Algorithms for Visualizing Phylogenetic Networks

Ioannis G. Tollis
Department of Computer Science, University of Crete, Greece

and

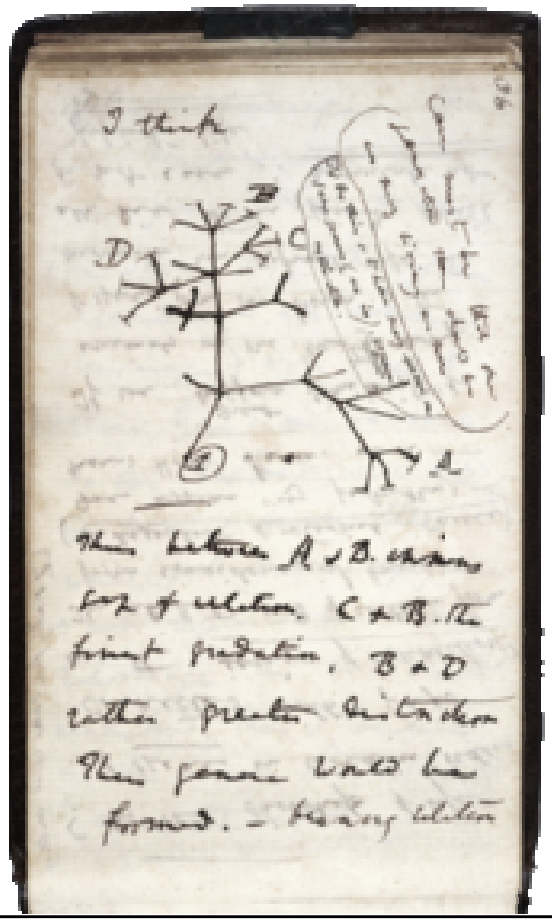
Konstantinos G. Kakoulis
Department of Mechanical and Industrial Design Engineering,
T.E.I. of West Macedonia, Greece.

PHYLOGENETIC TREE:

A model to represent evolution (of some species, or genes).

PHYLOGENETIC TREE:

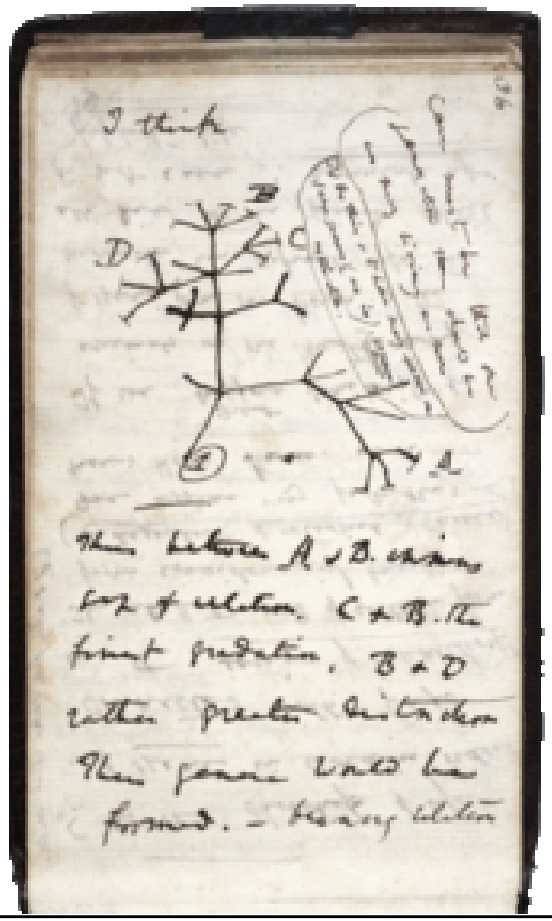
A model to represent evolution (of some species, or genes).



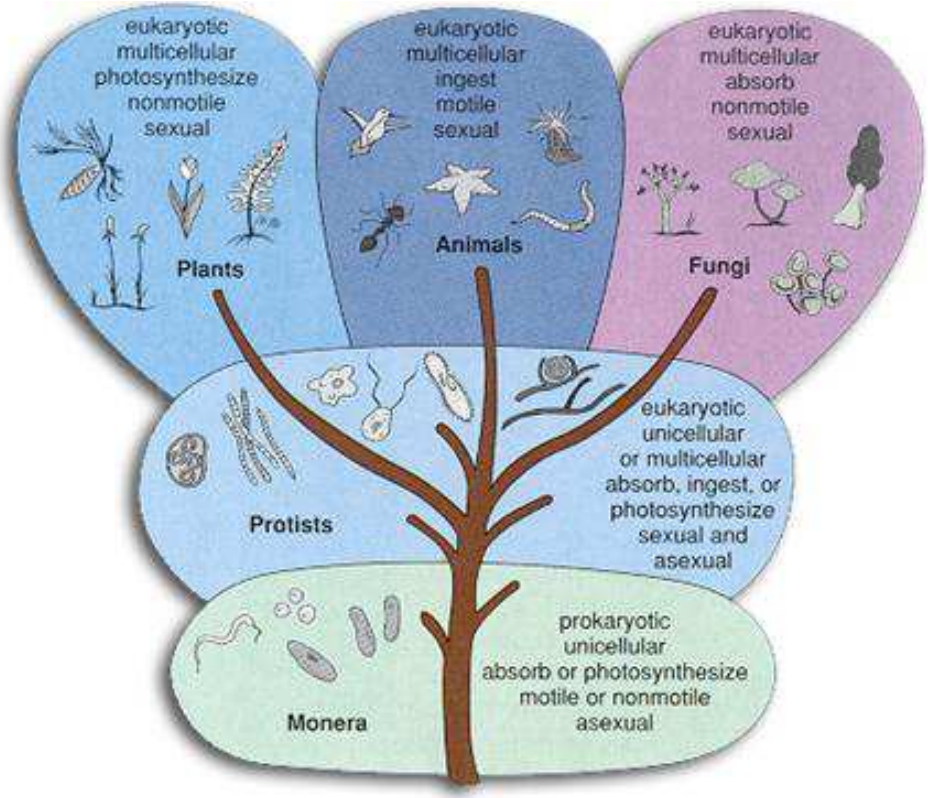
Darwin's first sketch of an evolutionary tree (1837)

PHYLOGENETIC TREE:

A model to represent evolution (of some species, or genes).



