# (十) 图论: 树 (Trees)

## 魏恒峰

hfwei@nju.edu.cn

2021年05月13日



# BLU-RAY" + DIGITAL 15TH ANNIVERSARY EDITION ROBIN WILLIAMS MATT DAMON ACADEMY AWARD NOMINATIONS BEST SUPPORTING ACTOR . BEST ORIGINAL SCREENPLAY



你, 真得, 看懂了吗?

Definition (Tree (树))

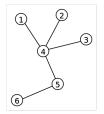
A tree is a connected acyclic undirected graph.

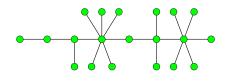
Definition (Tree (树))

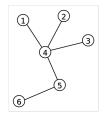
A tree is a connected acyclic undirected graph.

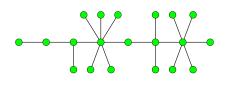
Definition (Forest (森林))

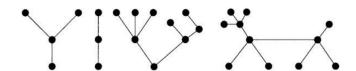
A forest is a acyclic undirected graph.

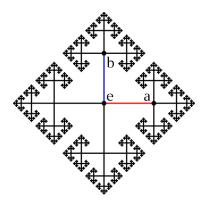












Cayley Graph (4-regular tree)

In a tree T with  $\geq 2$  vertices, for a vertex v in T, if

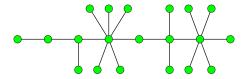
$$\deg(v) = 1$$

then v is called a leaf; otherwise, v is an internal vertex.

In a tree T with  $\geq 2$  vertices, for a vertex v in T, if

$$\deg(v) = 1$$

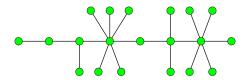
then v is called a leaf; otherwise, v is an internal vertex.



In a tree T with  $\geq 2$  vertices, for a vertex v in T, if

$$\deg(v) = 1$$

then v is called a leaf; otherwise, v is an internal vertex.



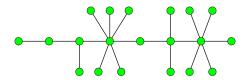
#### Lemma

Any tree T with  $\geq 2$  vertices contains  $\geq 1$  leaf.

In a tree T with  $\geq 2$  vertices, for a vertex v in T, if

$$\deg(v) = 1$$

then v is called a leaf; otherwise, v is an internal vertex.



#### Lemma

Any tree T with  $\geq 2$  vertices contains  $\geq 1$  leaf.

Otherwise,  $\forall v \in V. \deg(v) \geq 2 \implies T$  has cycles.

2021 年 05 月 13 日

6/53

Any tree T with  $\geq 2$  vertices contains  $\geq 2$  leaves.

7/53

Any tree T with  $\geq 2$  vertices contains  $\geq 2$  leaves.

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

Any tree T with  $\geq 2$  vertices contains  $\geq 2$  leaves.

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

Consider the two endpoints of any maximal (nontrivial) path in T.

7/53

Any tree T with  $\geq 2$  vertices contains  $\geq 2$  leaves.

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

Consider the two endpoints of any maximal (nontrivial) path in T. They are leaves of T.

Deleting a leaf from a tree T with n vertices produces a tree with n-1 vertices.

Deleting a leaf from a tree T with n vertices produces a tree with n-1 vertices.



G' = G - v is connected and acyclic.

Deleting a leaf from a tree T with n vertices produces a tree with n-1 vertices.



G' = G - v is connected and acyclic.

A leaf does not belong to any paths connecting two other vertices.

Deleting a leaf from a tree T with n vertices produces a tree with n-1 vertices.



G' = G - v is connected and acyclic.

A leaf does not belong to any paths connecting two other vertices.

This lemma can be used in induction for trees!

## Theorem ((We call it) Characterization of Trees)

Let T be an undirected graph with n vertices.

Then the following statements are equivalent:

- (1) T is a tree;
- (2) T is acyclic, and has m = n 1 edges;
- (3) T is connected, and has m = n 1 edges;
- (4) T is connected, and each edge is a bridge;
- (5) Any two vertices of T are connected by exactly one path;
- (6) T is acyclic, but the addition of any edge creates exactly one cycle.

