Graph theory: trees

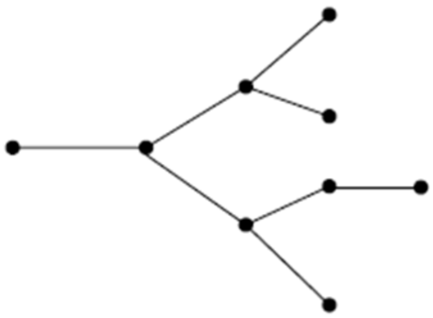
Yulia Burkatovskaya
Department of Information Technologies
Associate professor

5. Trees

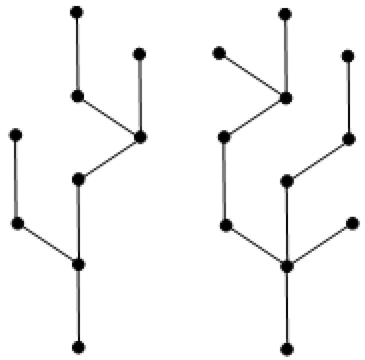
- Trees
- Minimum spanning tree
- Fundamental circuits and fundamental cut sets
- Rooted trees
- Maximum forest
- Search trees

5.1. Trees

- A *connected acyclic graph* is called a **tree**. In other words, a connected graph with no cycles is called a tree.



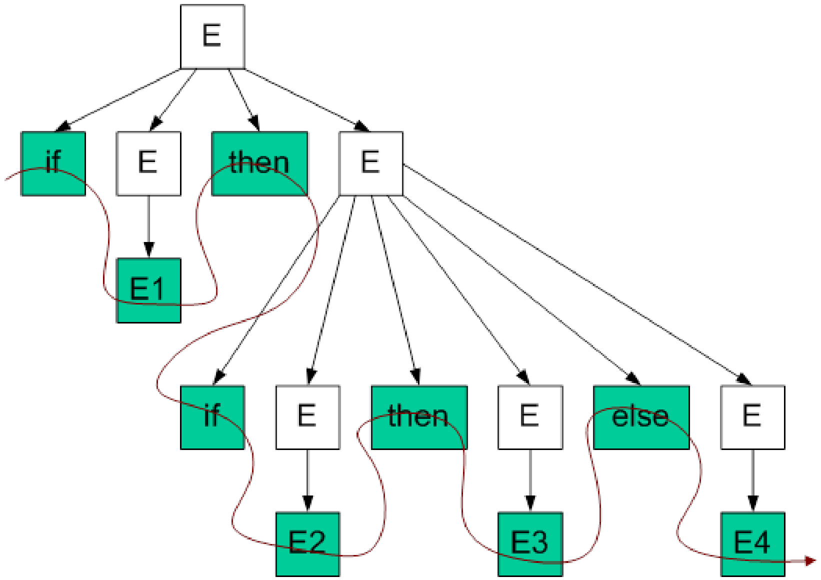
- An *acyclic graph* is called a **forest**.

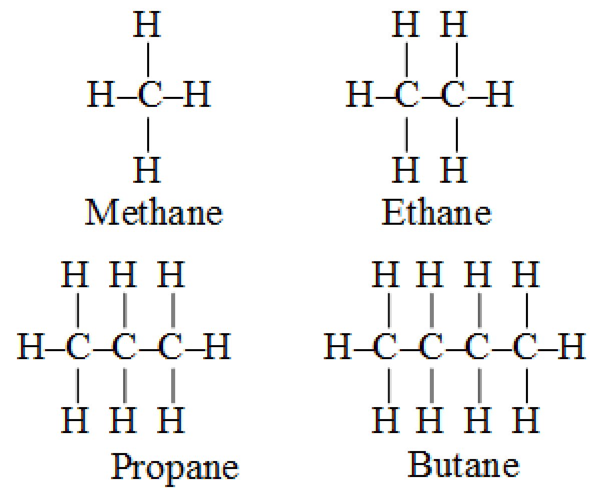


Application of trees

- **Chemistry** (saturated hydrocarbons)
- **Compilers** (parsing)