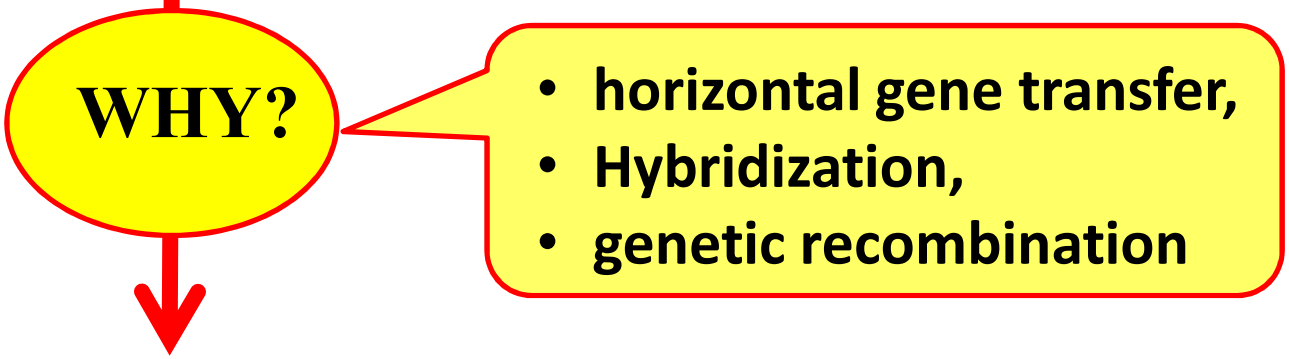
Darwin's first sketch of an evolutionary tree (1837)



Five Kingdom Classification (by R.H Whittaker,1969)

Evolution cannot be properly represented as a tree.

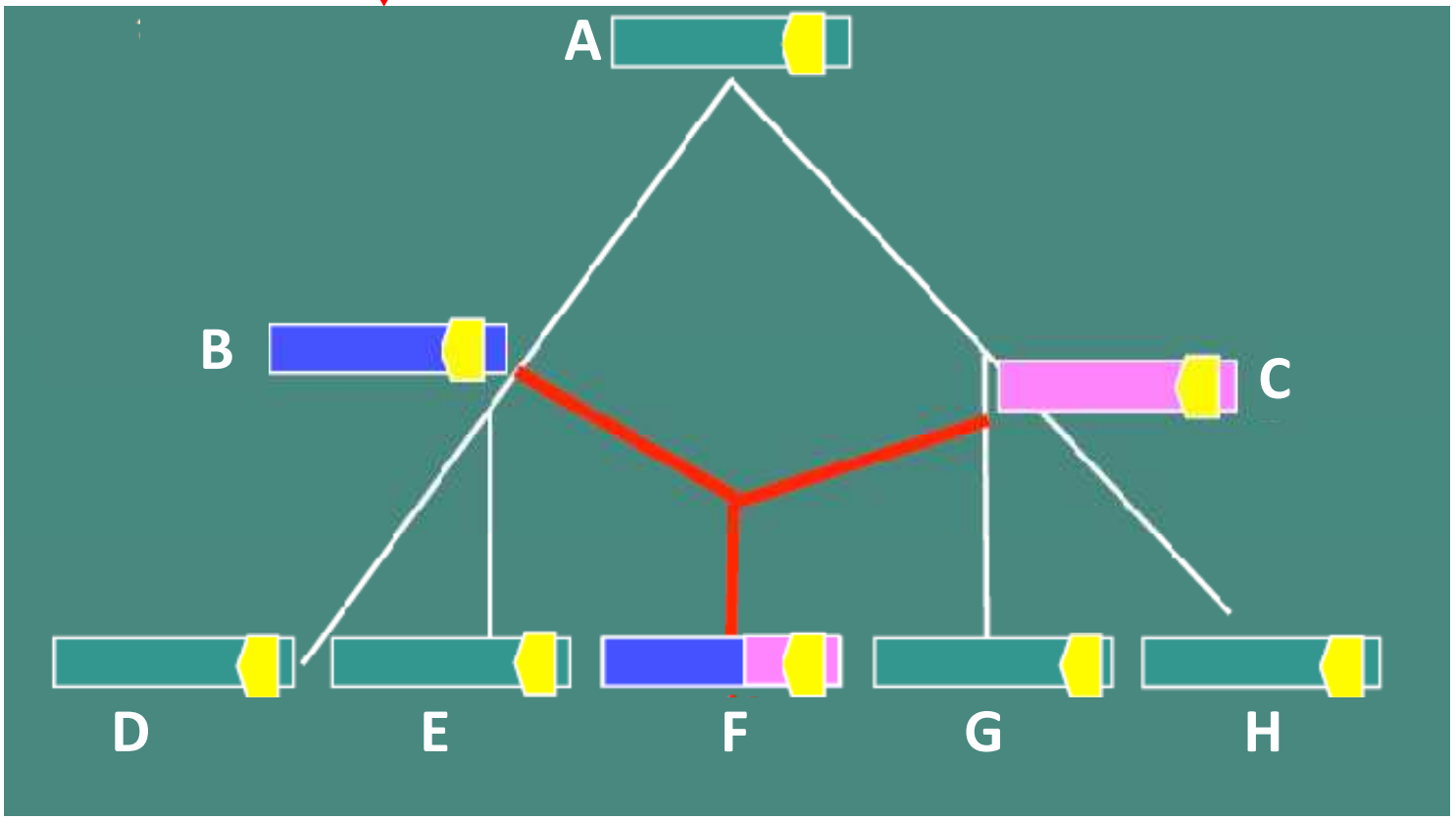
Evolution cannot be properly represented as a tree.



Evolution cannot be properly represented as a tree.

WHY?

