol  means that data have ended and that no further pointers need be followed.

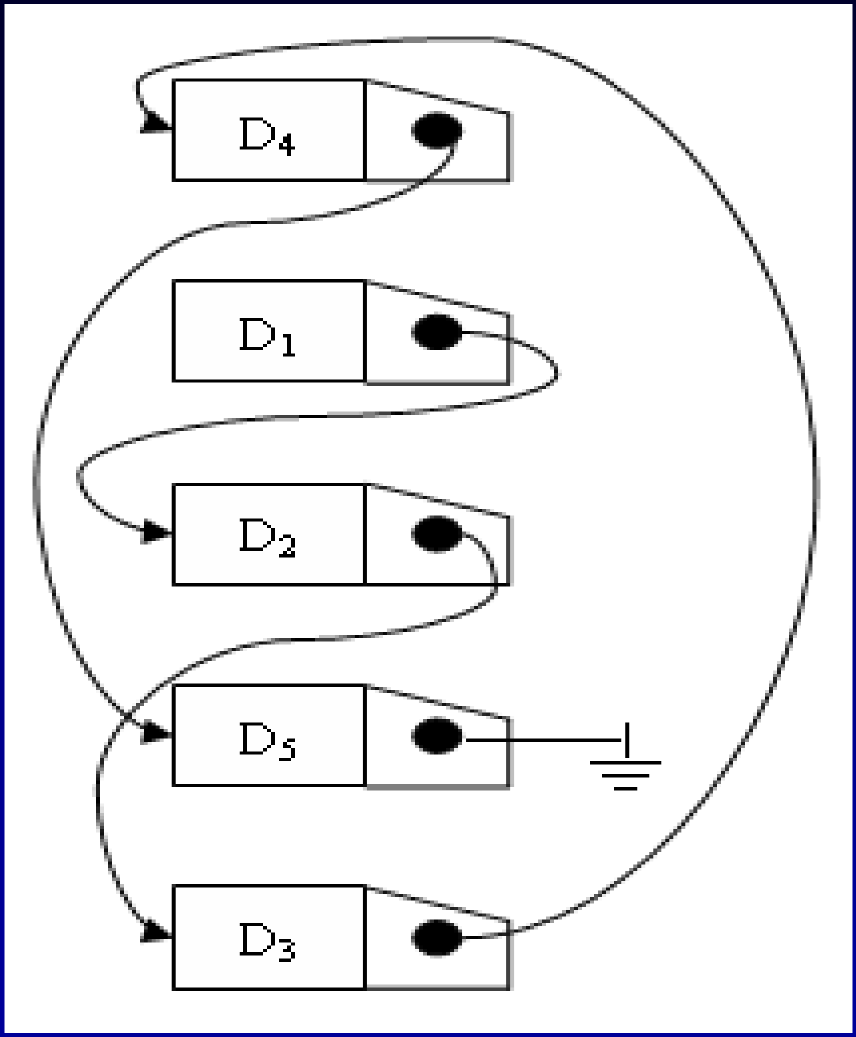


Figure 4.29

In practice, the concept of a linked list may be implemented using two linear arrays, a data array A and a pointer array P, as shown in Figure 4.30. We have accessed the data in location A[i], then the number in location P[i] gives, or points to, the index of A containing the next data item.

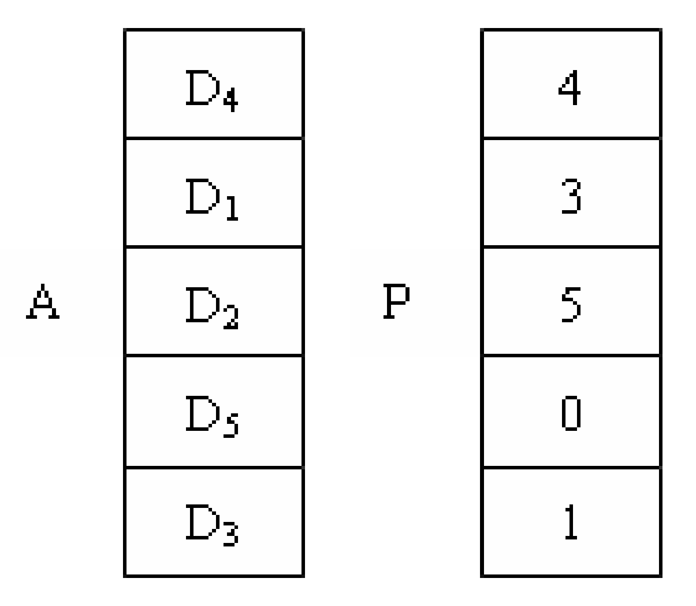


Figure 4.30

It does not matter how large the data item is, so A might actually be a two-dimensional array or matrix. The first row would hold several numbers describing the first data item, the second row would describe the next item.

