Lecture 15

n-ary Relations

Definition: An n-ary relation R on sets $A_1, ..., A_n$, written as $R: A_1, ..., A_n$, is a subset $R \subseteq A_1 \times \cdots \times A_n$.

- The sets $A_1, ..., A_n$ are called the domains of R.
- The degree of R is n.
- R is functional in domain A_i if it contains at most one n-tuple $(\cdots, a_i, \cdots)$ for any value a_i within domain A_i.

Transitive Relation and R^n

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

Proof:

- "if" part: In particular, $R^2 \subseteq R$. If $(a,b) \in R$ and $(b,c) \in R$, then by the definition of composition, we have $(a,c) \in R^2 \subseteq R$.
- "only if: part: by induction.
  - $n = 1$: $R^1 \subseteq R$
  - Suppose $R^n \subseteq R$:
    - $(a,c) \in R^{n+1} \triangleq R^n \circ R$: there is a $b \in A$ such that $(a,b) \in R$ and $(b,c) \in R^n \subseteq R$
    - Since R is transitive, $(a,b) \in R$ and $(b,c) \in R^n \subseteq R$ implies that $(a,c) \in R$

Relational Databases

A domain A_i is a primary key for the database if the relation R is functional in A_i.

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 composite key for the n-ary relation, assuming that no n-tuples are ever

Selection Operator

Let A be any n-ary domain $A = A_1 \times \cdots \times A_n$, and let $C: A \to \{T, F\}$ be any condition (predicate) on elements (n-tuples) of A.

The selection operator s_C is the operator that maps any (n-ary) relation R on A to the n-ary relation of all n-tuples from R that satisfy C.

$$\forall R \subseteq A, s_C(R) = R \cap \{a \in A | s_C(a) = T\} = \{a \in R | s_C(a) = T\}$$

Selection Operator: Example

Suppose that we have a domain

$$A = \text{StudentName} \times \text{Standing} \times \text{SocSecNos}$$

Suppose that we have a condition

$$UpperLevel(name, standing, ssn)$$

Projection Operator

$:\equiv [(standing = junior) \vee (standing = senior)]$

Let $A = A_1 \times \cdots \times A_n$ be any n-ary domain, and let $\{i_k\} = (i_1, ..., i_m)$ be a sequence of indices all falling in the range 1 to n. That is, where $1 \le i_k \le n$ for all $1 \le k \le m$.

Then the projection operator on n-tuples $P_{i_k}: A \to A_{i_1} \times \cdots \times A_{i_m}$ is defined by $P_{i_k}(a_1, \cdots, a_n) = (a_{i_1}, \cdots, a_{i_m})$

Example $P_{1,2}$

Student	Major	Course
Glauser	Biology	BI 290
Glauser	Biology	MS 475
Glauser	Biology	PY 410
Marcus	Mathematics	MS 511
Marcus	Mathematics	MS 603
Marcus	Mathematics	CS 322
Miller	Computer Science	MS 575
Miller	Computer Science	CS 455

Student	Major
Glauser	Biology
Marcus	Mathematics
Miller	Computer Science

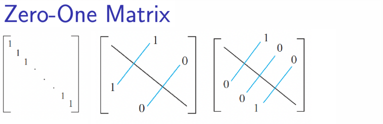
Join Operator $J(R_1, R_2)$

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.

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.

Zero-One Matrix



Reflexive Symmetric Antisymmetric

Join and Meet

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ zero-one matrices.

The join of A and B is the zero-one matrix with (i,j)-th entry $a_{ij} \vee b_{ij}$. The join of A and B is denoted by $A \vee B$.

The meet of A and B is the zero-one matrix with (i,j)-th entry $a_{ij} \wedge b_{ij}$. The meet of A and B is denoted by $A \wedge B$.

$$M_{R_1 \cup R_2} = M_{R_1} \vee M_{R_2}$$

$$M_{R_1 \cap R_2} = M_{R_1} \wedge M_{R_2}$$

Zero-One Matrix: Composite of Relations

Let $A = [a_{ij}]$ be an $m \times k$ zero-one matrix and $B = [b_{ij}]$ be a $k \times n$ zero-one matrix. Then, the Boolean product of A and B, denoted by $A \odot B$, is the $m \times n$ matrix with (i,j)-th entry c_{ij} where

$$c_{ij} = (a_{i1} \wedge b_{1j}) \vee (a_{i2} \wedge b_{2j}) \vee \cdots \vee (a_{ik} \wedge b_{kj}).$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$M_{S \circ R} = M_R \odot M_S$.

The ordered pair (a_i, c_j) belongs to $S \circ R$ 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.

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

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

Closures of Relations

Let $R = \{(1,1), (1,2), (2,1), (3,2)\}$ on $A = \{1,2,3\}$.

Is this relation R reflexive?

No. $(2,2)$ and $(3,3)$ are not in R.

The question is what is the minimal relation $S \supseteq R$ that is reflexive?

How to make R reflexive by minimum number of additions?

Add $(2,2)$ and $(3,3)$

Then $S = \{(1,1), (1,2), (2,1), (3,2), (2,2), (3,3)\} \supseteq R$.