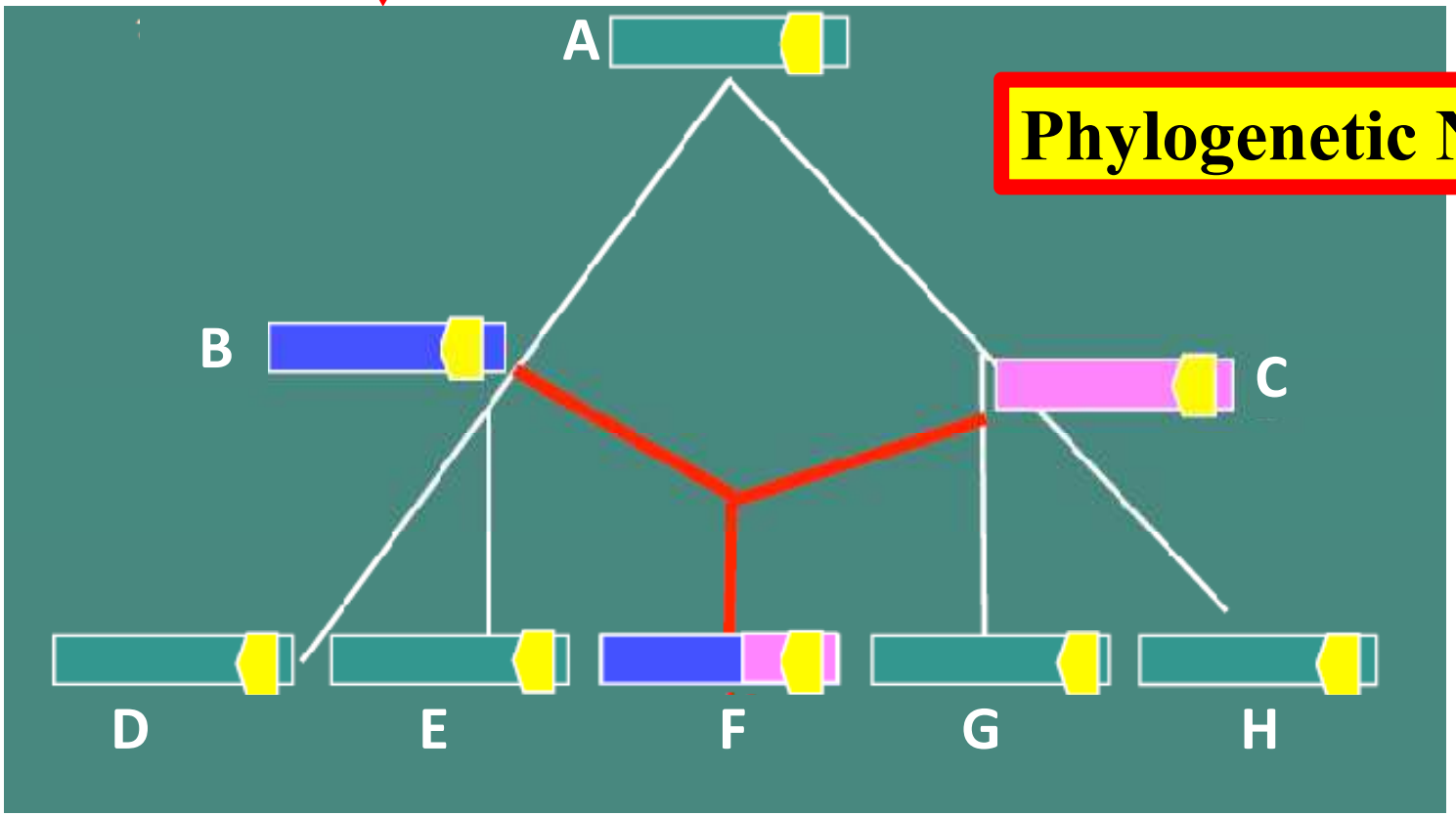
- **horizontal gene transfer,**
- **Hybridization,**
- **genetic recombination**



Evolution cannot be properly represented as a tree.

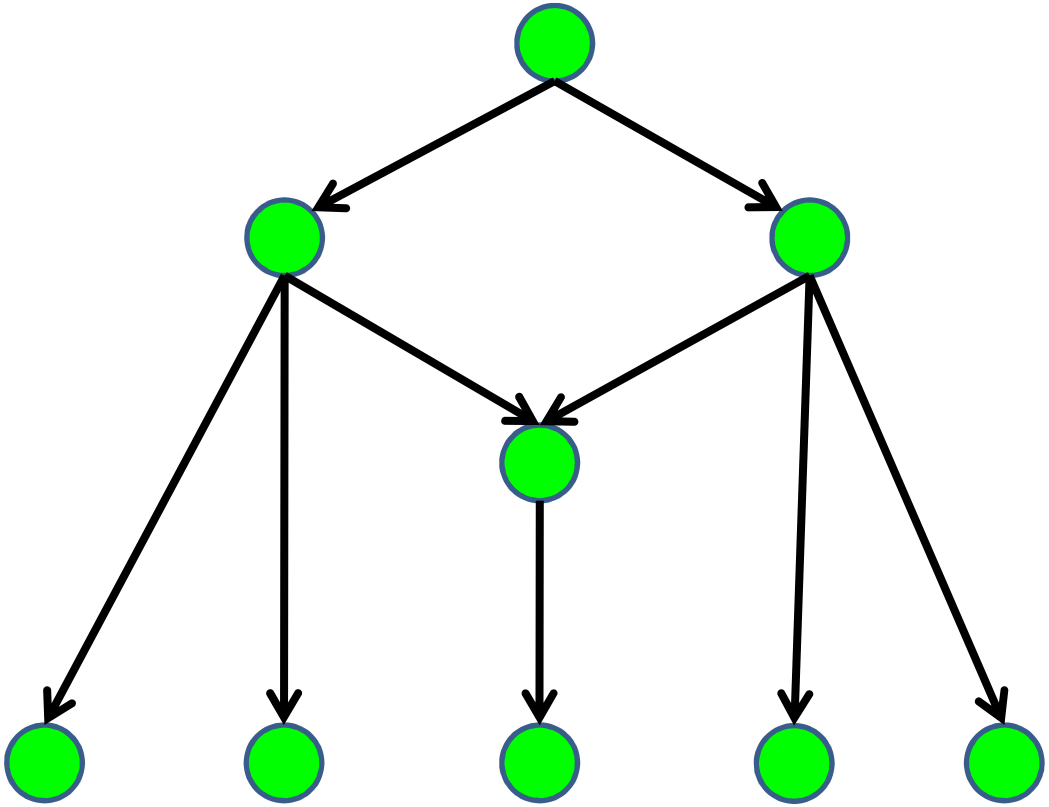
WHY?