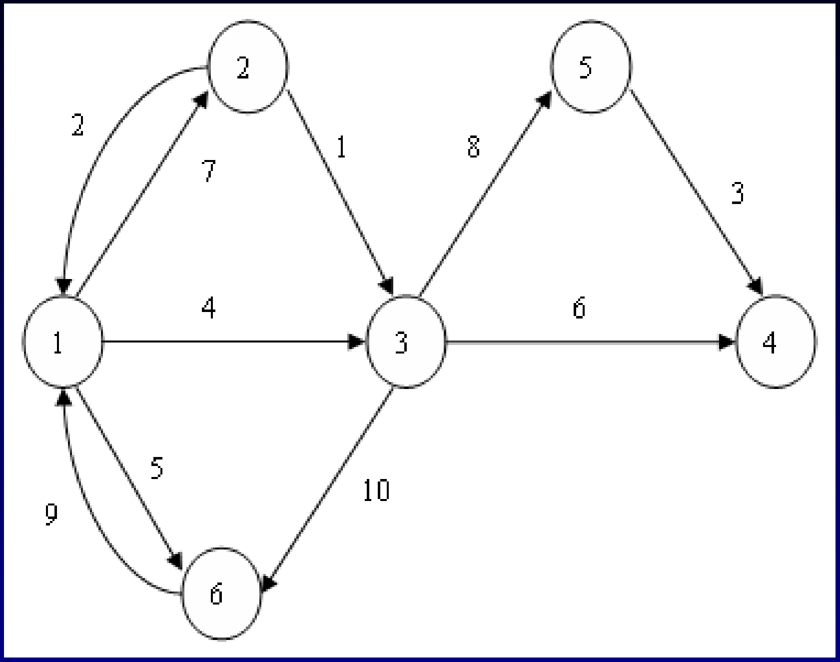
The problem of storing information to represent a relation or its digraph also has two solutions similar to those presented previously for simple data.

A relation R on A can be represented by an $n \times n$ matrix M_R if A has n elements.

Then a straightforward way of representing R in a computer would be by an $n \times n$ array having 0's and 1's stored in each location.

A second method of storing data for relations and digraphs uses the **linked list** idea. A linked list will be constructed that contains all the edges of the digraph. The data can be represented by two arrays, TAIL and HEAD, for all arrows. We will also need an array NEXT of pointers from each edge to the next edge.

Consider the relation whose digraph is shown in Figure 4.32. We can use a scheme such as that illustrated in Figure 4.33.



START	TAIL	HEAD	NEXT
2	1	3	9
	2	3	10
	2	1	4
	3	5	8
	5	4	1
	3	4	3
	3	6	0
	6	1	7
	1	6	6
	1	2	5

Figure 4.32 Figure 4.33

In many algorithms, it is efficient to locate a vertex and then immediately begin to investigate the edges that begin or end with this vertex. This is not possible in general with the storage mechanism shown in Figure 4.33, so we now give a modification of it. For each vertex I, VERT[I] is the index, in TAIL and HEAD, of the first edge we wish to consider leaving vertex I. Thus VERT, like NEXT, contains pointers to edges. For each vertex I, we must arrange the pointers in NEXT so that they link together all edges leaving I, starting with the edge pointed to by VERT[I].

This method is shown in Figure 4.34 for the digraph of Figure 4.32.

START	TAIL	HEAD	NEXT
10	1	2	0
2	2	3	3
4	2	1	0
0	3	5	6
5	5	4	0
8	3	4	7
	3	6	0
	6	1	0
	1	6	1
	1	3	9

Figure 4.34

We have (at least) two methods for storing the data for a relation or digraph, one using the matrix of the relation and one using linked lists. An analysis of such factors will determine which of the storage methods is superior. We will consider two cases.

Suppose that $A=\{1,2,\ldots,N\}$, and let R be a relation on A, whose matrix M_R is represented by the array MAT. R contains P ordered pairs so that MAT contains exactly P ones.

Adding (I,J) to R is accomplished by the statement

$$MAT[I,J] \leftarrow 1.$$

Consider the following algorithm, which assigns **RESULT** the value T (true) or F (false), depending on whether R is not transitive.

ALGORITHM TRANS

1. RESULT $\leftarrow$ T
2. FOR I=1 THROUGH N
   a. FOR J=1 THROUGH N
      1. IF(MAT[I,J]=1) THEN
         a. FOR K=1 THROUGH N
            1. IF(MAT[J,K]=1 and MAT[I,K]=0) THEN
               a. RESULT $\leftarrow$ F

END OF ALGORITHM TRANS

Here RESULT is originally set to T, and it is changed only if a situation is found where $(I,J) \in R$

and (J,K) $\in$ R, but (I,K) $\notin$ R.