## Theorem ((We call it) Characterization of Trees)

Let T be an undirected graph with n vertices.

Then the following statements are equivalent:

- (1) *T* is a tree;
- (2) T is acyclic, and has m = n 1 edges;
- (3) T is connected, and has m = n 1 edges;
- (4) T is connected, and each edge is a bridge;
- (5) Any two vertices of T are connected by exactly one path;
- (6) T is acyclic, but the addition of any edge creates exactly one cycle.

$$(1) \implies (2) \implies (3) \implies (4) \implies (5) \implies (6) \implies (1)$$



- (1) *T* is a tree;
- (2) T is acyclic, and has m = n 1 edges.

- (1) *T* is a tree;
- (2) T is acyclic, and has m = n 1 edges.

By induction on the number n of vertices of trees.

- (1) *T* is a tree;
- (2) T is acyclic, and has m = n 1 edges.

By induction on the number n of vertices of trees.

Basis Step: n = 1. m = 0 = n - 1.

- (1) *T* is a tree;
- (2) T is acyclic, and has m = n 1 edges.

By induction on the number n of vertices of trees.

Basis Step: n = 1. m = 0 = n - 1.

Induction Hypothesis: Any trees with n-1 vertices has n-2 edges.

- (1) *T* is a tree;
- (2) T is acyclic, and has m = n 1 edges.

By induction on the number n of vertices of trees.

Basis Step: n = 1. m = 0 = n - 1.

Induction Hypothesis: Any trees with n-1 vertices has n-2 edges.

Induction Step: Consider a tree T with  $n \geq 2$  vertices.

- (1) *T* is a tree;
- (2) T is acyclic, and has m = n 1 edges.

By induction on the number n of vertices of trees.

Basis Step: n = 1. m = 0 = n - 1.

Induction Hypothesis: Any trees with n-1 vertices has n-2 edges.

Induction Step: Consider a tree T with  $n \geq 2$  vertices.

T has a leaf v.

- (1) *T* is a tree;
- (2) T is acyclic, and has m = n 1 edges.

By induction on the number n of vertices of trees.

Basis Step: n = 1. m = 0 = n - 1.

Induction Hypothesis: Any trees with n-1 vertices has n-2 edges.

Induction Step: Consider a tree T with  $n \geq 2$  vertices.

T has a leaf v.

For 
$$T' = T - v$$
,  $m(T') = (n - 1) - 1 = n - 2$ .

- (1) T is a tree;
- (2) T is acyclic, and has m = n 1 edges.

By induction on the number n of vertices of trees.

Basis Step: 
$$n = 1$$
.  $m = 0 = n - 1$ .

Induction Hypothesis: Any trees with n-1 vertices has n-2 edges.

Induction Step: Consider a tree T with  $n \geq 2$  vertices.

T has a leaf v.

For 
$$T' = T - v$$
,  $m(T') = (n - 1) - 1 = n - 2$ .

$$m(T) = (n-2) + 1 = n - 1.$$



- (2) T is acyclic, and has n-1 edges;
- (3) T is connected, and has n-1 edges.

- (2) T is acyclic, and has n-1 edges;
- (3) T is connected, and has n-1 edges.

By Contradiction.

- (2) T is acyclic, and has n-1 edges;
- (3) T is connected, and has n-1 edges.

By Contradiction.

Suppose that T is disconnected.

- (2) T is acyclic, and has n-1 edges;
- (3) T is connected, and has n-1 edges.

## By Contradiction.

Suppose that T is disconnected.

T is a forest, consisting of  $k \geq 2$  trees  $T_1, T_2, \ldots$ 

- (2) T is acyclic, and has n-1 edges;
- (3) T is connected, and has n-1 edges.

## By Contradiction.

Suppose that T is disconnected.

T is a forest, consisting of  $k \geq 2$  trees  $T_1, T_2, \ldots$ 

By (2), for each  $T_i$ ,  $m(T_i) = n(T_i) - 1$ .