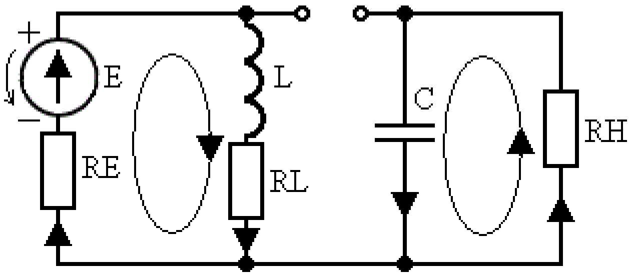
if E1 then if E2 then E3 else E4

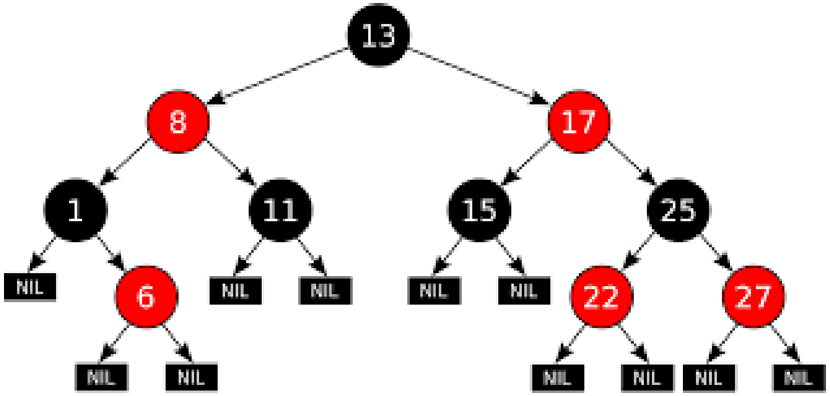


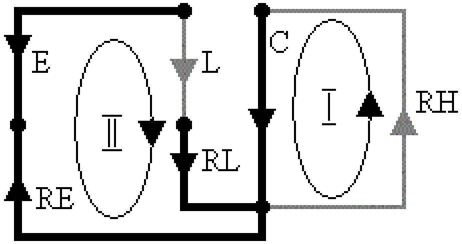


Application of trees

- **Physics** (electrical circuits)
- **Programming** (search trees)







Six different characterizations of a tree

- (1) T is a tree.
- (2) Any two vertices of *T* are connected by exactly one path.
- (3) T is connected, and every edge is a cut-edge.
- (4) T is connected and has $n - 1$ edges.
- (5) T contains no cycles and has $n - 1$ edges.
- (6) *T* contains no cycles, and for any new edge *e*, the graph $T+e$ has exactly one cycle.

Six different characterizations of a tree

(1)→(2)

- (1) *T* is a tree.
- (2) Any two vertices of *T* are connected by exactly one path.

- *T* is connected → any two vertices of *T* are connected.
- If there are two paths <*x*,*y*> then there is a cycle.

Six different characterizations of a tree

(2)→(3)

- (2) Any two vertices of *T* are connected by exactly one path.
- (3) *T* is connected, and every edge is a cut-edge.

- Any two vertices of *T* are connected → *T* is connected.
- If $e=(x,y)$ is not a cut-edge then $G\backslash(x,y)$ is connected; hence, there is a path $<x,y>$ which does not include *e*. So, there are two different paths $<x,y>$ in *G*.

Six different characterizations of a tree

(3)→(4)

- (3) T is connected, and every edge is a cut-edge.
- (4) T is connected and has $m=n-1$ edges.

Mathematical induction method.

- **Base:** for $n=1$ one has $m=0$.
- **Inductive step:** Let for graphs with $1,\dots,n-1$ vertices the statement is true. Consider graph T with n vertices. Delete any edge e from T. As e is a cut-edge, $T \backslash e = T_1 \cup T_2$. They are connected, every edge is a cut-edge and the numbers of vertices is less than n; so, $m_1=n_1-1$ and $m_2=n_2-1$. The number of edges in T is the sum of the numbers of edges in T_1 and T_2 plus the deleted edge; consequently

$$m = m_1+m_2 -1 = n_1 -1+n_2 -1+1 = n - 1.$$

Six different characterizations of a tree

(4)→(5)

- (4) T is connected and has $n - 1$ edges.
- (5) T contains no cycles and has $n - 1$ edges.

- Let T contain a cycle with s vertices and edges; then, to join other n −s vertices to the cycle one needs at least n −s edges. So, the total number of edges is

$$m \geq s + n - s = n > n - 1.$$

Six different characterizations of a tree

(5)→(6)

- (5) T contains no cycles and has $n - 1$ edges.
- (6) T contains no cycles, and for any new edge e, the graph $T+e$ has exactly one cycle.

- As T is acyclic than it is a forest with k trees. For a tree,

$$m=n-1.$$

The total number of edges in T is

$$m_1+\ldots+m_k = n_1 - 1+\ldots+n_k - 1= n - k.$$

Hence; k=1 and T is a tree. Every two vertices of a tree are joined by the only path; so, adding a new edge produces exactly one cycle.

Six different characterizations of a tree

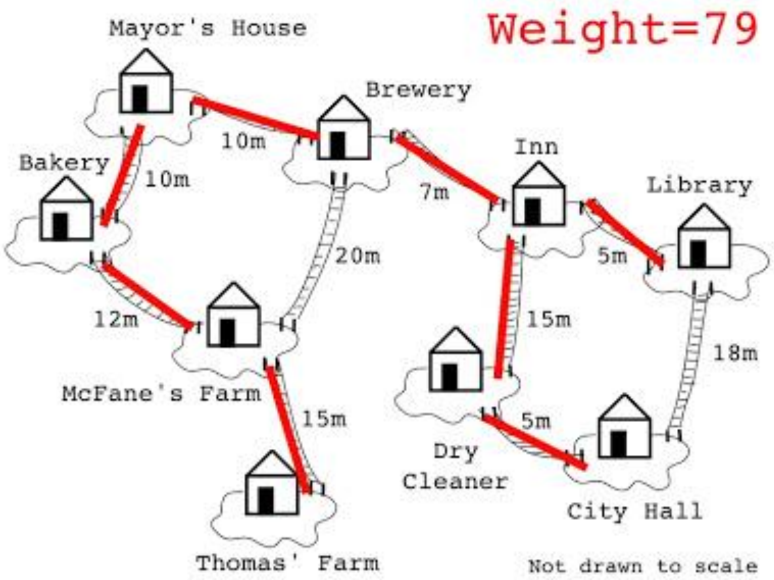
(6)$\rightarrow$(1)

- (6) T contains no cycles, and for any new edge e, the graph $T+e$ has exactly one cycle.
- (1) T is a tree.