We now provide a count of the number of steps required by algorithm TRANS. Observe that I and J each range from 1 to N. If (I,J) is not in R, we only perform the one test “IF MAT[I,J]=1,” which will be false, and the rest of the algorithm will not be executed. Since $N^2 - P$ ordered pairs do not belong to R, we have N^2-P steps that must be executed for such elements. If (I,J) $\in$ R, then the test “IF MAT[I,J]=1” will be true and an additional loop

a. FOR K=1 THROUGH N

1. IF (MAT[J,K]=1 and MAT[I,K]=0) THEN

a. RESULT←F

of N steps will be executed. Since R contains P ordered pairs, we have PN steps for such elements. Thus the total number of steps required by algorithm TRANS is

$$T_A=PN+(N^2-P).$$

Suppose that $P=kN^2$, where $0\leqslant k\leqslant 1$, since P must be between 0 and N2. Then algorithm TRANS tests for transitivity in

$$T_A=kN^3+(1-k)N^2$$

steps.

Now consider the same digraph represented by our linked-list scheme using VERT, TAIL, HEAD, and NEXT. First we deal with the problem of adding an edge (I,J).

ALGORITHM ADDEDGE

1. P ← P+1
2. TAIL[P]←I
3. HEAD[P]←J
4. NEXT[P]←VERT[I]
5. VERT[I]←P

END OF ALGORITHM ADDEDGE

Figure 4.36 shows the situation diagrammatically in pointer form, both before and after the addition of edge (I,J).

The matrix storage method has the advantage for the task for the task of adding an edge.

ALGORITHM NEWTRANS

1. RESULT ← T
2. FOR I=1 THROUGH N
   a. X ← VERT[I]
   b. WHILE(X ≠ 0)
      1. J ← HEAD[X]
      2. Y ← VERT[J]
      3. WHILE(Y ≠ 0)
         a. K ← HEAD[Y]
         b. TEST ← EDGE[I,K]
         c. IF (TEST) THEN
            1. Y ← NEXT[Y]

d. ELSE

1. RESULT$\leftarrow$F

2. Y$\leftarrow$NEXT[Y]

4. X$\leftarrow$NEXT[X]

END OF ALGORITHM NEWTRANS

Let us analyze the average number of steps that algorithm NEWTRANS takes to test for transitivity. Each of the P edges begins at unique vertex, so, on the average, P/N=D edges begin at a vertex. It is not hard to see that a function EDGE, such as needed in NEWTRANS, can be made to take an average of about D steps, since it must check all edges beginning at a particular vertex. The main FOR loop of NEWTRANS will be executed N times, and each subordinate WHILE statement will average about D executions. Since the last WHILE calls EDGE each time, we see that the entire algorithm will average about ND3 execution steps. As before, we suppose that $P=kN^2$ with $0 \leqslant k \leqslant 1$. Then NEWTRANS averages about

$$T_L = N\left(\frac{kN^2}{N}\right)^3 = k^3N^4 \quad \text{steps.}$$

Recall that algorithm TRANS, using matrix storage, required about $T_A = kN^3 + (1-k)N^2$ steps.

$$\frac{T_L}{T_A} = \frac{k^3 N^4}{kN^3 + (1-k)N^2} = \frac{k^2 N}{1 + \left(\frac{1}{k} - 1\right)\frac{1}{N}}.$$

When k is close to 1, that is, when there are many edges, then T_L/T_A is nearly N, so $T_L \approx T_A N$, and the linked-list method averages N times as many steps as the matrix-storage method. Thus the matrix-storage method is N times faster than the linked-list method in most cases.

On the other hand, if k is very small, then T_L/T_A may be nearly zero. This means that if the number of edges is small compared with N^2, it is, on average, considerably more efficient to test for transitivity in a linked-list storage method than with adjacency matrix storage.

4.7 OPERATIONS ON RELATIONS

Let R and S be relations from a set A to a set B. The complement of R, say $\overline{R}$, is referred to as the **complementary relation**. A relation from A to B that can be expressed simply in terms of R:

$$a \overline{R} b \text{ if and only if } a \not{R} b.$$

The intersection relation $R \cap S$: a $R \cap S$ b means that a R b and a S b.

The inverse relation R^{-1}: R^{-1} is a relation from B to A defined by

b R^{-1} a if and only if a R b.