The minimal set $S \supseteq R$ is called the reflexive closure of R.

The set S is called the reflexive closure of R if it:

- contains R
- is reflexive
- is minimal (is contained in every reflexive relation Q that contains R ($R \subseteq Q$), i.e., $S \subseteq Q$)

Relations can have different properties:

- reflexive
- symmetric
- transitive

We define:

- reflexive closures
- symmetric closures
- transitive closures

S is the minimal set containing R satisfying the property P.

Example: $R = \{(1,2), (2,3), (2,2)\}$ on $A = \{1,2,3\}$. What is the symmetric closure S of R?

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

What is the transitive closure S of R?

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

Transitive Closure

Example: $R = \{(1,2), (2,2), (2,3)\}$ on $A = \{1,2,3\}$. Transitive closure: $S = \{(1,2), (2,2), (2,3), (1,3)\}$

Paths in Directed Graphs

Definition: A path from a to b in the directed graph G is a sequence of edges $(x_0, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n)$ in G, where n is nonnegative integer and $x_0 = a$ and $x_n = b$.

A path of length $n \ge 1$ that begins and ends at the same vertex is called a circuit or cycle.

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

Connectivity Relation

Definition: Let R be a relation on a set A. The connectivity relation R^* consists of all pairs (a,b) such that there is a path (of any length) between a and b in R:

$$R^* = \bigcup_{k=1}^{\infty} R^k$$

$A = \{1,2,3,4\}$

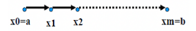
$R = \{(1,2), (1,3), (1,4), (2,3), (3,4)\}$, $R^2 = \{(1,3), (2,4), (1,4)\}$

$R^3 = \{(1,4)\}$, $R^4 = \emptyset$

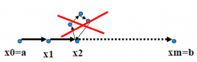
$R^* = \{(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)\}$

Lemma: Let A be a set with n elements, and R a relation on A. If there is a path from a to b with $a \ne b$, then there exists a path of length $\le n-1$.

Proof (by intuition): There are at most n different elements we can visit on a path if the path does not have loops:



Loops may increase the length but the same node is visited more than once



Proof: Suppose there is a path from a to b in R. Let m be the length of the shortest such path. Suppose that $x_0, x1, x2, ..., x_m$, where $x_0 = a$ and $x_m = b$, is such a path.

Suppose that $a \ne b$ and that $m \ge n$. The $m+1$ vertices are from n elements. According to the pigeonhole principle and $a \ne b$, at least two of the vertices $x_0, x_1, ..., x_{m-1}$ are equal.

There is a circuit that can be deleted until the length is $< n$.

Lemma: Let A be a set with n elements, and R a relation on A. If there is a path from a to b with $a \ne b$, then there exists a path of length $\le n-1$.

Lemma: 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.

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

$$R^* = \bigcup_{k=1}^{\infty} R^k$$

- R^* is transitive

  If $(a,b) \in R^*$ and $(b,c) \in R^*$, then there are paths from a to b and from b to c in R. Thus, there is a path from a to c in R. This means that $(a,c) \in R^*$.
- $R^* \subseteq S$ whenever S is a transitive relation containing R
  - Suppose that S is a transitive relation containing R.
  - $S^n \subseteq S$ for integer $n \ge 1$. (Recall S is transitive iff $S^n \subseteq S$).
  - We have $S^* \subseteq S$.
  - If $R \subseteq S$, then $R^* \subseteq S^*$, because any path in R is also a path in S.
  - Thus, $R^* \subseteq S^* \subseteq S$.

Find Transitive Closure

Recall that 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. Thus,

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

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

$$M_{R^*} = M_R \vee M_R^{[2]} \vee M_R^{[3]} \vee \cdots \vee M_R^{[n]},$$

where $M_R^{[n]} = \underbrace{M_R \odot M_R \odot \cdots \odot M_R}_{n\ M_R's}$

ALGORITHM 1 A Procedure for Computing the Transitive Closure.

```
procedure transitive closure (M_R : zero–one n × n matrix)
A := M_R
B := A
for i := 2 to n
    A := A ⊙ M_R
    B := B ∨ A
return B{B is the zero–one matrix for R*}
```

- $n-1$ Boolean products
- Each of these Boolean products use $n^2(2n-1)$ bit operations.
- $O(n^4)$ bit operations.

16 Equivalence Relation

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

Equivalence Class

Definition: Let R be an equivalence relation on a set A. The set of all elements that are related to an element a of A is called the equivalence class of a, denoted by $[a]_R$. When only one relation is considered, we use the notation $[a]$.

$$[a]_R = \{b : (a,b) \in R\}$$

Theorem: Let R be an equivalence relation on a set A. The following statements are equivalent:

(i) aRb (ii) $[a] = [b]$ (iii) $[a] \cap [b] \ne \emptyset$

Partition of a Set S

Definition: Let S be a set. A collection of nonempty subsets of S, i.e. A_1, A_2, . . . , A_k, is called a partition of S if:

$$A_i \cap A_j = \emptyset, i \ne j \text{ and } S = \bigcup_{i=1}^{k} A_i$$

Theorem: Let R be an equivalence relation on a set A. Then, union of all the equivalence classes of R is A:

$$A = \bigcup_{a \in A} [a]_R$$

Theorem: The equivalence classes form a partition of A.