- If after adding any new edge a cycle appears then any two vertices are joined by a path; so, T is connected. This together with absence of cycles gives a tree.

5.2. Minimum spanning tree

- How to join all houses and to minimize the length of the communications?
- In a weighted graph, the **minimum spanning tree** is the set of edges with the minimum total weight such that they connect all of the nodes.
- **Applications of MST problem**: https://www.geeksforgeeks.org/?p=11110 .



http://computationaltales.blogspot.ru/2011/08/minimum-spanning-trees-prims-algorithm.html

Greedy algorithms

- A **greedy algorithm** is an algorithmic paradigm that follows the problem solving heuristic of making the locally optimal choice at each stage with the hope of finding a global optimum.
- In many problems, a greedy strategy does not in general produce an optimal solution.
- But for the **minimum spanning tree** problem, greedy algorithms produce a **global optimum**.

Kruskal's algorithm

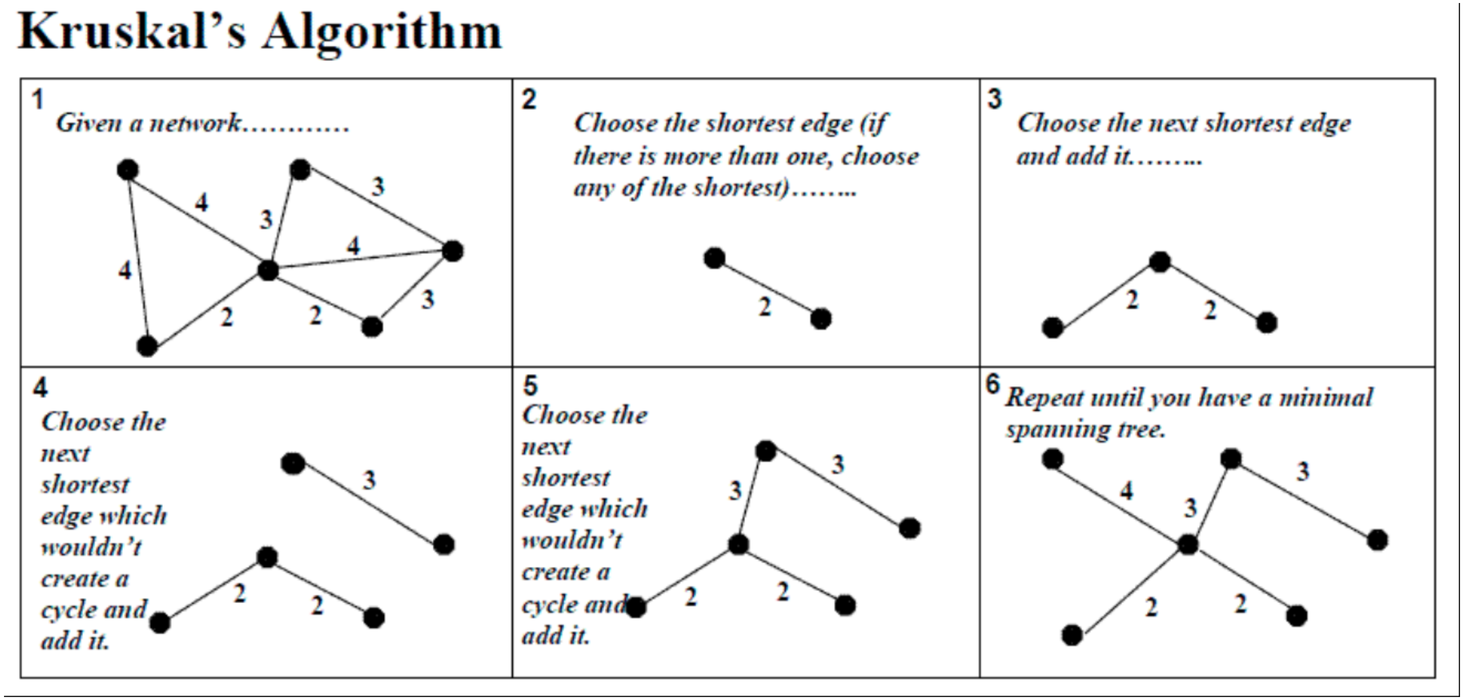
- Sort all the edges from low weight to high
- Take the edge with the lowest weight and add it to the spanning tree. If adding the edge created a cycle, then reject this edge.
- Keep adding edges until we have $p-1$ edges.

- https://www.programiz.com/dsa/kruskal-algorithm
- https://youtu.be/71UQH7Pr9kU

Kruskal's algorithm

Example.