Example Let A={1,2,3} and let R and S be relations on A. Suppose that the matrices of R and S are

$$M_R = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad M_S = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then we can verify that

$$M_{\overline{R}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad M_{R^{-1}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix},$$

$$M_{R \cap S} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad M_{R \cup S} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

We can show (Exercise 27) that if R and S are relations on set A, then

$$M_{R \cap S} = M_R \wedge M_S$$

$$M_{R \cup S} = M_R \vee M_S$$

$$M_R^{-1} = (M_R)^T.$$

If M is a Boolean matrix, we define the **complement** $\overline{M}$ of M as the matrix obtained from M by replacing every 1 in M by a 0 and every 0 by a 1.

$$M_{\overline{R}} = \overline{M}_R$$

A symmetric relation is a relation R such that $M_R=(M_R)^T$, R is symmetric if and only if $R=R^{-1}$.

Theorem 1: Suppose that R and S are relations from A to B.

(a) If $R \subseteq S$, then $R^{-1} \subseteq S^{-1}$.

(b) If $R \subseteq S$, then $\overline{S} \subseteq \overline{R}$.

(c) $(R \cap S)^{-1} = R^{-1} \cap S^{-1}$ and $(R \cup S)^{-1} = R^{-1} \cup S^{-1}$.

(d) $\overline{R \cap S} = \overline{R} \cup \overline{S}$ and $\overline{R \cup S} = \overline{R} \cap \overline{S}$.

Theorem 2: Let R and S be relations on a set A.

(a) If R is reflexive, so is R^{-1}.

(b) If R and S are reflexive, then so are $R \cap S$ and $R \cup S$.

(c) R is reflexive if and only if $\overline{R}$ is irreflexive.

Theorem 3: Let R be a relation on a set A. Then

(a) R is symmetric if and only if $R = R^{-1}$.

(b) R is antisymmetric if and only if $R \cap R^{-1} \subseteq \Delta$.

(c) R is asymmetric if and only if $R \cap R^{-1} = \varnothing$.

Theorem 4: Let R and S be relations on A.

(a) If R is symmetric, so are R^{-1} and $\overline{R}$.

(b) If R and S are symmetric, so are $R \cap S$ and $R \cup S$.

Theorem 5: Let R and S be relations on A.

(a) $(R \cap S)^2 \subseteq R^2 \cap S^2$.

(b) If R and S are transitive, so is $R \cap S$.

(c) If R and S are equivalence relations, so is $R \cap S$.

- Closures（闭包）

If R is a relation on a set A, it may well happen that R lacks some of the important relational properties. If R does not possess a particular property, we may wish to add pairs to R until we get a relation that does have the required property. We want to add as few pairs as possible, so what we need to find is the *smallest* relation R_1 on A that contains R and possesses the property we desire. If a relation such as R_1 does exist, we call it the **closure**（闭包）of R with respect to the property in question.

Example (1) The **reflexive closure** of R is $R \cup \triangle$.

(2) $R \cup R^{-1}$ is the **symmetric closure** of R.

The graph of the symmetric closure of R is simply the digraph of R with all edges made bidirectional.

The **transitive closure** of a relation R is the smallest transitive relation containing R.

Composition（复合）

A, B, and C are sets, R is a relation form A to B, and S is a relation from B to C. We can then define a new relation, the **composition** of R and S, written $S \circ R$. The relation $S \circ R$ is a relation from A to C and is defined as follows. If x is in A and z is in C, then $x\ (S \circ R)\ z$ if and only if for some y in B, we have $x\ R\ y$ and $y\ S\ z$.

Theorem 6: Let R be a relation from A to B and S a relation from B to C. Then, if A_1 is any subset of A, we have

$$(S \circ R)(A_1)=S(R(A_1)).$$

Let A, B, and C be finite sets with n, p and m elements, respectively, let R be a relation from A to B, and let S be a relation from B to C. Then R and S have Boolean matrices M_R and M_S with respective sizes $n \times p$ and $p \times m$. $M_R \odot M_S$ can be computed, and it equals $M_{S \circ R}$.

To see this let $A=\{a_1,\ldots,a_n\}$, $B=\{b_1,\ldots,b_p\}$, and $C=\{c_1,\ldots,c_m\}$. Also, suppose that $M_R=[r_{ij}]$, $M_S=[s_{ij}]$, and $M_{S \circ R}=[t_{ij}]$. Then $t_{ij}=1$ if and only if $(a_i,c_j) \in S \odot R$, which means that for some k, $(a_i,b_k) \in R$ and $(b_k,c_j) \in S$. In other words, $r_{ik}=1$and $s_{kj}=1$ for some k between 1 and p. This condition is identical to the condition needed for $M_R \odot M_S$ to have a 1 in position i, j, and thus $M_{S \circ R}$ and $M_R \odot M_S$ are equal.