Theorem: Let $\{A_1, A_2, ..., A_i, ...\}$ be a partition of S. Then, there is an equivalence relation R on S, that has the sets A_i as its equivalence classes.

Partial Ordering

Definition: A relation R on a set S is called a partial ordering, or partial order, if it is reflexive, antisymmetric, and transitive.

A set S together with a partial ordering R is called a partially ordered set, or poset, denoted by (S, R). Members of S are called elements of the poset.

$S = \{1,2,3,4,5,6\}$, R denotes the "$|$" relation

- Is R reflexive? Yes
- Is R antisymmetric? Yes
- Is R transitive? Yes

R is a partial ordering

Comparability

The notation $a \preccurlyeq b$ is used to denote that $(a,b) \in R$ in an arbitrary poset (S, R).

The notation $a \prec b$ denotes that $a \preccurlyeq b$, but $a \ne b$.

Definition: The elements a and b of a poset $(S, \preccurlyeq)$ are comparable if either $a \preccurlyeq b$ or $b \preccurlyeq a$. Otherwise, a and b are called incomparable.

Example: $S = \{1,2,3,4,5,6\}$, R denotes the "$|$" relation. $2,4$ are comparable, $3,5$ are incomparable.

Total Ordering

Definition: If $(S, \preccurlyeq)$ is a poset and every two elements of S are comparable, S is called a totally ordered set, and $\preccurlyeq$ is called a total order or a linear order. A totally ordered set is also called a chain.

Example: $S = \{1,2,3,4,5,6\}$, R denotes the "$\ge$" relation S is a chain.

Well-Ordered Set

$(S, \preccurlyeq)$ is a well-ordered set if it is a poset such that $\preccurlyeq$ is a total ordering and every nonempty subset of S has a least element.

Example: The set of ordered pairs of positive integers, $\mathbf{Z}^+ \times \mathbf{Z}^+$, with (a_1, a_2), (b_1, b_2) if $a_1 < b_1$, or if $a_1 = b_1$ and $a_2 \le b_2$ (the lexicographic ordering), is a well-ordered set.

The set Z, with the usual $\le$ ordering, is not well-ordered because the set of negative integers, which is a subset of $\mathbf{Z}$, has no least element.

Partial Ordering

- reflexive, antisymmetric, and transitive.
- e.g., $\le$, $|$
- $a \preccurlyeq b$ denotes $(a,b) \in R$ in a poset (S, R); $(S, \preccurlyeq)$

Comparable

- if either $a \preccurlyeq b$ or $b \preccurlyeq a$
- e.g., $S = \{1,2,3,4,5,6\}$, R denotes the "$|$" relation: $2,4$ are comparable, $3,5$ are incomparable.

Total Ordering

- $(S, \preccurlyeq)$ is a poset and every two elements of S are comparable
- "$\le$" is a total order, "$|$" is not a totally order

Well-ordered set

- total ordering; every nonempty subset of S has a least element
- e.g., $a \preccurlyeq b \preccurlyeq c...$; exists an a such that $(a,b) \in R$ for all $b \in S$

The Principle of Well-Ordered Induction

The Principle of Well-Ordered Induction: Suppose that $(S, \preccurlyeq)$ is a well-ordered set. Then $P(x)$ is true for all $x \in S$, if

Inductive Step: For every $y \in S$, if $P(x)$ is true for all $x \in S$ with $x \prec y$, then $P(y)$ is true.

Note: Suppose x_0 is the least element of a well ordered set, the inductive step tells us that $P(x_0)$ is true. We do not need a basis step.

Proof: Suppose it is not the case that $P(x)$ is true for all $x \in S$. Then there is an element $y \in S$ such that $P(y)$ is false.

Consequently, the set $A = \{x \in S | P(x) \text{is false}\}$ is nonempty. Because S is well ordered, A has a least element a.

By the choice of a as a least element of A, we know that $P(x)$ is true for all $x \in S$ with $x \prec a$. By the inductive step, $P(a)$ is true.

This contradiction shows that $P(x)$ must be true for all $x \in S$.

Lexicographic Ordering

Definition: Given two posets $(A_1, \preccurlyeq_1)$ and $(A_2, \preccurlyeq_2)$, the lexicographic ordering on $A_1 \times A_2$ is defined by specifying that (a_1, a_2) is less than (b_1, b_2), i.e., $(a_1, a_2) \prec (b_1, b_2)$, either if $a_1 \prec_1 b_1$ or if $a_1 = b_1$ then $a_2 \prec_2 b_2$.

Example: Consider strings of lowercase English letters. A lexicographic ordering can be defined using the ordering of the letters in the alphabet.This is the same ordering as that used in dictionaries.

- discreet $\prec$ discrete
- discreet $\prec$ discreetness

Hasse Diagram

Remove the loops (a,a) present at every vertex due to the reflexive property.