- http://nadide.github.io/assets/img/algo-image/MST/kruskal.png



Prim's algorithm

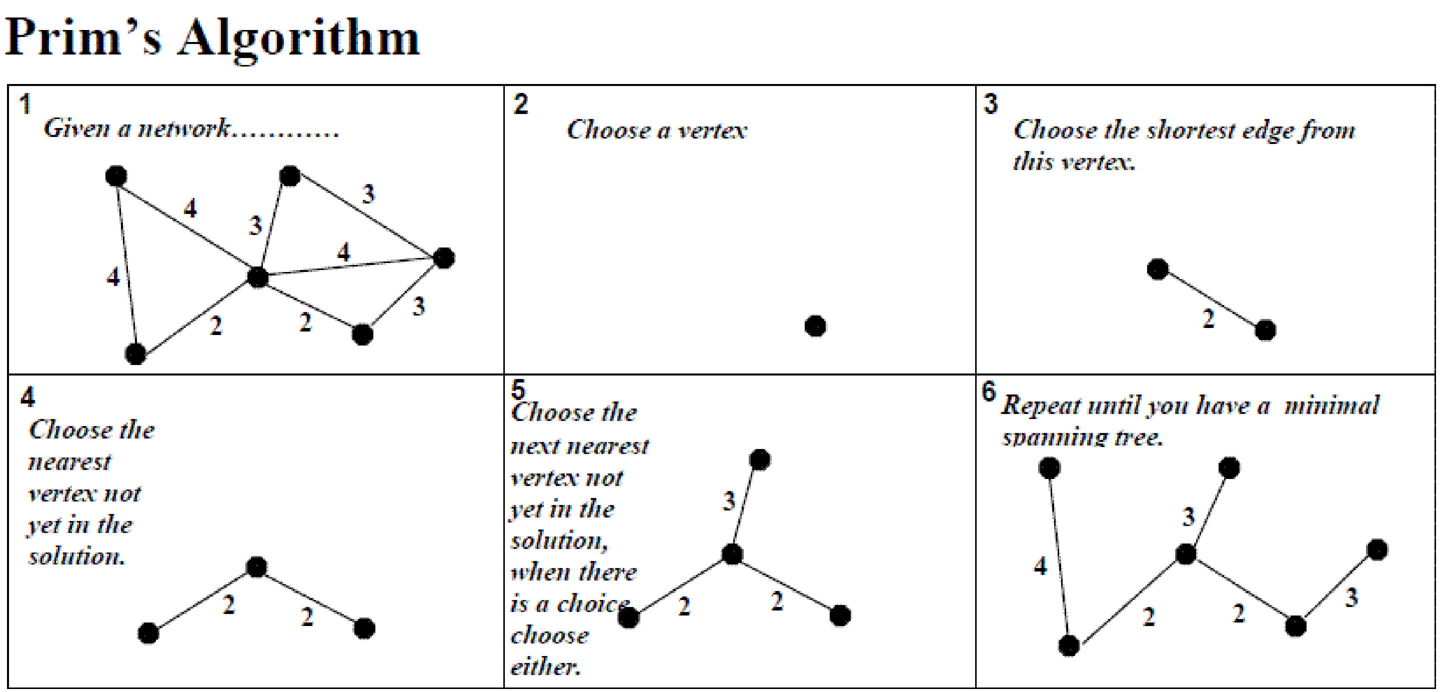
- Initialize the minimum spanning tree with a vertex chosen at random.
- Find all the edges that connect the tree to new vertices, find the minimum and add it to the tree
- Keep adding edges until we have $p-1$ edges.

- https://www.programiz.com/dsa/prim-algorithm
- https://youtu.be/cplfcGZmX7I

Prim’s algorithm

Example.

- https://www.thestudentroom.co.uk/attachment.php?attachmentid=23572&stc=1&d=1148396387



5.3. Fundamental circuits and fundamental cut sets

Let $G(V,E)$ be a multigraph with n vertices, m edges and k connected components.

- **Cocyclomatic number** of the graph $G(V,E)$ is $\rho(G)=n-k$.

It is the total number of edges in spanning trees of all connected components of the graph.

- **Cyclomatic number** of the graph $G(V,E)$ is $\nu(G)=m-n+k$.

It indicates ho many edges need to be removed in order to the graph became a forest with k connected components.

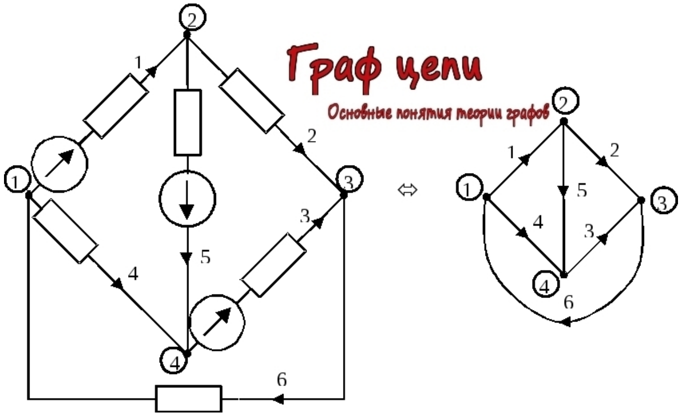
Cyclomatic and cocyclomatic numbers

Example.

Cyclomatic and cocyclomatic numbers

In the **theory of electrical circuits**, the numbers have a definite physical meaning.

- The **cyclomatic number** is equal to the largest number of independent circuits in the electric circuit graph, i.e. the largest number of independent circular currents that can flow in the circuit.
- The **cocyclomatic number** is equal to the number of independent potential differences between the nodes of the circuit.



Fundamental circuits

Any circuit or cycle can be represented by the set of its edges.