In the special case where R and S are equal, we have $S \circ R = R^2$ and $M_{S \circ R} = M_{R^2} = M_R \odot M_S$, as was shown in Section 4.3.

Theorem 7: Let A, B, C, and D be sets, R a relation from A to B, S a relation from B to C, and T a relation from C to D. Then

$$T \circ (S \circ R) = (T \circ S) \circ R.$$

Proof: The relations R, S, and T are determined by their Boolean matrices M_R, M_S and M_T, respectively. As we showed after Example 11, the matrix of the composition is the Boolean matrix product; that is, $M_{S \circ R} = M_R \odot M_S$. Thus

$$M_{T\circ(S\circ R)}=M_{S\circ R}\odot M_T=(M_R\odot M_S)\odot M_T.$$

Similarly,

$$M_{T\circ(S\circ R)}=M_R\odot(M_S\odot M_T).$$

Since Boolean matrix multiplication is associative [see Exercise 37 of Section 1.5], we must have

$$(M_R\odot M_S)\odot M_T=M_R\odot(M_S\odot M_T),$$

and therefore

$$M_{T\circ(S\circ R)}=M_{(T\circ S)\circ R}.$$

Then

$$T\circ(S\circ R)=(T\circ S)\circ R$$

since these relations have the same matrices.

Theorem 8: Let A, B, and C be sets, R a relation from A to B, and S a relation from B to C. Then

$$(S \circ R)^{-1} = R^{-1} \circ S^{-1}.$$

Proof: Let $c \in C$ and $a \in A$. Then $(c,a) \in (S \circ R)^{-1}$ if and only if $(a,c) \in S \circ R$, that is, if and only if there is a $b \in B$ with $(a,b) \in R$ and $(b,c) \in S$. Finally, this is equivalent to the statement that $(c,b) \in S^{-1}$ and $(b,a) \in R^{-1}$; that is, $(c,a) \in R^{-1} \circ S^{-1}$.

4.8 TRANSITIVE CLOSURE AND WARSHALL'S ALGORITHM

- Transitive Closure

Theorem 1: Let R be a relation on a set A. Then R^{∞} is the transitive closure of R.

Proof: We recall that if a and b are in the set A, then a R^{∞} b if and only if there is a path in R^{∞} from a to b. Now R^{∞} is certainly transitive since, if a R^{∞} b and b R^{∞} c, the composition of the paths from a to b and then from b to c forms a path from a to c in R, and so a R^{∞} c. To show that R is the smallest transitive relation containing R, we must show that if S is any transitive relation on A and $R \subseteq S$, then $R^{\infty} \subseteq S$. Theorem 1 of Section 4.4 tells us that if S is transitive, then $S^n \subseteq S$ for all n; that is, if a and b are connected by a path of length n, then a S b. It follows that $S^{\infty} = \bigcup_{n=1}^{\infty} S^n \subseteq S$. It is also true that if $R \subseteq S$, then $R^{\infty} \subseteq S^{\infty}$, since any path in R is also a path in S. Putting these facts together, we see that if $R \subseteq S$ and S is transitive on A, then $R^{\infty} \subseteq S^{\infty} \subseteq S$. This means that R^{∞} is the smallest of all transitive relations on A that contain R.

From a geometric point of view, R^{∞} is called the **connectivity relation**. If we include the relation $\triangle$ (see Section 4.4), then $R^{\infty} \cup \triangle$ is the reachability relation R* (see Section 4.3). On the other hand, from the algebraic point of view, R^{∞} is the transitive closure of R.

Example   Let A={1,2,3,4}, and let

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

Find the transitive closure of R.

METHOD 1: The digraph of R is shown in Figure 4.42. Since R^∞ is the transitive closure, we can proceed geometrically by computing all paths. So we have

$R^\infty=\{(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(2,3),(2,4),(3,4)\}$.