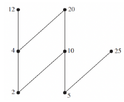
Remove all edges (x,y) for which there is an element $z \in S$ s.t. $x \prec z$ and $z \prec y$. These are the edges that must be present due to the transitive property.

Arrange each edge so that its initial vertex is below the terminal vertex. Remove all the arrows, because all edges point upwards toward their terminal vertex.

Maximal and Minimal Elements

Definition: a is a maximal (resp. minimal) element in poset $(S, \preccurlyeq)$ if there is no $b \in S$ such that $a \prec b$ (resp. $b \prec a$).

Example: Which elements of the poset $(\{2,4,5,10,12,20,25\}, |)$ are maximal, and which are minimal?



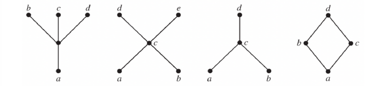
The maximal elements are 12, 20, and 25.

The minimal elements are 2 and 5.

A poset can have more than one maximal element and more than one minimal element.

Greatest and Least Elements

Definition: a is the greatest (resp. least) element of the poset $(S, \preccurlyeq)$ if $b \preccurlyeq a$ (resp. $a \preccurlyeq b$) for all $b \in S$.



Upper and Lower Bound

Definition: Let A be a subset of a poset $(S, \preccurlyeq)$.

- $u \in S$ is called an upper bound (resp. lower bound) of A if $a \preccurlyeq u$ (resp. $u \preccurlyeq a$) for all $a \in A$.
- $x \in S$ is called the least upper bound (resp. greatest lower bound) of A if x is an upper bound (resp. lower bound) that is less than any other upper bounds (resp. lower bounds) of A.

Find the greatest lower bound and the least upper bound of $\{b, d, g\}$, if they exist.

g is the least upper bound, b is the greatest lower bound.

Lattices

Definition: A partial ordered set in which every pair of elements has both a least upper bound and a greatest lower bound is called a lattice.

Determine whether the posets $(\{1,2,3,4,5\}, |)$ and $(\{1,2,4,8,16\}, |)$ are lattices.

Solution: Because 2 and 3 have no upper bounds, they certainly do not have a least upper bound. Hence, the first poset is not a lattice.

Every two elements of the second poset have both a least upper bound and a greatest lower bound.

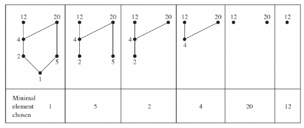
- The least upper bound of two elements in this poset is the larger of the elements
- The greatest lower bound of two elements is the smaller of the elements

Hence, this second poset is a lattice.

Topological Sorting

Topological Sorting for Finite Posets

Find a compatible total ordering for the poset $(\{1,2,4,5,12,20\}, |)$.



This produces the total ordering

$$1 \prec 5 \prec 2 \prec 4 \prec 20 \prec 12$$

Lecture 17:

Definition of a Graph

Definition: A graph $G = (V, E)$ consists of a nonempty set V of vertices (or nodes) and a set E of edges. Each edge has either one or two vertices associated with it, called its endpoints. An edge is said to be incident to (or connect) its endpoints.

simple graph: A graph in which each edge connects two different vertices and where no two edges connect the same pair of vertices.

Multigraph: Graphs that may have multiple edges connecting the same vertices.

Pseudograph: Graphs that may include loops, and possibly multiple edges connecting the same pair of vertices or a vertex to itself.

Directed and Undirected Graph

A directed graph (or digraph) (V, E) consists of a nonempty set of vertices V and a set of directed edges (or arcs) E. The directed edge associated with the ordered pair (u, v) is said to start at u and end at v.

Undirected Graphs

Definition: Two vertices u, v in an undirected graph G are called adjacent (or neighbors) in G if there is an edge e between u and v. Such an edge e is called incident with the vertices u and v and e is said to connect u and v.

Definition: The set of all neighbors of a vertex v of $G = (V, E)$, denoted by $N(v)$, is called the neighborhood of v.

If A is a subset of V, we denote by $N(A)$ the set of all vertices in G that are adjacent to at least one vertex in A.

Definition: The degree of a vertex in an undirected graph is the number of edges incident with it, except that a loop at a vertex contributes two to the degree of that vertex. The degree of the vertex v is denoted by $deg(v)$.

Theorem (Handshaking Theorem): If $G = (V, E)$ is an undirected graph with m edges, then

$$2m = \sum_{v \in V} deg(v)$$

(Note that this applies even if multiple edges and loops are present.)

Theorem: An undirected graph has an even number of vertices of odd degree.

Proof: Let V_1 be the vertices of even degree and V_2 be the vertices of odd degree.

$$2m = \sum_{v \in V} deg(v) = \sum_{v \in V_1} deg(v) + \sum_{v \in V_2} deg(v)$$

Directed Graphs

Definition: An directed graph $G = (V, E)$ consists of V, a nonempty set of vertices, and E, a set of directed edges.

Each edge is an ordered pair of vertices. The directed edge (u, v) is said to start at u and end at v.

Directed Graphs

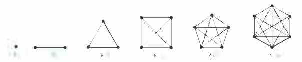
Definition: The in-degree of a vertex v, denoted by $deg^-(v)$ is the number of edges which terminate at v. The out-degree of v, denoted by $deg^+(v)$, is the number of edges with v as their initial vertex.