- **Modulo 2 addition (XOR):**

$$\mu_1 \oplus \mu_2 = \{e : e \in \mu_1, e \notin \mu_2\} \cup \{e : e \in \mu_2, e \notin \mu_1\}$$

- **Conjunction (AND):**

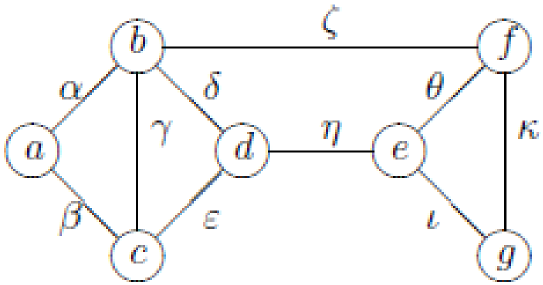
$$0\mu = \emptyset, \quad 1\mu = \mu.$$

- **Linear combination:**

$$\mu = \bigoplus_{i=1}^{n} a_i \mu_i, \quad a_i \in \{0, 1\}.$$

Fundamental circuits

Example.

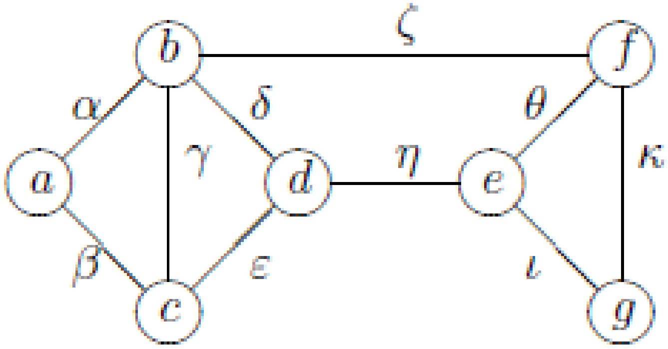


$\mu_1 = \{\delta, \zeta, \eta, \theta\};$
$\mu_2 = \{\gamma, \delta, \varepsilon\};$
$\mu_3 = \{\alpha, \beta, \gamma\};$
$1\mu_1 \oplus 1\mu_2 \oplus 0\mu_3 = \{\delta, \zeta, \eta, \theta\} \oplus \{\gamma, \delta, \varepsilon\} \oplus \emptyset = \{\gamma, \varepsilon, \zeta, \eta, \theta\}.$

Fundamental circuits

A set of circuits is **independent** if any circuit is not a linear combination of others; otherwise, the set is **dependent**.

Example. Set $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ is dependent as $\mu_4 = \mu_2 + \mu_3$; set $\{\mu_1, \mu_2, \mu_4\}$ is independent

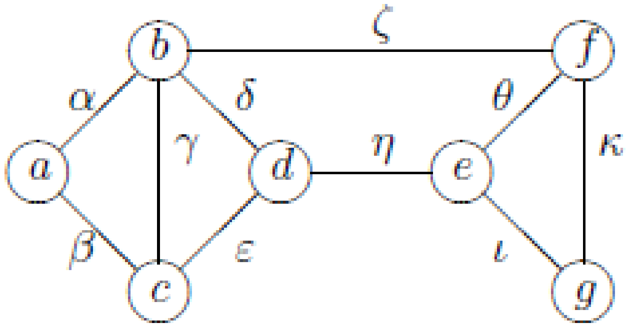


$$\mu_1 = \{\delta, \zeta, \eta, \theta\};$$
$$\mu_2 = \{\gamma, \delta, \varepsilon\};$$
$$\mu_3 = \{\alpha, \beta, \gamma\};$$
$$\mu_4 = \{\alpha, \beta, \delta, \varepsilon\}.$$

Fundamental circuits

An independent set of circuits is a **system of fundamental circuits** if in contains the greatest possible number of circuits; the circuits of this set are **fundamental**.

Example. Set $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ is independent; any circuits is a linear combination of the circuits from the set.



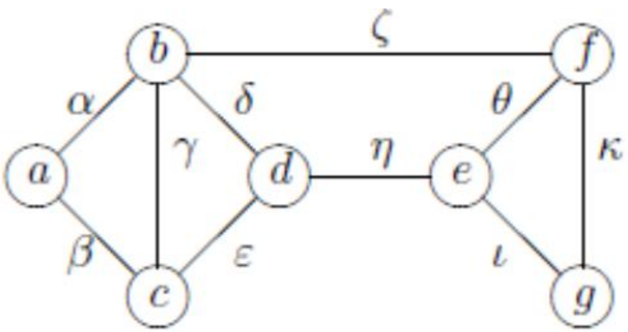
$$\mu_1 = \{\delta, \zeta, \eta, \theta\};$$
$$\mu_2 = \{\gamma, \delta, \varepsilon\};$$
$$\mu_3 = \{\alpha, \beta, \gamma\};$$
$$\mu_4 = \{\theta, \iota, \kappa\}.$$

$$\{\alpha, \beta, \delta, \varepsilon\} = \mu_2 \oplus \mu_3;$$
$$\{\delta, \zeta, \eta, \theta, \iota, \kappa\} = \mu_1 \oplus \mu_4;$$
$$\{\alpha, \beta, \varepsilon, \zeta, \eta, \theta\} = \mu_1 \oplus \mu_2 \oplus \mu_3.$$

Fundamental circuits theorem

Theorem. For a simple connected graph, the number of fundamental circuits is equal to $\nu(G)=m-n+1$.