METHOD 2: The matrix of R is

$$M_R = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

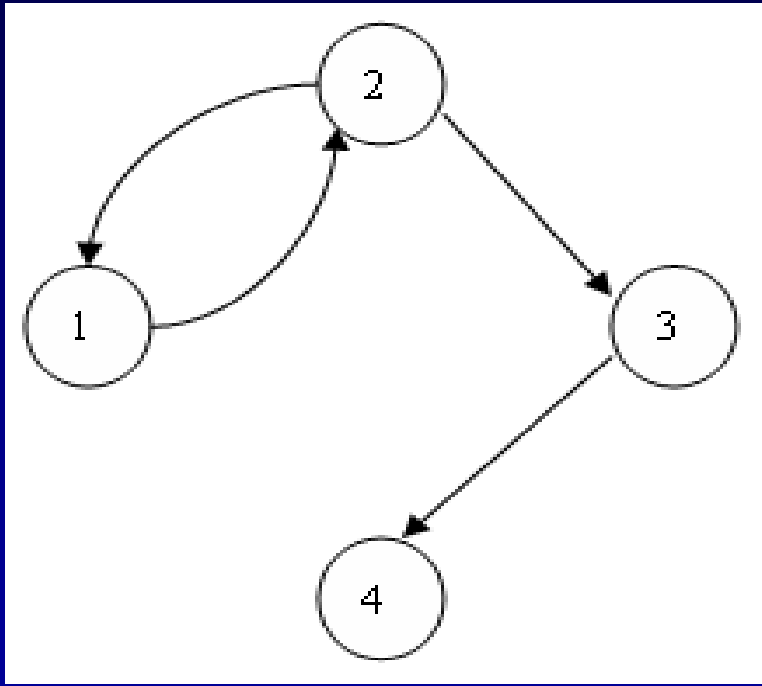


Figure 4.42

We may proceed algebraically and compute the powers of M_R. Thus

$$(M_R)^2_\odot = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (M_R)^3_\odot = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$(M_R)^4_\odot = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Continuing in this way, we can see that $(M_R)^n_\odot$ equals $(M_R)^2_\odot$ if n is **even** and equals $(M_R)^3_\odot$ if n is **odd** and greater than 1. Thus

$$M_{R^{\infty}} = M_R \vee (M_R)^2_{\odot} \vee (M_R)^3_{\odot} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and this gives the same relation as Method 1.

Theorem 2: Let A be a set with $|A|=n$, and let R be a relation on A. Then

$$R^{\infty}=R \cup R^2 \cup \cdots \cup R^n.$$

In other words, powers of R greater than n are not needed to compute R^{∞}.

Proof: Let a and b be in A, and suppose that a, x_1, x_2, …, x_m, b is a path from a to b in R; that is, (a,x_1), (x_1, x_2),…,(x_m, b) are all in R. If x_i and x_j are the

same vertex, say $i<j$, then the path can be divided into three sections. First, a path from a to x_i, then a path from x_i to x_j, and finally a path from x_j to b. The middle path is a cycle, since $x_i=x_j$, so we simply leave it out and put the remaining two paths together. This gives us a shorter path from a to b (see Figure 4.43).

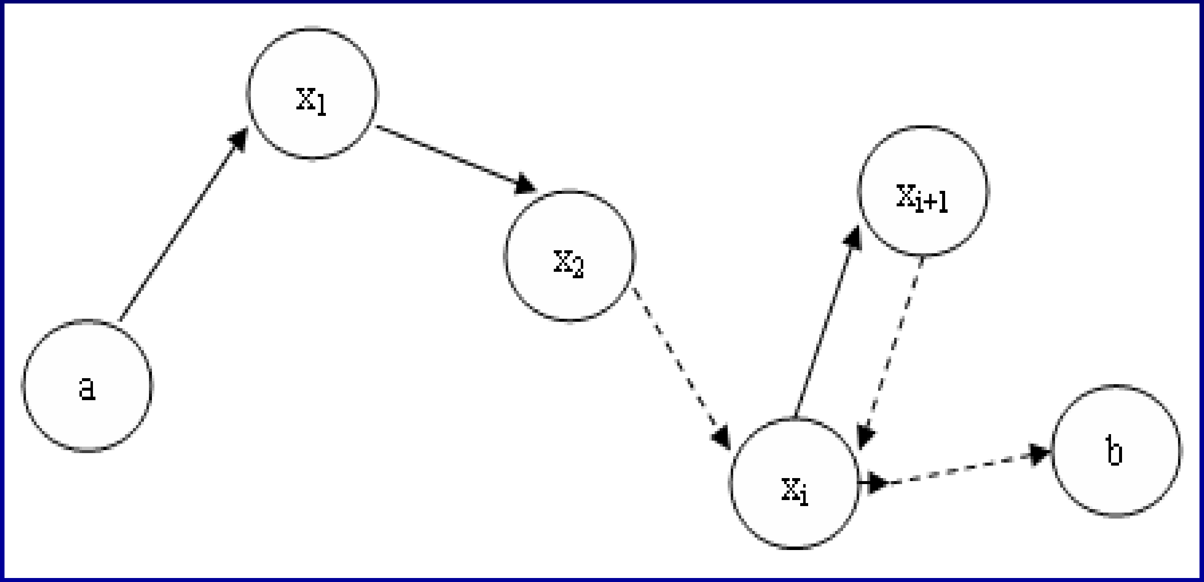


Figure 4.43

Now let a, x_1, x_2, …, x_k, b be the shortest path from a to b. If $a \neq b$, then all vertices a, x_1, x_2, …, x_k, b are distinct. Otherwise, the preceding discussion shows that we could find a shorter path. Thus the length of the path is at most n-1 (since $|A|=n$). If $a=b$, then for similar reasons, the vertices a, x_1, x_2, …, x_k are distinct, so the length of the path is at most n. In other words, if a R^{∞} b, then a R^k b, for some k, $1 \leqslant k \leqslant n$. Thus $R^{\infty}=R \cup R^2 \cup \cdots \cup R^n$.