- **horizontal gene transfer,**
- **Hybridization,**
- **genetic recombination**

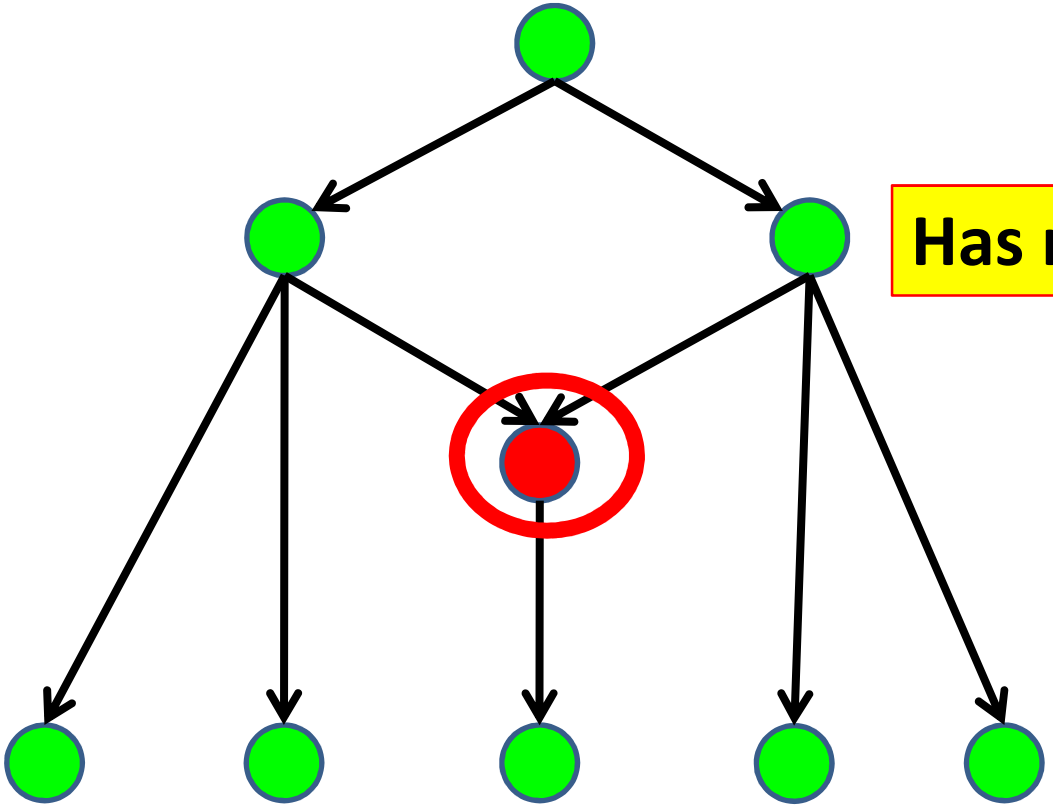


Phylogenetic Networks

Phylogenetic Networks



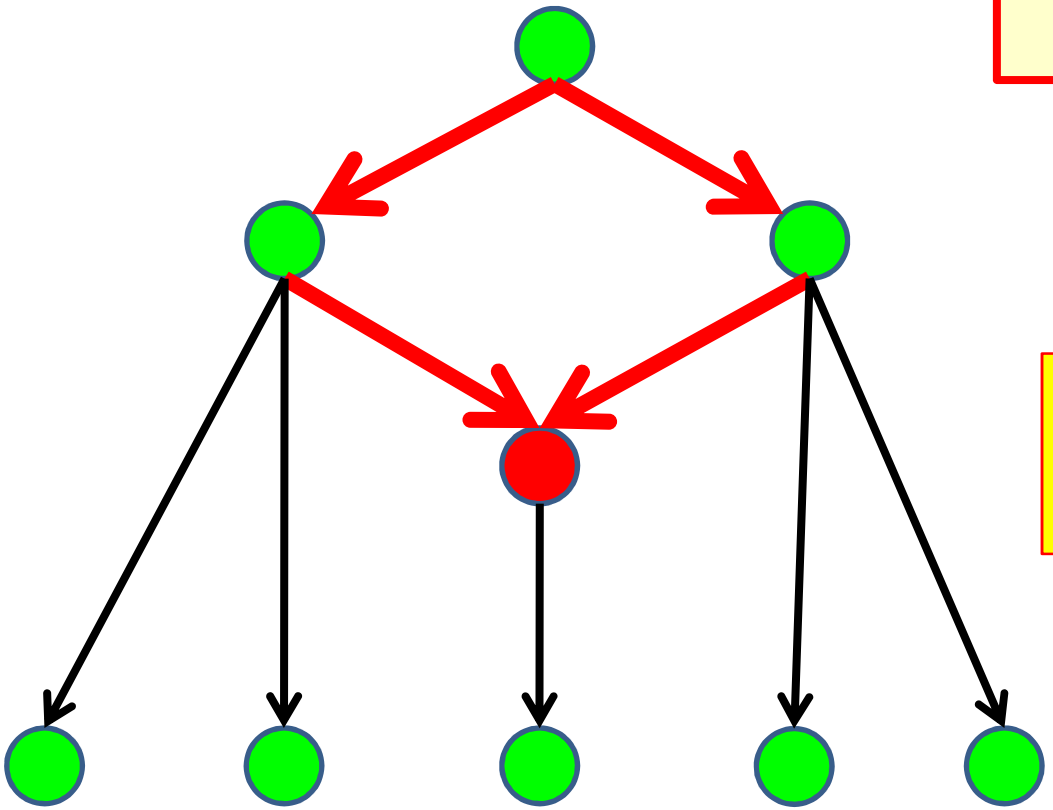
Phylogenetic Networks



Reticulation node

Has more than one ancestors.

Phylogenetic Networks



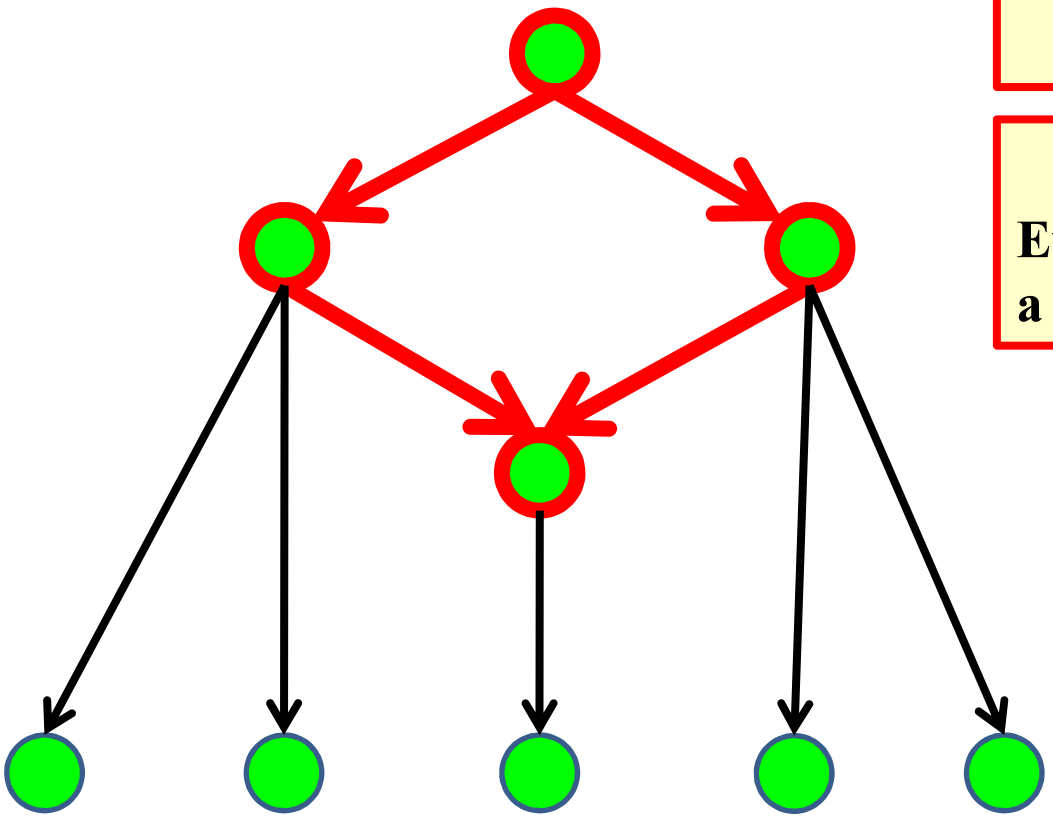
Reticulation node

Has more than one ancestors.