Example. Here n=7, m=10; there are 4=10−7+1 independent circuits.



$$\mu_1 = \{\delta, \zeta, \eta, \theta\};$$
$$\mu_2 = \{\gamma, \delta, \varepsilon\};$$
$$\mu_3 = \{\alpha, \beta, \gamma\};$$
$$\mu_4 = \{\theta, \iota, \kappa\}.$$

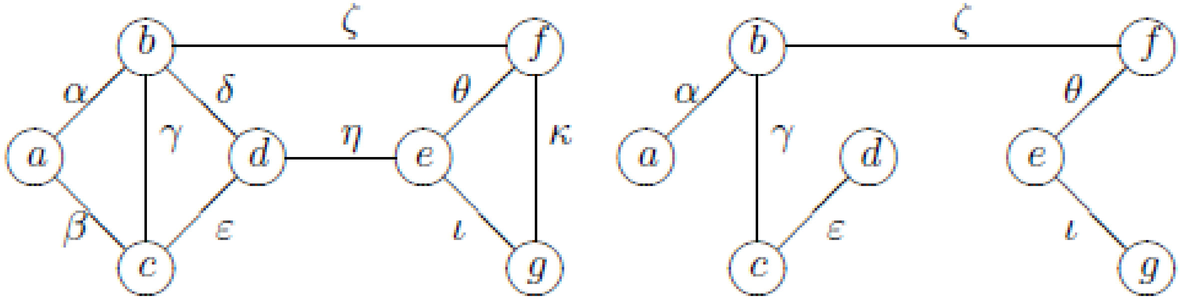
Fundamental cycles construction

Algorithm

- *Start.* There is graph $G(V,E)$.
- *Step 1*. Construct any spanning tree $T(V,E')$. Set $j=0$.
- *Step 2*. If $j= m-n+1$ then go to End; else set $j=j+1$.
- *Step 3.* Choose the next edge $e_j=(v_j,u_j)$ not included into the spanning tree.
- *Step 4*. Find the path $< v_j,u_j >$ in the spanning tree; together with the edge (v_j,u_j), it gives cycle Z_j. Go to Step 2.
- *End*. $\{Z_j\}$ is a system of fundamental cycles.

Fundamental cycles construction

Example.

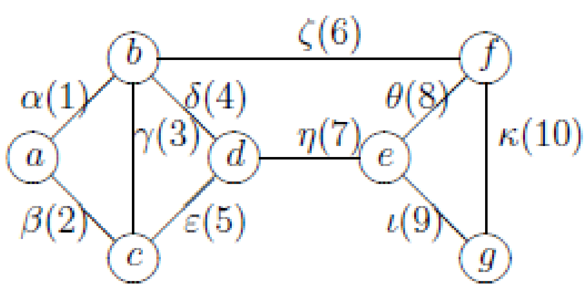


$$
\begin{aligned}
e_1 &= \beta = (a,c), & <a,c> &= abc, & Z_1 &= \{\beta, \alpha, \gamma\};\\
e_2 &= \delta = (b,d), & <b,d> &= bcd, & Z_2 &= \{\delta, \gamma, \varepsilon\};\\
e_3 &= \eta = (d,e), & <d,e> &= dcbfe, & Z_3 &= \{\eta, \varepsilon, \gamma, \zeta, \theta\};\\
e_4 &= \kappa = (f,g), & <f,g> &= feg, & Z_4 &= \{\kappa, \theta, \iota\}.
\end{aligned}
$$

Matrix of fundamental circuits

- Rows correspond to fundamental circuits, columns correspond to edges; an element is equal to 1 iff the edge belongs to the circuit.

Example.



$Z_1 = \{\beta, \alpha, \gamma\};$
$Z_2 = \{\delta, \gamma, \varepsilon\};$
$Z_3 = \{\eta, \varepsilon, \gamma, \zeta, \theta\};$
$Z_4 = \{\kappa, \theta, \iota\}.$

$\Phi(G) =$

	α	β	γ	δ	ε	ζ	η	θ	ι	κ
Z_1	1	1	1	0	0	0	0	0	0	0
Z_2	0	0	1	1	1	0	0	0	0	0
Z_3	0	0	1	0	1	1	1	1	0	0
Z_4	0	0	0	0	0	0	0	1	1	1

Matrix of fundamental circuits

- **Module 2 product** of matrices $A{:}n\times k$ and $B{:}k\times m$ is matrix $C{:}n\times m$ calculated as follows

$$C_{ij} = \bigoplus_{l=1}^{k} A_{ik} B_{kj}.$$

- **Theorem.** If $I(G)$ is the incidence matrix of graph $G(V,E)$, $\Phi(G)$ is its matrix of fundamental circuits, then

$$I \oplus \Phi^T = 0$$

Fundamental cuts

Consider graph G(V,E) and two subsets

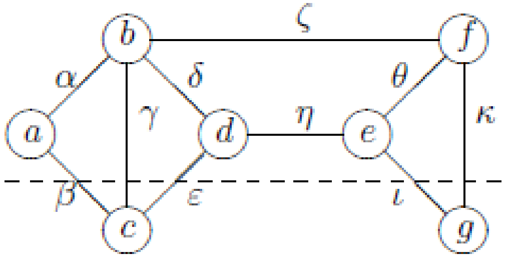
$$V = V_1 \cup V_2,\ V_1 \cap V_2 = \emptyset.$$

Cut (**cocycle**) $P(V_1,V_2)$ is the set of edges joining vertices from V_1 with vertices from V_2, i.e.

$$P(V_1, V_2) = \{(v_1, v_2) \in E : v_1 \in V_1, v_2 \in V_2\}.$$

A cut is **proper** if after removal of any its subset the graph is connected.