- (2) T is acyclic, and has n-1 edges;
- (3) T is connected, and has n-1 edges.

## By Contradiction.

Suppose that T is disconnected.

T is a forest, consisting of  $k \geq 2$  trees  $T_1, T_2, \ldots$ 

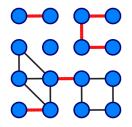
By (2), for each 
$$T_i$$
,  $m(T_i) = n(T_i) - 1$ .

$$m(T) = \sum_{i=1}^{k} m(T_i) = n - k \neq n - 1.$$



- (3) T is connected, and has n-1 edges;
- (4) T is connected, and each edge is a bridge.

- (3) T is connected, and has n-1 edges;
- (4) T is connected, and each edge is a bridge.



# Definition (Bridge (桥))

A bridge of a graph G is an edge e such that

$$c(G - e) > c(G)$$
.

- (3) T is connected, and has n-1 edges;
- (4) T is connected, and each edge is a bridge.

- (3) T is connected, and has n-1 edges;
- (4) T is connected, and each edge is a bridge.

Consider any edge e of T.

- (3) T is connected, and has n-1 edges;
- (4) T is connected, and each edge is a bridge.

Consider any edge e of T.

$$m(T - e) = (n - 1) - 1 = n - 2.$$

- (3) T is connected, and has n-1 edges;
- (4) T is connected, and each edge is a bridge.

Consider any edge e of T.

$$m(T - e) = (n - 1) - 1 = n - 2.$$

T - e must be disconnected.

- (4) T is connected, and each edge is a bridge;
- (5) Any two vertices of T are connected by exactly one path.

- (4) T is connected, and each edge is a bridge;
- (5) Any two vertices of T are connected by exactly one path.

Consider any two vertices u and v.

- (4) T is connected, and each edge is a bridge;
- (5) Any two vertices of T are connected by exactly one path.

Consider any two vertices u and v.

T is connected  $\implies u$  and v are connected by  $\geq 1$  path.

- (4) T is connected, and each edge is a bridge;
- (5) Any two vertices of T are connected by exactly one path.

Consider any two vertices u and v.

T is connected  $\implies u$  and v are connected by  $\geq 1$  path.

If u and v are connected by two paths, the edges on these two paths are not bridges.

- (5) Any two vertices of T are connected by exactly one path;
- (6) T is acyclic, but the addition of any edge creates exactly one cycle.

- (5) Any two vertices of T are connected by exactly one path;
- (6) T is acyclic, but the addition of any edge creates exactly one cycle.

If T has a cycle C,

any two vertices in C is connected by  $\geq 2$  paths.

- (5) Any two vertices of T are connected by exactly one path;
- (6) T is acyclic, but the addition of any edge creates exactly one cycle.

### If T has a cycle C,

any two vertices in C is connected by  $\geq 2$  paths.

Consider the addition of edge  $\{u, v\}$  to T.

It creates a cycle, consisting of  $\{u, v\}$  and the path from u to v.

- (5) Any two vertices of T are connected by exactly one path;
- (6) T is acyclic, but the addition of any edge creates exactly one cycle.

### If T has a cycle C,

any two vertices in C is connected by  $\geq 2$  paths.

Consider the addition of edge  $\{u, v\}$  to T.

It creates a cycle, consisting of  $\{u, v\}$  and the path from u to v.

#### Lemma

If two distinct cycles of a graph G share a common edge e, then G has a cycle that does not contain e.

- $(6)\ T\ is\ acyclic,\ but\ the\ addition\ of\ any\ edge\ creates\ exactly\ one\ cycle;$
- (1) T is a tree.

- (6) T is acyclic, but the addition of any edge creates exactly one cycle;
- (1) T is a tree.

Suppose that T is disconnected.

- (6) T is acyclic, but the addition of any edge creates exactly one cycle;
- (1) T is a tree.

Suppose that T is disconnected.

T is a forest, consisting of  $\geq 2$  trees  $T_1, T_2, \ldots$ 

- (6) T is acyclic, but the addition of any edge creates exactly one cycle;
- (1) T is a tree.