Warshall's Algorithm

Let R be be a relation on a set $A=\{a_1, a_2,\ldots,a_n\}$. If $x_1, x_2,\ldots,x_m$ is a path in R, then any vertices other than x_1 and x_m are called **interior vertices** of the path. For $1\leqslant k\leqslant n$, we define a Boolean matrix W_k as follows. W_k has a 1 in position i, j if and only if there is a path from a_i to a_j in R whose interior vertices, if any, come from the set $\{a_1,a_2,\ldots,a_k\}$.

The matrix W_n has a 1 in position i, j if and only if some path in R connects a_i with a_j. In other words, $W_n=M_R^{\infty}$. If we define W_0 to be M_R, then we will have a sequence $W_0, W_1, \ldots,W_n$ whose first term is M_R and whose last term is M_R^{∞}.

This procedure is called **Warshall's algorithm**.

The matrices W_k are **different** from the powers of the matrix M_R.

Suppose that $W_k=[t_{ij}]$ and $W_{k-1}=[s_{ij}]$. If $t_{ij}=1$, then there must be a path form a_i to a_j whose interior vertices come from the set $\{a_1, a_2,\ldots,a_k\}$. If the vertex a_k is not an interior vertex of this path, then all interior vertices must actually come from the set $\{a_1, a_2,\ldots,a_{k-1}\}$, so $s_{ij}=1$. If a_k is an interior vertex of the path, then the situation is as shown in Figure 4.44. We may assume that all interior vertices are distinct. Thus a_k appears only once in the path, so all interior vertices of subpaths 1 and 2 must come from the set $\{a_1, a_2,\ldots,a_{k-1}\}$. This means that $s_{ik}=1$

and $s_{kj}=1$.

Thus $t_{ij}=1$ if and only if either

(1) $s_{ij}=1$ or

(2) $s_{ik}=1$ and $s_{kj}=1$.

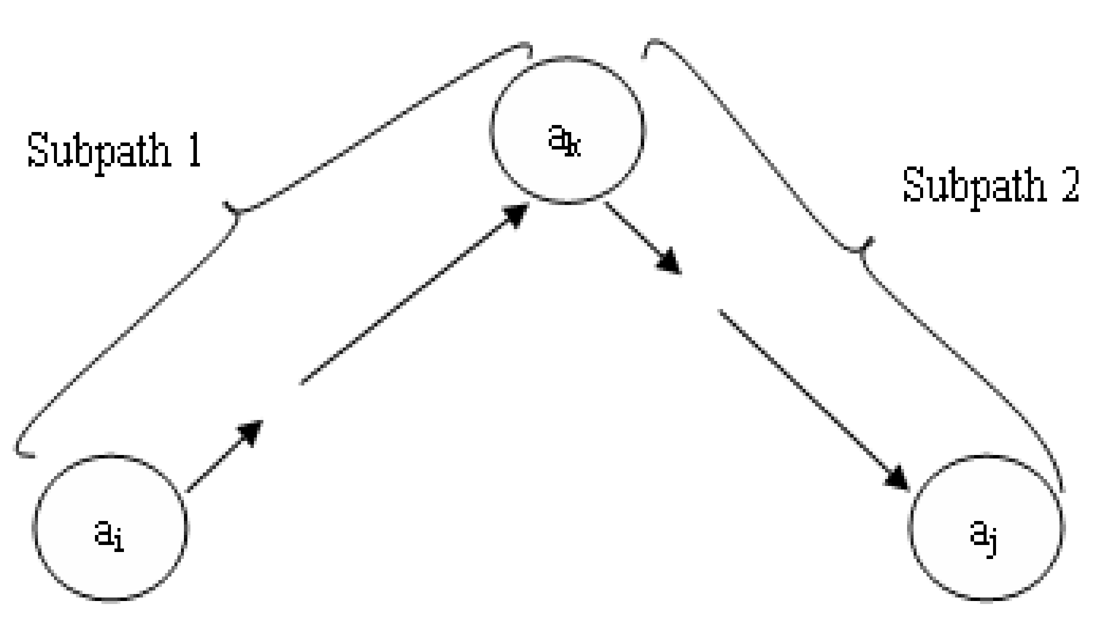


Figure 4.44

This is the basis for Warshall's algorithm. If W_{k-1} has a 1 in position i, j, then, by (1), so will W_k. By (2), a new 1 can be added in position i, j of W_k if and only if column k of W_{k-1} has a 1 in position i and row k of W_{k-1} has a 1 in position j. Thus we have the following procedure for computing W_k from W_{k-1}.