Theorem: Let $G=(V,E)$ be a graph with directed edges. Then ,

$$|E|=\sum_{v\in V} deg^-(v)=\sum_{v\in V} deg^+(v)$$

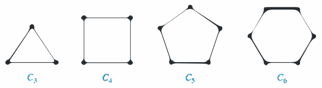
Complete Graphs

A complete graph on n vertices, denoted by K_n, is the simple graph that contains exactly one edge between each pair of distinct vertices .



Cycles

A cycle C_n for $n\ge 3$ consists of n vertices $v_1, v_2, \dots, v_n$, and edges $\{v_1,v_2\}$ $\{v_2, v_3\}$ $\{v_{n-1},v_n\}$ $\{v_n, v_1\}$

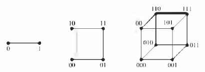


Wheels

A wheel W_n is obtained by adding an additional vertex to a cycle C_n

N-dimensional Hypercube

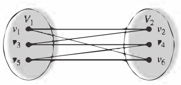
An n-dimensional hypercube or n-cube, Q_n is a graph with 2^n vertices representing all bit strings of length n, where there is an edge between two vertices that differ in exactly one bit position



Bipartite Graphs

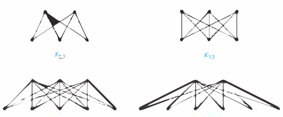
Definition A simple graph G is bipartite if V can be partitioned into two disjoint subsets V_1 and V_2 such that every edge connects a vertex in V_1 and a vertex in V_2

An equivalent definition of a bipartite graph is a graph where it is possible to color the vertices red or blue so that no two adjacent vertices are of the same color



Complete Bipartite Graphs

Definition: A complete bipartite graph $K_{m,n}$ is a graph that has its vertex set partitioned into two subsets V_1 of size m and V_2 of size n such that there is an edge from every vertex in V_1 to every vertex in V_2 .



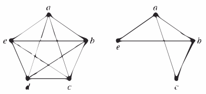
Bipartite Graphs and Matchings

Theorem (Hall's Marriage Theorem) The bipartite graph $G=(V,E)$ with bipartition (V_1,V_2) has a complete matching from V_1 to V_2 if and only if $|N(A)|\ge|A|$ for all subsets A of V_1

Subgraphs

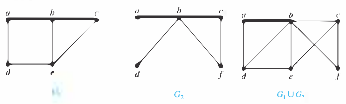
Definition: A subgraph of a graph $G=(V,E)$ is a graph (W,F) where $W\subseteq V$ and $F\subseteq E$.

A subgraph H of G is a proper subgraph of G if $H\ne G$



Union of Graphs

Definition: The union of two simple graphs $G_1=(V_1,E_1)$ and $G_2=(V_2,E_2)$ is the simple graph with vertex set $V_1\cup V_2$ and edge set $E_1\cup E_2$ denoted by $G_1\cup G_2$



Lecture 18

一些定义和记号

- 如果两个顶点之间存在边，那么这两个顶点是adjacent的或者说是neighbors
- $N(v)$: 如果v是$G=(V,E)$中的一个顶点，那么$N(v)$是与v相邻的顶点的集合。
- $N(A)$: 如果A是$G=(V,E)$的一个子集，那么$N(A)$是与A中的顶点相邻的顶点的集合。
- $deg(v)$: 无向图的degree是指与v相邻的顶点的个数，但是一个环对degree的贡献是2。

Theorem (Handshaking Theorem)

If $G=(V,E)$ is an undirected graph with m edges, then

$$2m=\sum_{v\in V} deg(v)$$

如果一个无向图有m条边，那么这个图中所有顶点的度数之和为$2m$（即使是有多重边或环的图）。

Theorem

An undirected graph has an even number of vertices of odd degree.

一个无向图中，度数为奇数的顶点的个数为偶数。

证明：

假设V_{odd}是所有度数为奇数的顶点的集合，V_{even}是所有度数为偶数的顶点的集合，那么

$$2m=\sum_{v\in V} deg(v)=\sum_{v\in V_{odd}} deg(v)+\sum_{v\in V_{even}} deg(v)$$

由于$2m$是偶数，$\sum_{v\in V_{even}} deg(v)$也是偶数，所以$\sum_{v\in V_{odd}} deg(v)$必须也是偶数，而$\sum_{v\in V_{odd}} deg(v)$是所有度数为奇数的顶点的度数之和，所以度数为奇数的顶点的个数为偶数。

Directed Graph

一些定义和记号

- 每一条边都是一个有序对(u,v)，这条边的方向是从u指向v
- 假设(u,v)是$G=(V,E)$中的一条边，那么u是initial vertex并且adjacent to v，v是terminal vertex并且adjacent from u
- $deg^-(v)$: in-degree of v，指向v的边的条数
- $deg^+(v)$: out-degree of v，从v出发的边的条数
- 环对in-degree和out-degree的贡献都是1

Theorem

Let $G=(V,E)$ be a graph with directed edges. Then,

$$|E|=\sum_{v\in V} deg^-(v)=\sum_{v\in V} deg^+(v)$$

有向图的边的条数等于所有顶点的in-degree之和，也等于所有顶点的out-degree之和。

Bipartite Graphs and Matchings

Matching是指把一个集合中的元素和另一个集合中的元素匹配起来。一个matching是边集的一个子集，使得任意两条边都不与同一个顶点关联。换句话说，一个matching是边集的一个子集，使得如果$\{s,t\}$和$\{u,v\}$是matching的两条边，那么s,t,u,v都是不同的。

Job assignments：顶点代表工作和员工，边连接员工和他们被训练过的工作。一个常见的目标是把工作分配给员工，使得完成的工作最多。

A maximum matching is a matching with the largest number of edges.