# Suppose that T is disconnected.

T is a forest, consisting of  $\geq 2$  trees  $T_1, T_2, \ldots$ 

Choose 
$$u \in V(T_1), v \in V(T_2)$$
.

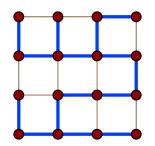
 $T + \{u, v\}$  does **not** create cycles.

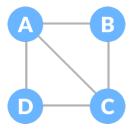
Definition (Spanning Tree (生成树))

A spanning tree T of an undirected graph G is a subgraph that is a tree with all vertices of G.

# Definition (Spanning Tree (生成树))

A spanning tree T of an undirected graph G is a subgraph that is a tree with all vertices of G.





Definition (Subgraph (子图))

Definition (Subgraph (子图))

Definition (Induced Subgraph (诱导子图))

### Theorem

Every connected undirected graph G admits a spanning tree.

### Theorem

Every connected undirected graph G admits a spanning tree.

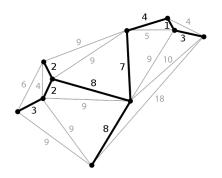
Repeatedly deleting vertices in cycles until the graph is acyclic.

# Definition (Minimum Spanning Tree (MST; 最小生成树))

A minimum spanning tree T of an edge-weighted undirected graph G is a spanning tree with minimum total weight of edges.

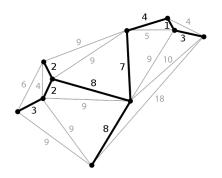
# Definition (Minimum Spanning Tree (MST; 最小生成树))

A minimum spanning tree T of an edge-weighted undirected graph G is a spanning tree with minimum total weight of edges.



# Definition (Minimum Spanning Tree (MST; 最小生成树))

A minimum spanning tree T of an edge-weighted undirected graph G is a spanning tree with minimum total weight of edges.



Existence?

Uniqueness?

Algorithms?

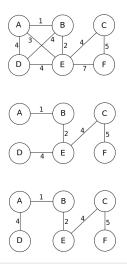
20/53

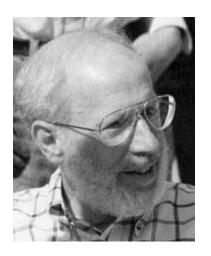
### Theorem

Every weighted connected undirected graph G admits a minimum spanning tree.

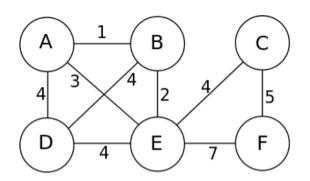
### Theorem

Every weighted connected undirected graph G admits a minimum spanning tree.



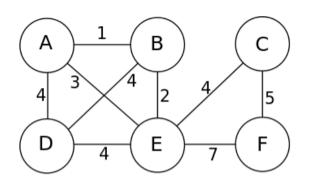


Joseph Kruskal (1928  $\sim 2010)$ 





Robert C. Prim  $(1921 \sim)$ 



Cut Property

Cut Property (Version I)

X: A part of some MST  $T_1$  of G

 $(S, V \setminus S)$ : A *cut* such that X does *not* cross  $(S, V \setminus S)$ 

e: A lightest edge across  $(S, V \setminus S)$ 

Cut Property (Version I)

X: A part of some MST  $T_1$  of G

 $(S, V \setminus S)$ : A *cut* such that X does *not* cross  $(S, V \setminus S)$ 

e: A lightest edge across  $(S, V \setminus S)$ 

Then  $X \cup \{e\}$  is a part of some MST  $T_2$  of G.

Cut Property (Version I)

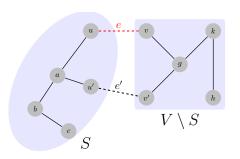
X: A part of some MST  $T_1$  of G

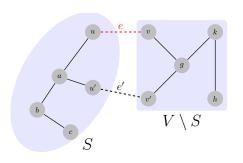
 $(S, V \setminus S)$ : A *cut* such that X does *not* cross  $(S, V \setminus S)$ 

e: A lightest edge across  $(S, V \setminus S)$ 

Then  $X \cup \{e\}$  is a part of some MST  $T_2$  of G.