Reticulation cycle

Every reticulation node belongs to a cycle.

Phylogenetic Networks



Reticulation node

Has more than one ancestors.

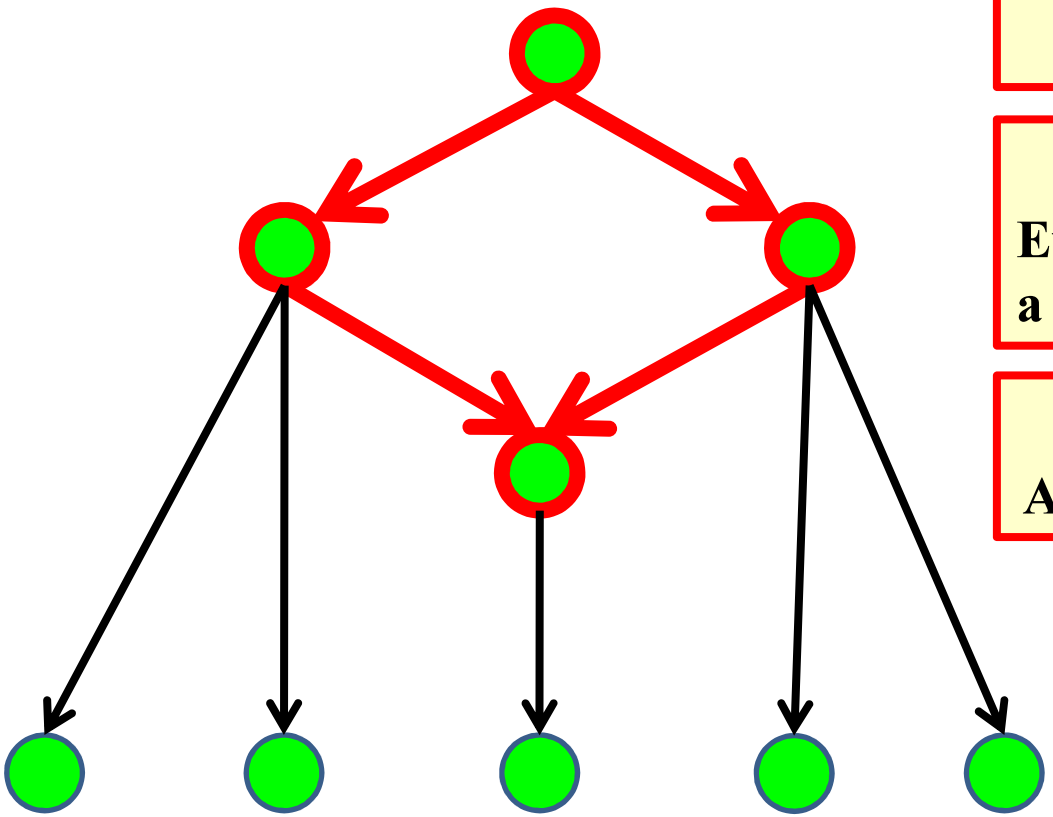
Reticulation cycle

Every reticulation node belongs to a cycle.

GALL

A single (isolated) reticulation cycle

Phylogenetic Networks



Reticulation node

Has more than one ancestors.

Reticulation cycle

Every reticulation node belongs to a cycle.