Example. Non-proper cut P is union of proper cuts P_1 and P_2.



$$P(V_1, V_2) = \{\beta, \gamma, \varepsilon, \iota, \kappa\}$$

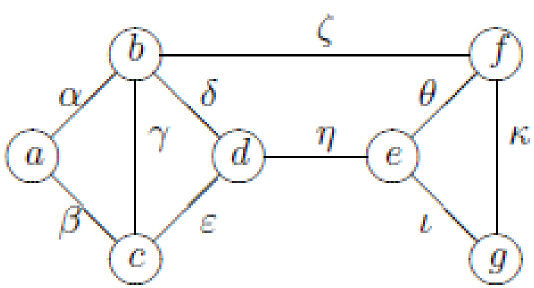
$$P_1 = \{\beta, \gamma, \varepsilon\} \text{ и } P_2 = \{\iota, \kappa\}.$$

Fundamental cuts

Lemma. Any non-proper cut is a union of disjoint proper cuts.

A set of cuts is **independent** if any cut is not a linear combination of others; otherwise, the set is **dependent**.

Example. Set $\{\psi_1, \psi_2, \psi_3, \psi_4\}$ is dependent as $\psi_3 = \psi_1 + \psi_2$; set $\{\psi_1, \psi_2, \psi_4\}$ is independent

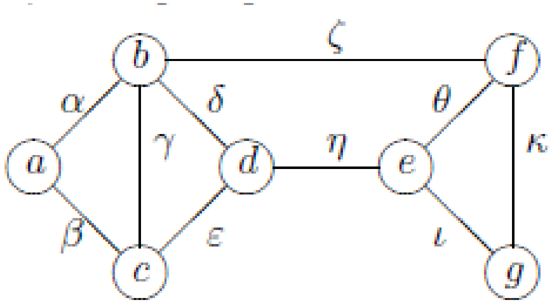


$$\psi_1 = \{\beta, \gamma, \varepsilon, \zeta, \eta\};$$
$$\psi_2 = \{\zeta, \eta, \iota, \kappa\};$$
$$\psi_3 = \{\beta, \gamma, \varepsilon, \iota, \kappa\};$$
$$\psi_4 = \{\zeta, \eta, \theta\}.$$

Fundamental cuts

An independent set of cuts is a **system of fundamental cuts** if in contains the greatest possible number of cuts; the cuts of this set are **fundamental**.

Example. Set is independent; any cut is a linear combination of the cuts from the set.



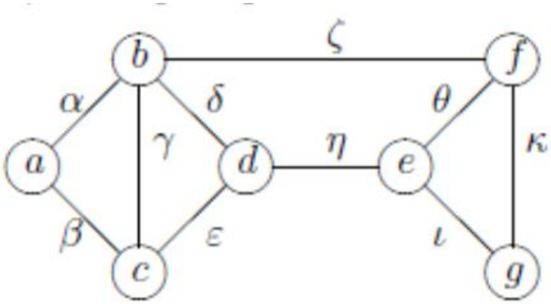
$$\psi_1 = \{\alpha, \beta\};$$
$$\psi_2 = \{\delta, \varepsilon, \zeta\};$$
$$\psi_3 = \{\zeta, \eta\};$$
$$\psi_4 = \{\zeta, \theta, \iota\};$$
$$\psi_5 = \{\iota, \kappa\};$$
$$\psi_6 = \{\beta, \gamma, \varepsilon\}.$$

$$\{\alpha, \gamma, \delta, \theta, \kappa\} = \psi_1 \oplus \psi_2 \oplus \psi_4 \oplus \psi_5 \oplus \psi_6;$$
$$\{\alpha, \gamma, \varepsilon\} = \psi_1 \oplus \psi_6;$$
$$\{\beta, \gamma, \varepsilon, \iota, \kappa\} = \psi_5 \oplus \psi_5.$$

Fundamental cuts theorem

Theorem. For a simple connected graph, the number of fundamental cus is equal to ρ(*G*)=*n*+1.

Example. Here *n*=7; there are 6=7−1 independent cuts.



$\psi_1 = \{\alpha, \beta\};$
$\psi_2 = \{\delta, \varepsilon, \zeta\};$
$\psi_3 = \{\zeta, \eta\};$
$\psi_4 = \{\zeta, \theta, \iota\};$
$\psi_5 = \{\iota, \kappa\};$
$\psi_6 = \{\beta, \gamma, \varepsilon\}.$

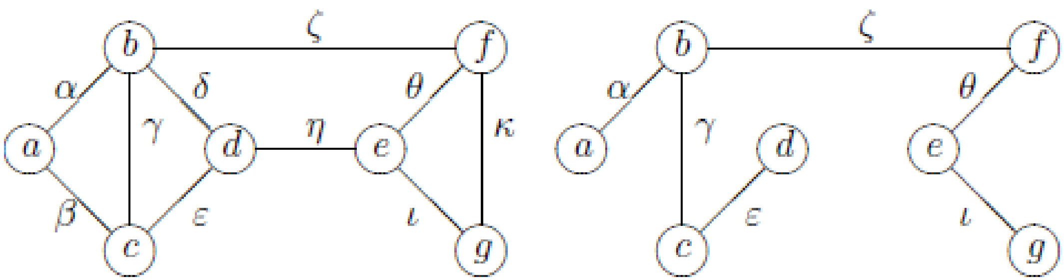
Fundamental cuts construction

Algorithm

- *Start.* There is graph $G(V,E)$.
- *Step 1*. Construct any spanning tree $T(V,E')$. Set j=0.
- *Step 2*. If $j = n-1$ then go to End; else set $j=j+1$.
- *Step 3.* Choose the next edge $e_j=(w_j,u_j)$ included into the spanning tree. Remove it from the tree and obtain a forest from two trees with the sets of vertices W_j and U_j.
- *Step 4*. Find the cut $Y_j=P(W_j, U_j)$. Go to Step 2.
- *End*. $\{Y_j\}$ is a system of fundamental cuts.