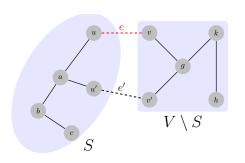
Correctness of Prim's and Kruskal's algorithms.





$$T' = \underbrace{T}_{X \subseteq T} + \{e\} - \{e'\}$$
if  $e \notin T$ 

28 / 53



$$T' = \underbrace{T}_{X \subseteq T} + \{e\} - \{e'\}$$
if  $e \notin T$ 

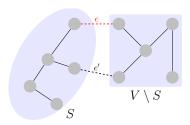
"a"  $\rightarrow$  "the"  $\Longrightarrow$  "some"  $\rightarrow$  "all"

## Cut Property (Version II)

A cut  $(S, V \setminus S)$ 

Let e = (u, v) be <u>a</u> lightest edge across  $(S, V \setminus S)$ 

 $\exists$  MST T of  $G: e \in T$ 

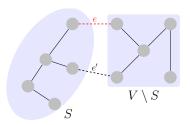


#### Cut Property (Version II)

A cut 
$$(S, V \setminus S)$$

Let e = (u, v) be <u>a</u> lightest edge across  $(S, V \setminus S)$ 

 $\exists$  MST T of  $G: e \in T$ 



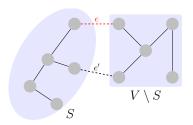
$$T' = \underbrace{T + \{e\}}_{\text{if } e \notin T} - \{e'\}$$

#### Cut Property (Version II)

A cut 
$$(S, V \setminus S)$$

Let e = (u, v) be a lightest edge across  $(S, V \setminus S)$ 

#### $\exists$ MST T of $G: e \in T$



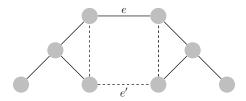
$$T' = \underbrace{T + \{e\}}_{\text{if } e \notin T} - \{e'\}$$

"a" 
$$\rightarrow$$
 "the"  $\Longrightarrow$  " $\exists$ "  $\rightarrow$  " $\forall$ "



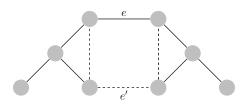
- ightharpoonup Let C be any cycle in G
- ▶ Let e = (u, v) be **a** maximum-weight edge in C

Then  $\exists$  MST T of  $G: e \notin T$ .



- $\blacktriangleright$  Let C be any cycle in G
- Let e = (u, v) be a maximum-weight edge in C

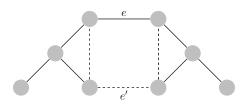
Then  $\exists$  MST T of  $G: e \notin T$ .



$$T' = \underbrace{T - \{e\}}_{\text{if } e \in T} + \{e'\}$$

- $\blacktriangleright$  Let C be any cycle in G
- ▶ Let e = (u, v) be a maximum-weight edge in C

Then  $\exists$  MST T of  $G: e \notin T$ .



$$T' = \underbrace{T - \{e\}}_{\text{if } e \in T} + \{e'\}$$

"a" 
$$\rightarrow$$
 "the"  $\Longrightarrow$  " $\exists$ "  $\rightarrow$  " $\forall$ "



Joseph Kruskal (1928  $\sim 2010)$ 

### Anti-Kruskal Algorithm

Reverse-delete algorithm (wiki; clickable)

#### Anti-Kruskal Algorithm

#### Reverse-delete algorithm (wiki; clickable)

Cycle Property

$$T \subseteq F \implies \exists T' : T' \subseteq F - \{e\}$$

### Anti-Kruskal Algorithm

#### Reverse-delete algorithm (wiki; clickable)

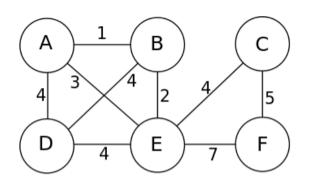
## Cycle Property

$$T \subseteq F \implies \exists T' : T' \subseteq F - \{e\}$$

"On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem"

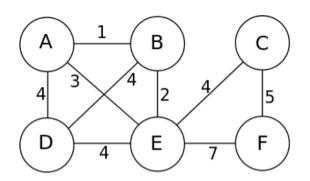
— Kruskal, 1956.