一个maximum matching是一个matching，它的边数最多。

A matching M in a bipartite graph $G=(V,E)$ with bipartition (V_1,V_2) is a complete matching from V_1 to V_2 if every vertex in V_1 is the endpoint of an edge in the matching, or equivalently, if $|M|=|V_1|$.

一个matching M是一个complete matching，如果M是从V_1到V_2的matching，并且V_1中的每个顶点都是M中一条边的端点，或者等价地，如果$|M|=|V_1|$。

Theorem (Hall's Marriage Theorem): The bipartite graph $G=(V,E)$ with bipartition (V_1,V_2) has a complete matching from V_1 to V_2 if and only if $|N(A)|\ge|A|$ for all subsets A of V_1.

Hall's Marriage Theorem：如果一个二分图$G=(V,E)$，它的顶点集被划分为两个子集V_1和V_2，那么G有一个从V_1到V_2的complete matching，当且仅当对于V_1的任意子集A，$|N(A)|\ge|A|$。

图的表示

- adjacency list (邻接表)
- adjacency matrix (邻接矩阵)
- incidence matrix (关联矩阵)

Adjacency List (邻接表)

定义：adjacency list (邻接表)可以用来表示一个没有重复边的图，它指定了每个顶点的邻接顶点。

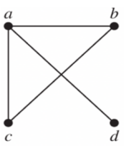
Adjacency Matrix (邻接矩阵)

简单图的邻接矩阵

定义：假设$G=(V,E)$是一个简单图，$|V|=n$。任意地把G的顶点列出来，$v_1,v_2,\cdots,v_n$。G的adjacency matrix A_G是一个$n\times n$的0-1矩阵，当v_i和v_j是adjacent的时候，A_G的(i,j)位置是1，当v_i和v_j不是adjacent的时候，A_G的(i,j)位置是0。

$$A_G=[a_{ij}], \text{where}$$

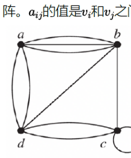
$$a_{ij}=\begin{cases}1 & \text{if } v_i \text{ and } v_j \text{ are adjacent}\\ 0 & \text{if } v_i \text{ and } v_j \text{ are not adjacent}\end{cases}$$



$$\begin{bmatrix}0&1&1&1\\1&0&1&0\\1&1&0&0\\1&0&0&0\end{bmatrix}$$

有环或多重边的图的邻接矩阵

邻接矩阵也可以用来表示有环或多重边的图。邻接矩阵不再是0-1矩阵。a_{ij}的值是v_i和v_j之间的边的条数。

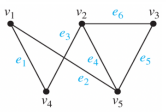


$$\begin{bmatrix}0&3&0&2\\3&0&1&1\\0&1&1&2\\2&1&2&0\end{bmatrix}$$

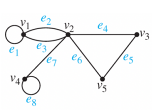
Incidence Matrix (关联矩阵)

定义：假设$G=(V,E)$是一个无向图，顶点$v_1,v_2,\cdots,v_n$，边$e_1,e_2,\cdots,e_m$。G的incidence matrix M_G是一个$n\times m$的0-1矩阵$M=[m_{ij}]$，

$$m_{ij}=\begin{cases}1 & \text{if } e_j \text{ is incident with } v_i\\ 0 & \text{otherwise}\end{cases}$$



$$\begin{bmatrix}1&1&0&0&0&0\\0&0&1&1&0&1\\0&0&0&0&1&1\\1&0&1&0&0&0\\0&1&0&1&1&0\end{bmatrix}$$

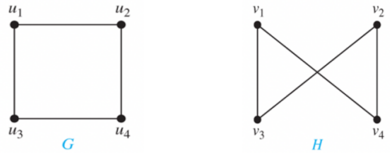


$$\begin{bmatrix}1&1&1&0&0&0&0&0\\0&1&1&1&0&1&1&0\\0&0&0&1&1&0&0&0\\0&0&0&0&0&0&1&1\\0&0&0&0&1&1&0&0\end{bmatrix}$$

Isomorphism of Graphs (同构)

定义：简单图$G_1=(V_1,E_1)$和$G_2=(V_2,E_2)$是isomorphic的，如果存在一个从V_1到V_2的双射，并且满足对于V_1中的任意两个顶点a和b，a和b是adjacent的当且仅当$f(a)$和$f(b)$是adjacent的。这样的函数f被称为isomorphism。