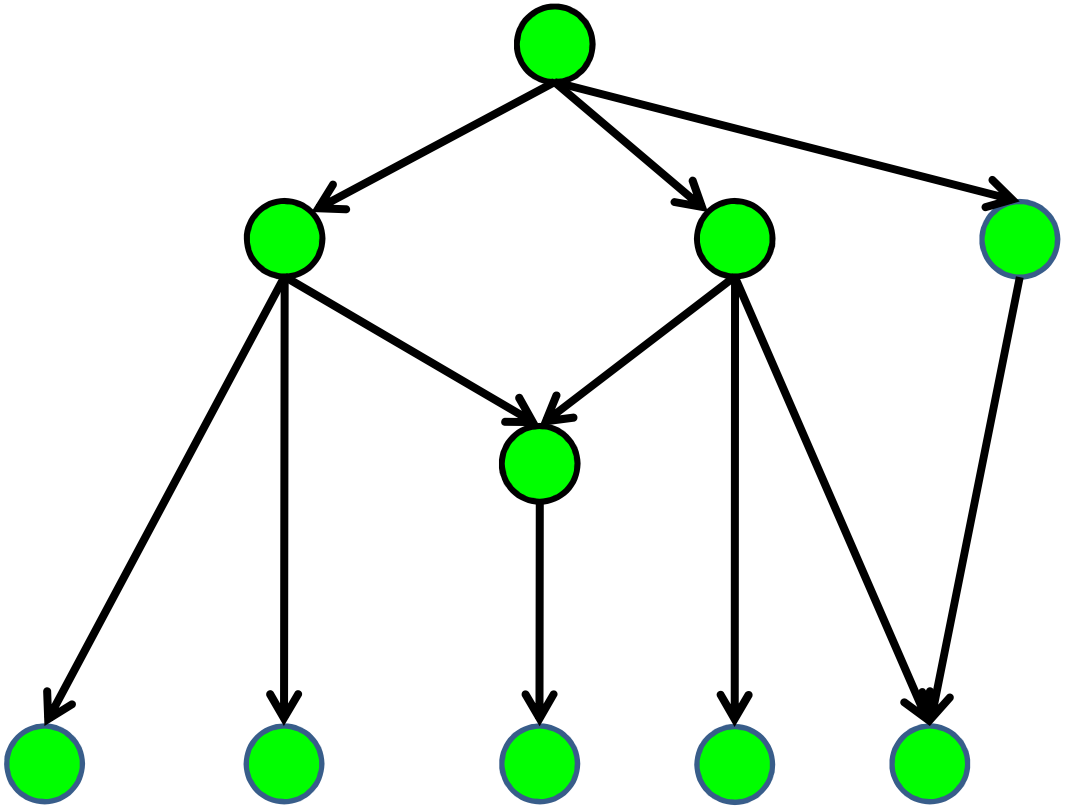
Gall

A single (isolated) reticulation cycle

Galled tree

A network in which the galls do not share edges or nodes

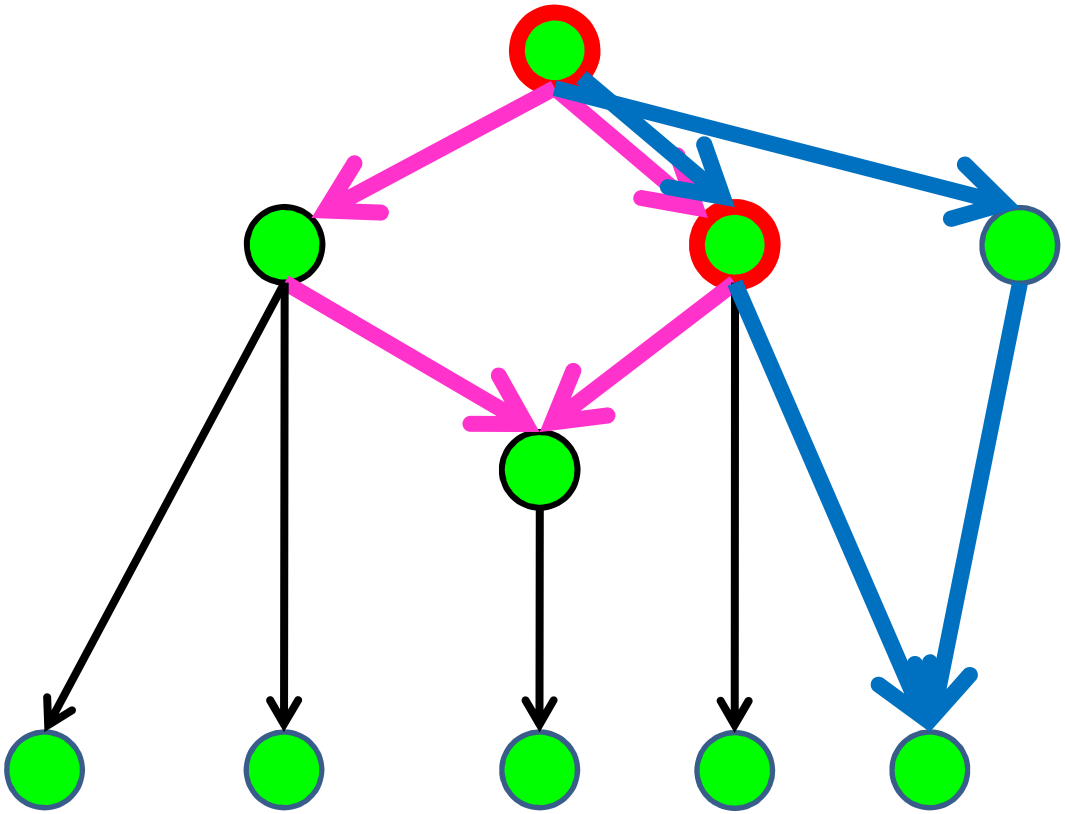
Phylogenetic Networks



Galled network

A network in which the galls can share edges but not reticulation nodes

Phylogenetic Networks

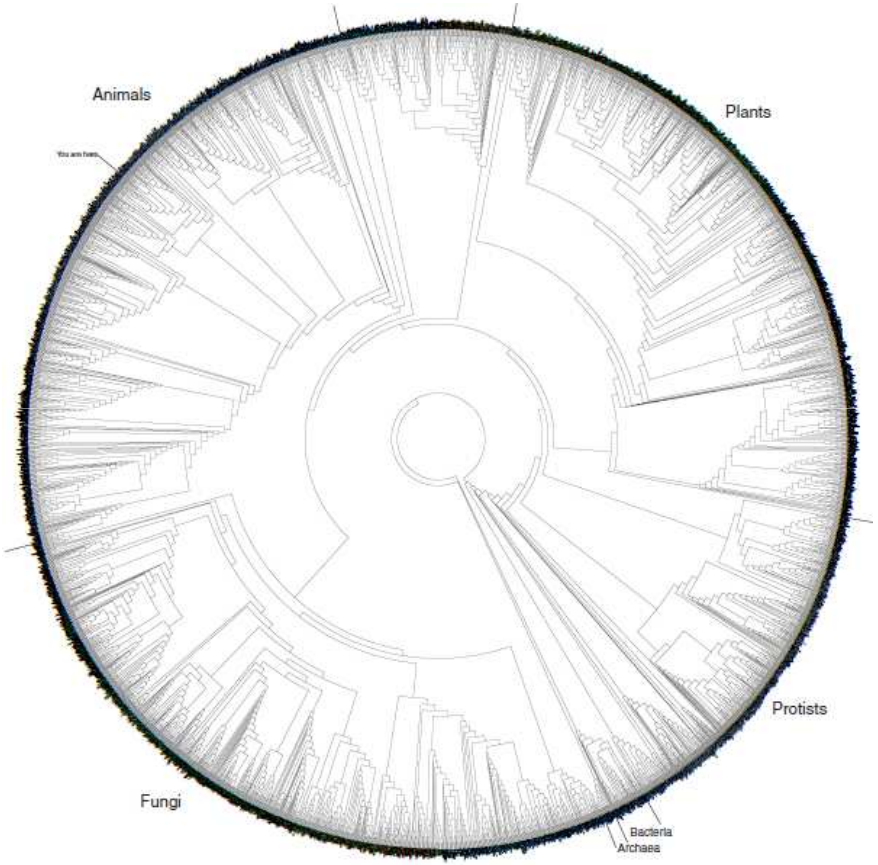


Galled network

A network in which the galls can share edges but not reticulation nodes

Visualization of Phylogenetic Trees and Networks

Visualization of Phylogenetic Trees and Networks

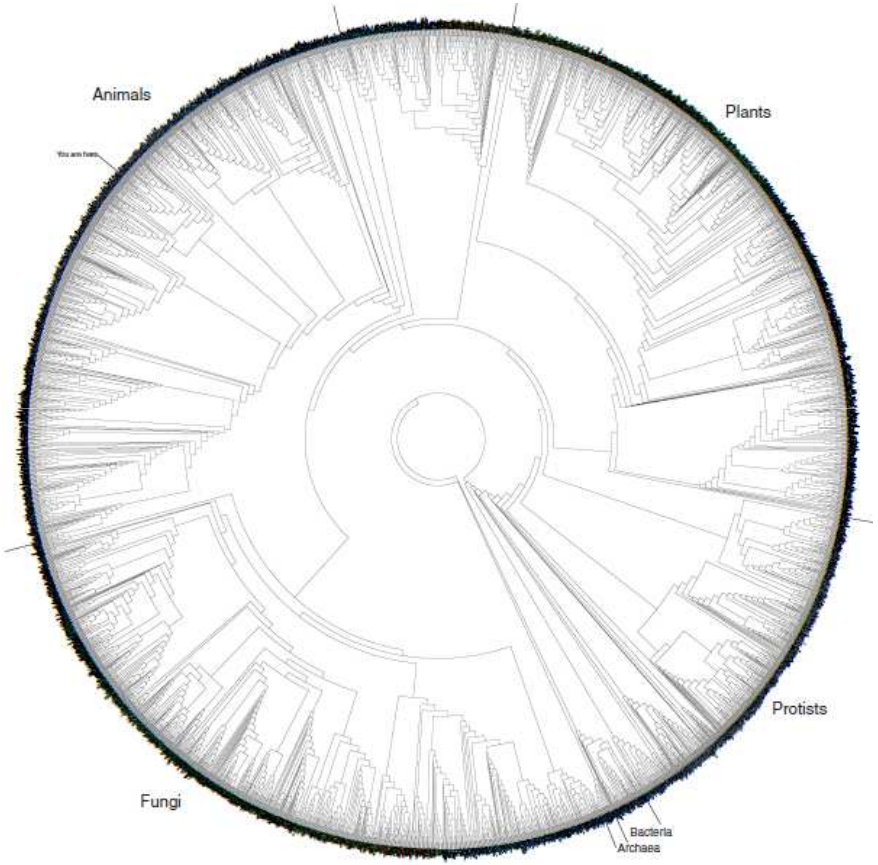