33 / 53





Otakar Borůvka (1899  $\sim 1995)$ 



 $Let \ G \ be \ an \ edge\text{-}weighted \ undirected \ graph.$ 

If each edge has a distinct weight, then there is a unique MST of G.

Let G be an edge-weighted undirected graph. If each edge has a distinct weight, then there is a unique MST of G.

Let G be an edge-weighted undirected graph. If each edge has a distinct weight, then there is a unique MST of G.

By Contradiction.

 $\exists$  MSTs  $T_1 \neq T_2$ 

 $Let \ G \ be \ an \ edge\text{-}weighted \ undirected \ graph.$ 

If each edge has a distinct weight, then there is a unique MST of G.

$$\exists \text{ MSTs } T_1 \neq T_2$$

$$\Delta E = \{ e \mid e \in T_1 \setminus T_2 \lor e \in T_2 \setminus T_1 \}$$

 $Let \ G \ be \ an \ edge\text{-}weighted \ undirected \ graph.$ 

If each edge has a distinct weight, then there is a unique MST of G.

$$\exists$$
 MSTs  $T_1 \neq T_2$ 

$$\Delta E = \{ e \mid e \in T_1 \setminus T_2 \lor e \in T_2 \setminus T_1 \}$$

$$e = \min \Delta E$$

Let G be an edge-weighted undirected graph.

If each edge has a distinct weight, then there is a unique MST of G.

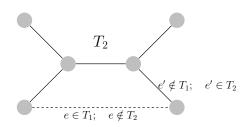
$$\exists$$
 MSTs  $T_1 \neq T_2$ 

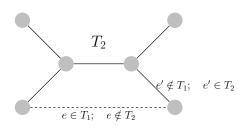
$$\Delta E = \{ e \mid e \in T_1 \setminus T_2 \lor e \in T_2 \setminus T_1 \}$$

$$e = \min \Delta E$$

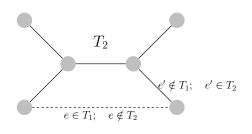
$$e \in T_1 \setminus T_2 \ (w.l.o.g)$$





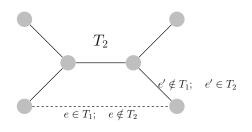


$$T_2 + \{e\} \implies C$$



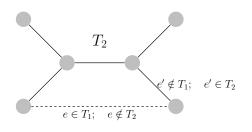
$$T_2 + \{e\} \implies C$$

$$\exists (e' \in C) \notin T_1$$



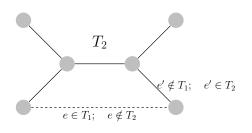
$$T_2 + \{e\} \implies C$$

$$\exists (e' \in C) \notin T_1 \implies e' \in T_2 \setminus T_1 \implies e' \in \Delta E$$



$$T_2 + \{e\} \implies C$$

$$\exists (e' \in C) \notin T_1 \implies e' \in T_2 \setminus T_1 \implies e' \in \Delta E \implies w(e') > w(e)$$



$$T_2 + \{e\} \implies C$$

$$\exists (e' \in C) \notin T_1 \implies e' \in T_2 \setminus T_1 \implies e' \in \Delta E \implies w(e') > w(e)$$

$$T' = T_2 + \{e\} - \{e'\} \implies w(T') < w(T_2)$$

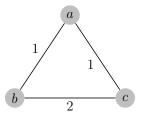


Condition for Uniqueness of MST

Unique MST  $\implies$  Distinct weights

### Condition for Uniqueness of MST

# Unique MST $\implies$ Distinct weights



Rooted Trees in Computer Science

Definition (Rooted Trees (有根树))

bfs

42 / 53

dfs: in-order, pre-order, post-order

search trees

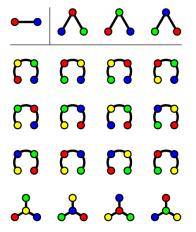


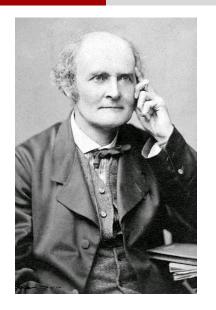
# Theorem (Cayley's Formula)