举个例子，下面的两个图是isomorphic的



双射函数可以是$f(u_1)=v_1, f(u_2)=v_4, f(u_3)=v_3, f(u_4)=v_2$。

至此，我们来总结一下几个isomorphic invariants：

- number of vertices (顶点的个数)
- number of edges (边的条数)
- degree sequence (度数序列)
- existence of simple circuits of various lengths (长度为k的simple circuit的存在)

Path: Undirected Graph

Definition: Let n be a nonnegative integer and G an undirected graph. A path of length n from u to v in G is a sequence of n edges $e_1,e_2,\cdots,e_n$ of G for which there exists a sequence $x_0=u,x_1,\cdots,x_{n-1},x_n=v$ of vertices such that e_i has the endpoints x_{i-1} and x_i for $i=1,\cdots,n$.

定义：n是一个非负整数，G是一个无向图。G中从u到v的长度为n的path是一个边的序列$e_1,e_2,\cdots,e_n$，满足存在一个顶点的序列$x_0=u,x_1,\cdots,x_{n-1},x_n=v$，使得$e_i$的端点是$x_{i-1}$和$x_i$，$i=1,\cdots,n$。

一个path被称作circuit或cycle，如果它的起点和终点是同一个顶点，但长度大于0。

一个path或者circuit是simple的，如果它不包含重复的edge。

path的长度=path里面的边的条数

Path: Directed Graph

Definition: Let n be a nonnegative integer and G an directed graph. A path of length n from u to v in G is a sequence of n edges $e_1,e_2,\cdots,e_n$ of G for which there exists a sequence $x_0=u,x_1,\cdots,x_{n-1},x_n=v$ of vertices such that e_i is associated with initial vertex x_{i-1} and terminal vertex x_i for $i=1,\cdots,n$.

定义：n是一个非负整数，G是一个有向图。G中从u到v的长度为n的path是一个边的序列$e_1,e_2,\cdots,e_n$，满足存在一个顶点的序列$x_0=u,x_1,\cdots,x_{n-1},x_n=v$，使得$e_i$的initial vertex是$x_{i-1}$，terminal vertex是$x_i$，$i=1,\cdots,n$。

cycle和simple path的定义和无向图中的一样。

Connectivity

一个无向图是connected的，如果图中任意两个不同的顶点之间都存在一条路径。

一个不是connected的无向图是disconnected的。

Lemma: 如果图G中两个不同的顶点x和y之间存在一条路径，那么G中x和y之间存在一条simple path。

Proof:

删除里面的circuit即可。

Theorem: 一个connected的无向图中任意两个不同的顶点之间都存在一条simple path。

一个图G的connected component是一个connected的子图，它不是G的另一个connected的子图的proper subgraph。

也就是说，connected component要满足两个条件：

- 连通性：它是connected的
- 最大性：这个子图是最大的连通子图，意味着它不是另一个更大的连通子图的真子图（即，不是另一个连通子图的一部分）

Connectedness in Directed Graphs

定义：

- 一个有向图是strongly connected的，如果对于图中的任意两个顶点a和b，a到b有一条path，b到a也有一条path。
- 一个有向图是weakly connected的，如果它的underlying undirected graph是connected的。

Cut Vertices and Cut Edges

Sometimes the removal from a graph of a vertex and all incident edges disconnect the graph.

Such vertices are called cut vertices. Similarly we may define cut edges.

有时候，从一个图中删除一个顶点和所有与它关联的边会使得图不再是connected的。这样的顶点被称为cut vertices。类似地，我们也可以定义cut edges。

一个图G的edge connectivity $\lambda(G)$是一个edge cut中边的最小条数。就是最少需要删除多少条边，才能使得图不再是connected的。这个数值就是edge connectivity。

Counting Paths between Vertices

Theorem: 假设G是一个图，A是G的adjacency matrix，顶点的顺序是$v_1,v_2,\cdots,v_n$。从v_i到v_j的长度为r的不同的path的个数，其中r是一个正整数，等于A^r的(i,j)位置的值。

Euler Paths and Circuits

引入：Ko¨nigsberg seven-bridge problem: 有人想知道是否可以从城镇的某个位置出发，穿过所有的桥一次而不重复，然后回到起点。

定义: 一个图G中的Euler circuit是一个包含G中所有边的simple circuit。一个图G中的Euler path是一个包含G中所有边的simple path。

Necessary Conditions for Euler Circuits and Paths

- Euler Circuit: 每个顶点的度数都是偶数
- Euler Path: 除了两个顶点的度数是奇数，其他顶点的度数都是偶数。这条path的起点和终点是这两个度数为奇数的顶点。

定义：一个rooted tree中顶点v的level是指从root到v的唯一的path的长度。

一个rooted tree的height是指它的顶点的level的最大值。

Balanced m-ary Trees: 一个高度为h的rooted m-ary tree是balanced的，如果所有的leaves都在level h或$h-1$。

Theorem: 一个高度为h的rooted m-ary tree最多有m^h个leaves。

Corollary: 如果一个高度为h的rooted m-ary tree有l个leaves，那么$h\ge\lceil\log_m l\rceil$。如果这个m-ary tree是full和balanced的，那么$h=\lceil\log_m l\rceil$。