Tree of life for 3,000 of the 1.8 million known species based on RNA sequences.

node-link representation

- Huge graphs.
- Visual clutter.

Visualization of Phylogenetic Trees and Networks



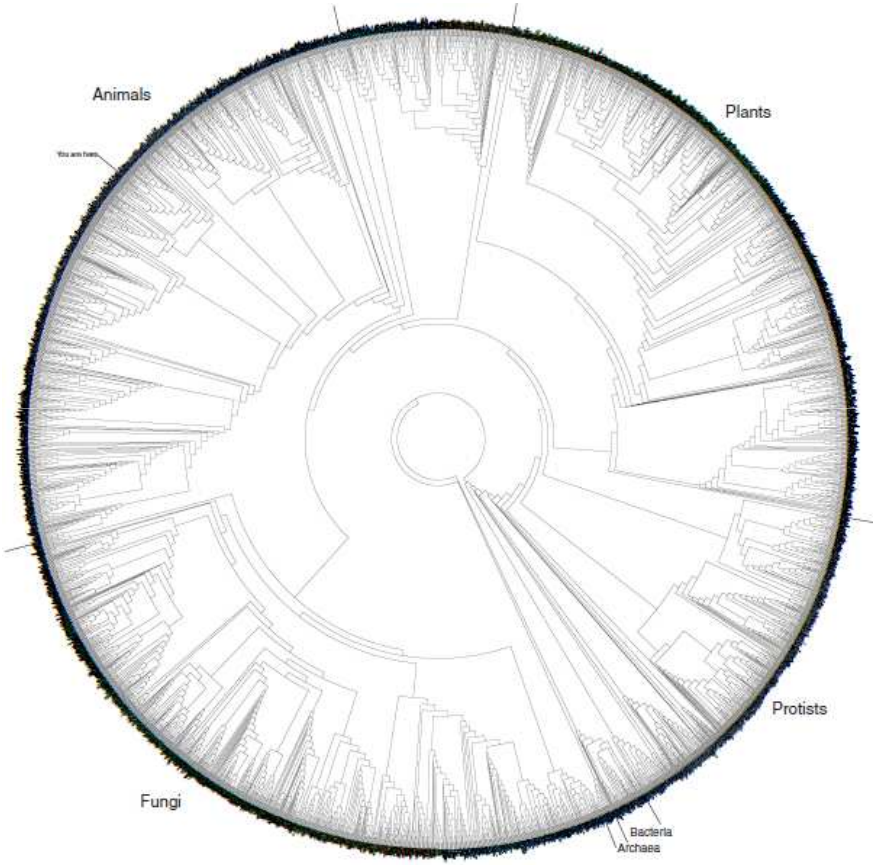
Tree of life for 3,000 of the 1.8 million known species based on RNA sequences.

node-link representation

- Huge graphs.
- Visual clutter.

alternative visualization ???

Visualization of Phylogenetic Trees and Networks



Tree of life for 3,000 of the 1.8 million known species based on RNA sequences.

node-link representation

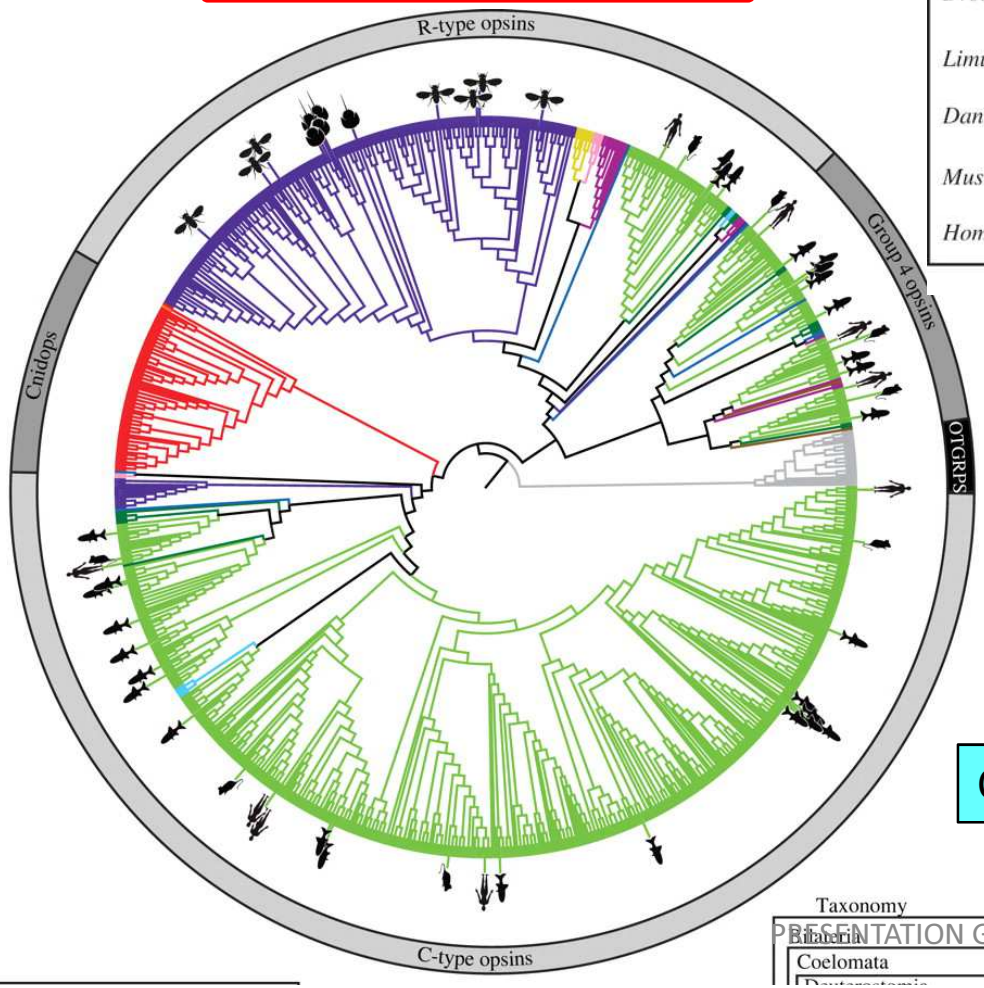
- Huge graphs.
- Visual clutter.