The number  $T_n$  of labeled trees on  $n \ge 2$  vertices is  $n^{n-2}$ .

### Theorem (Cayley's Formula)

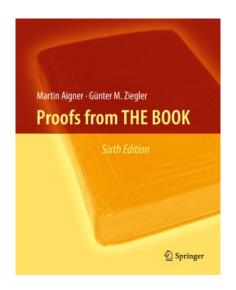
The number  $T_n$  of labeled trees on  $n \ge 2$  vertices is  $n^{n-2}$ .





Arthur Cayley (1821  $\sim 1895)$ 

47/53



Chapter 33: Cayley's formula for the number of trees

# By Double Counting.

— Jim Pitman

# By Double Counting.

— Jim Pitman

https://en.wikipedia.org/wiki/Double\_counting\_(proof\_technique)#Counting\_trees

# By Double Counting.

— Jim Pitman

https://en.wikipedia.org/wiki/Double\_counting\_(proof\_technique)#Counting\_trees

How many ways are there of forming a rooted tree from an empty graph by adding directed edges one by one? Choose one of the  $T_n$  labeled trees on n vertices.

Choose one of the  $T_n$  labeled trees on n vertices.

Choose one of its n vertices as root.

Choose one of the  $T_n$  labeled trees on n vertices.

Choose one of its n vertices as root.

Choose one of the (n-1)! possible sequences in which to add its n-1 directed edges.

$$\frac{T_n n(n-1)!}{T_n n!} = T_n n!$$

Suppose that we have added n-k directed edges.

Suppose that we have added n-k directed edges.

We obtain a rooted forest with k trees.

Suppose that we have added n - k directed edges.

We obtain a rooted forest with k trees.

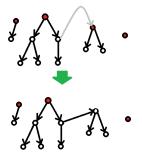
There are n(k-1) choices for the next edge to add.

51/53

Suppose that we have added n-k directed edges.

We obtain a rooted forest with k trees.

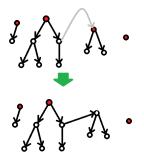
There are n(k-1) choices for the next edge to add.



Suppose that we have added n-k directed edges.

We obtain a rooted forest with k trees.

There are n(k-1) choices for the next edge to add.



$$\prod_{k=2}^{n} n(k-1) = n^{n-1}(n-1)! = n^{n-2}n!$$

$$T_n n! = n^{n-2} n!$$

$$T_n n! = n^{n-2} n!$$

$$T_n = n^{n-2}$$

$$T_n n! = n^{n-2} n!$$

$$T_n = n^{n-2}$$

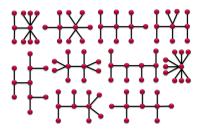


An irreducible tree is a tree T where

$$\forall v \in V(T). \deg(v) \neq 2.$$

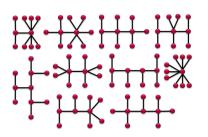
An irreducible tree is a tree T where

$$\forall v \in V(T). \deg(v) \neq 2.$$



An irreducible tree is a tree T where

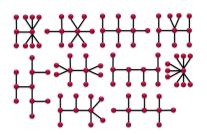
$$\forall v \in V(T). \deg(v) \neq 2.$$





An irreducible tree is a tree T where

$$\forall v \in V(T). \deg(v) \neq 2.$$





Homeomorphically Irreducible Trees of size n = 10

# Thank You!



Office 302

Mailbox: H016

hfwei@nju.edu.cn