Hamilton Paths and Circuits

Euler path是每个边都只经过一次

Hamilton path是每个顶点都只经过一次

Necessary Conditions for Hamilton Circuits and Paths

没有已知的充要条件可以判断存在Hamilton circuit或path。

但是有一些sufficient conditions：

- Dirac's Theorem: 如果G是一个简单图，$|V|\ge 3$，并且G中每个顶点的度数都大于等于$\frac{|V|}{2}$，那么G有一个Hamilton circuit。
- Ore's Theorem: 如果G是一个简单图，$|V|\ge 3$，并且对于G中的任意两个不相邻的顶点u和v，u和v的度数之和都大于等于$|V|$，那么G有一个Hamilton circuit。
  例子：证明K_n有Hamilton circuit

Hamilton path问题是NP-completez的

Planar Graphs

Definition: A graph is called planar if it can be drawn in the plane without any edges crossing. Such a drawing is called a planar representation of the graph.

定义：如果一个图可以在平面上画出来，而且没有边相交，那么这个图是planar的。这样的画法被称为这个图的planar representation。

关于怎么找planar representation，可以试试先画一个闭合的多边形，然后再把剩下的顶点一个一个加进去

Euler's Formula

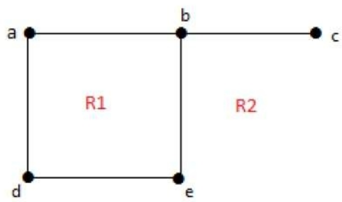
一个图的planar representation把平面分成了一些区域，包括一个无界区域。

Theorem (Euler's Formula): 假设G是一个connected的planar simple graph，e是边的条数，v是顶点的个数，r是G的planar representation中的区域的个数。那么，$r=e-v+2$。

The Degree of Regions

定义：一个region的degree是指这个region的边的条数。当一个边在边界上出现两次的时候，它对degree的贡献是2。

例2：上面那个比较正常，我们看个抽象的：



degree of region 1 = 4

degree of region 2 = 6 (这里bc被计算了两次)

Corollary 1

如果G是一个connected的planar simple graph，e是边的条数，v是顶点的个数，$v\ge 3$，那么$e\le 3v-6$。

Corollary 2

如果G是一个connected的planar simple graph，那么G有一个顶点的度数不超过5。

Corollary 3

如果G是一个connected的planar simple graph，e是边的条数，v是顶点的个数，$v\ge 3$，并且G中没有长度为3的circuit，那么$e\le 2v-4$。

Four-color theorem

Four-color theorem: 给定一个平面，把它分成一些连续的区域，产生一个叫做map的图形，最多只需要四种颜色来给map中的区域染色，使得任意两个相邻的区域的颜色不同。

Chromatic number

色数是指给图中的顶点染色，使得相邻的顶点颜色不同，所需的最少颜色数。

根据Four-color theorem，平面图的色数不超过4。

记号：$\chi(G)$表示图G的色数。

K_n的色数是n。

K_{mn}的色数是2。

C_n的色数是2（当n是偶数的时候）或者3（当n是奇数的时候）。

Trees

定义：一个tree是一个connected的无向图，它没有simple circuit。

Theorem: 一个无向图是一个tree，当且仅当它的任意两个顶点之间都有唯一的simple path。

Rooted Trees

定义：一个rooted tree是一个tree，其中一个顶点被指定为root，每条边都是从root指向其他顶点的。

我们可以通过选择任意一个顶点作为root，把一个unrooted tree变成一个rooted tree。

在一个rooted tree中，边的方向可以省略，因为root的选择决定了边的方向。

定义：顶点v的parent是指唯一的顶点u，使得u到v有一条有向边。

当u是v的parent的时候，v被称为u的child。

有相同parent的顶点被称为siblings。

顶点v的ancestors是指从root到v的路径上的顶点，不包括v，包括root。

顶点v的descendants是指v的祖先。

一个rooted tree的顶点被称为leaf，如果它没有child。

有child的顶点被称为internal vertices。

以a为root的subtree包括a和a的descendants，以及所有与这些descendants关联的边。

定义：如果一个rooted tree的每个internal vertex都有不超过m个children，那么这个rooted tree被称为m-ary tree。如果每个internal vertex都有m个children，那么这个rooted tree被称为full m-ary tree。特别地，如果$m=2$，那么这个rooted tree被称为binary tree。

Theorem: 一个有n个顶点的tree有$n-1$条边。

Theorem: 一个有i个internal vertices的full m-ary tree有$n=mi+1$个顶点。

Theorem: 一个有n个顶点的full m-ary tree有$i=\frac{n-1}{m}$个internal vertices和$l=\frac{(m-1)n+1}{m}$个leaves。

Theorem: 一个有i个internal vertices的full m-ary tree有$n=mi+1$个顶点和$l=(m-1)i+1$个leaves。

Theorem: 一个有l个leaves的full m-ary tree有$n=\frac{ml-1}{m-1}$个顶点和$i=\frac{l-1}{m-1}$个internal vertices。