Step 1: First transfer to W_k all 1's in W_{k-1}.

Step 2: List the locations p_1, p_2, …, in column k of W_{k-1}, where the entry is 1, and the locations q_1, q_2, …, in row k of W_{k-1}, where the entry in 1.

Step 3: Put 1's in all the positions p_i, q_j of W_k (if they are not already there).

Consider the relation R defined in Example 1. Then

$$W_0 = M_R = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad W_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad W_2 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$W_3 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad M_{R^\infty} = W_4 = W_3 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

ALGORITHM WARSHALL

1. CLOSURE $\leftarrow$ MAT
2. FOR K=1 THROUGH N
   a. FOR I=1 THROUGH N
      1. FOR J=1 THROUGH N
         a. CLOSURE[I,J] $\leftarrow$ CLOSURE[I,J] $\vee$ (CLOSURE[I,K] $\wedge$ CLOSURE[K,J])

END OF ALGORITHM WARSHALL

If we think of the testing and assignment line as one step, then algorithm WARSHALL requires n^3 steps in all. The Boolean product of two $n \times n$ Boolean matrices A and B also requires n^3 steps. To compute all products $(M_R)_{\odot}^{2}, (M_R)_{\odot}^{3}, \cdots, (M_R)_{\odot}^{n}$, we require $n^3(n-1)$ steps, since we will need n-1 matrix multiplications. The formula

$$M_{R^\infty} = M_R \vee \left(M_R\right)_\odot^2 \vee \cdots \vee \left(M_R\right)_\odot^n \qquad (1)$$

would require about n^4 steps without the final joins.

An interesting application of the transitive closure is to equivalence relations.

Theorem 3: If R and S are equivalence relations on a set A, then the smallest equivalence relation containing both R and S is $(R \cup S)^\infty$.

Proof: Recall that $\triangle$ is the relation of equality on A and that a relation is reflexive if and only if it contains $\triangle$. Then $\Delta \subseteq R, \Delta \subseteq S$ since both are reflexive, so $\Delta \subseteq R \cup S \subseteq (R \cup S)^{\infty}$, and $(R \cup S)^{\infty}$ is also reflexive.

Since R and S are symmetric, $R=R^{-1}$ and $S=S^{-1}$, so $(R \cup S)^{-1}=R^{-1} \cup S^{-1}=R \cup S$, and $R \cup S$ is also symmetric. Because of this, all paths in $R \cup S$ are "two-way streets," and it follows from the definitions that $(R \cup S)^{\infty}$ must also be symmetric. Since we already know that $(R \cup S)^{\infty}$ is transitive, it is an equivalence relation containing $R \cup S$. It is the smallest one, because no smaller set containing $R \cup S$ can be transitive, by definition of the transitive closure.

Example Let $A=\{1,2,3,4,5\}$, $R=\{(1,1), (1,2), (2,1), (2,2), (3,3), (3,4), (4,3), (4,4), (5,5)\}$, and $S=\{(1,1), (2,2), (3,3), (4,4), (4,5), (5,4), (5,5)\}$. We may verify that both R and S are equivalence relations. The partition A/R of A corresponding to R is $\{\{1,2\}, \{3,4\}, \{5\}\}$, and the partition A/S of A corresponding to S is $\{\{1\},\{2\},\{3\},\{4,5\}\}$. Find the smallest equivalence relation containing R and S, and compute the partition of A that it produces.

Solution: We have

$$M_R = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } M_S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix},$$

so

$$M_{R \cup S} = M_R \vee M_S = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

We now compute $M_{(R \cup S)}^{\infty}$ by Warshall's algorithm.

First, $W_0=M_{R \cup S}$. $W_1=W_0$. $W_2=W_1$. $W_3=W_2$.

$$W_4 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

$W_5=W_4$.

$(R \cup S)^{\infty}=\{(1,1),(1,2),(2,1),(2,2),(3,3),(3,4),(3,5),$
$(4,3),(4,4),(4,5),(5,3),(5,4),(5,5)\}$.

The corresponding partition of A is then (verify)

{{1,2},{3,4,5}}.

The end !