Fundamental cuts construction

Example.

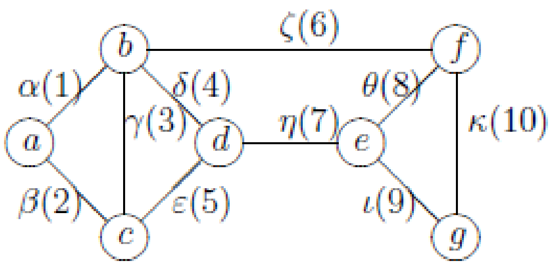


$$
\begin{array}{llll}
e_1 = \alpha = (a,b), & U_1 = \{a\}, & W_1 = \{b,c,d,e,f,g\}, & Y_1 = \{\alpha,\beta\}; \\
e_2 = \gamma = (b,c), & U_2 = \{a,b,f,e,g\}, & W_2 = \{c,d\}, & Y_2 = \{\beta,\gamma,\delta,\eta\}; \\
e_3 = \varepsilon = (c,d), & U_3 = \{a,b,c,e,f,g\}, & W_3 = \{d\}, & Y_3 = \{\delta,\varepsilon,\eta\}; \\
e_4 = \zeta = (b,f), & U_4 = \{a,b,c,d\}, & W_4 = \{e,f,g\}, & Y_4 = \{\zeta,\eta\}; \\
e_5 = \theta = (e,f), & U_5 = \{e,g\}, & W_5 = \{a,b,c,d,f\}, & Y_5 = \{\eta,\theta,\kappa\}; \\
e_6 = \iota = (e,g), & U_6 = \{a,b,c,d,e,f\}, & W_6 = \{g\}, & Y_6 = \{\iota,\kappa\}.
\end{array}
$$

Matrix of fundamental cuts

- Rows correspond to fundamental cuts, columns correspond to edges; an element is equal to 1 iff the edge belongs to the cut.

Example.



$Y_1 = \{\alpha, \beta\};$
$Y_2 = \{\beta, \gamma, \delta, \eta\};$
$Y_3 = \{\delta, \varepsilon, \eta\};$
$Y_4 = \{\zeta, \eta\};$
$Y_5 = \{\eta, \theta, \kappa\};$
$Y_6 = \{\iota, \kappa\}.$

$\Psi(G) =$

	α	β	γ	δ	ε	ζ	η	θ	ι	κ
Y_1	1	1	0	0	0	0	0	0	0	0
Y_2	0	1	1	1	0	0	1	0	0	0
Y_3	0	0	0	1	1	0	1	0	0	0
Y_4	0	0	0	0	0	1	1	0	0	0
Y_5	0	0	0	0	0	0	1	1	0	1
Y_6	0	0	0	0	0	0	0	0	1	1

Matrix of fundamental cuts

- **Theorem.** If *Ψ*(*G*) is the matrix of fundamental cuts of graph *G*(*V*,*E*), *Φ*(*G*) is its matrix of fundamental circuits, then

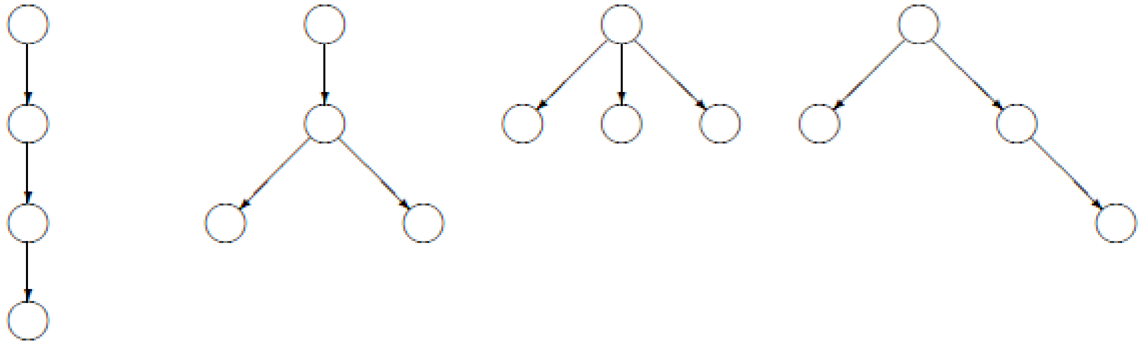
$$\Phi \oplus \Psi^T = 0$$

5.4. Rooted trees

Rooted tree is a digraph with the following properties:

- there is a single node *v* with in-degree equal to 0 (it is called **root**);
- the in-degrees of all other nodes are equal to 1;
- each node is reachable from the root.

Example. All rooted trees with four vertices.



Properties of directed trees

Theorem. Any directed tree has the following properties:

- $m = n - 1$;
- the in-degrees of all other nodes are equal to 1;
- each node is reachable from the root.

Example. All directed trees with four vertices.