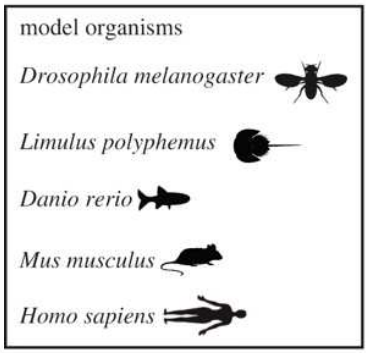
alternative visualization ???

- **Space filling techniques**
  - ✓**Treemaps**
  - ✓**DAGmaps**

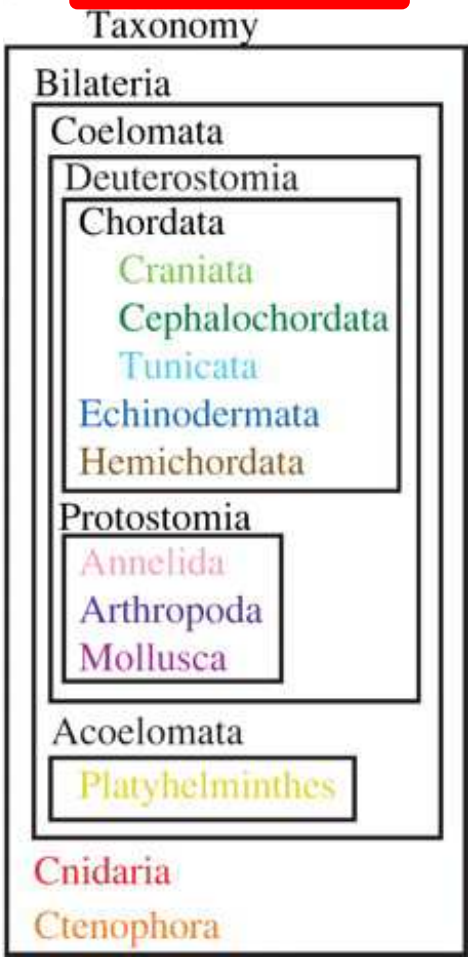
Visualization of Phylogenetic Trees and Networks

node-link



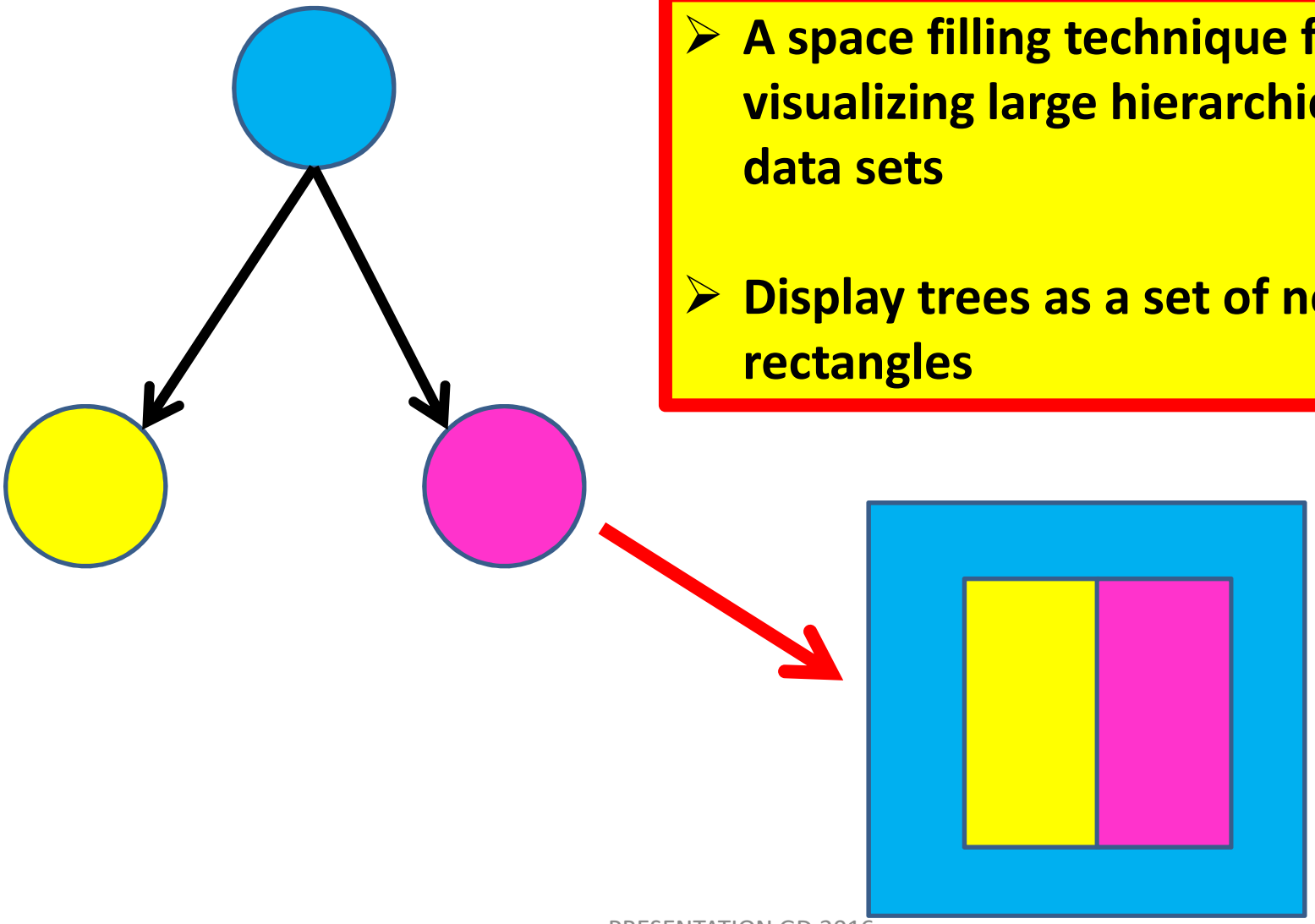


Treemap



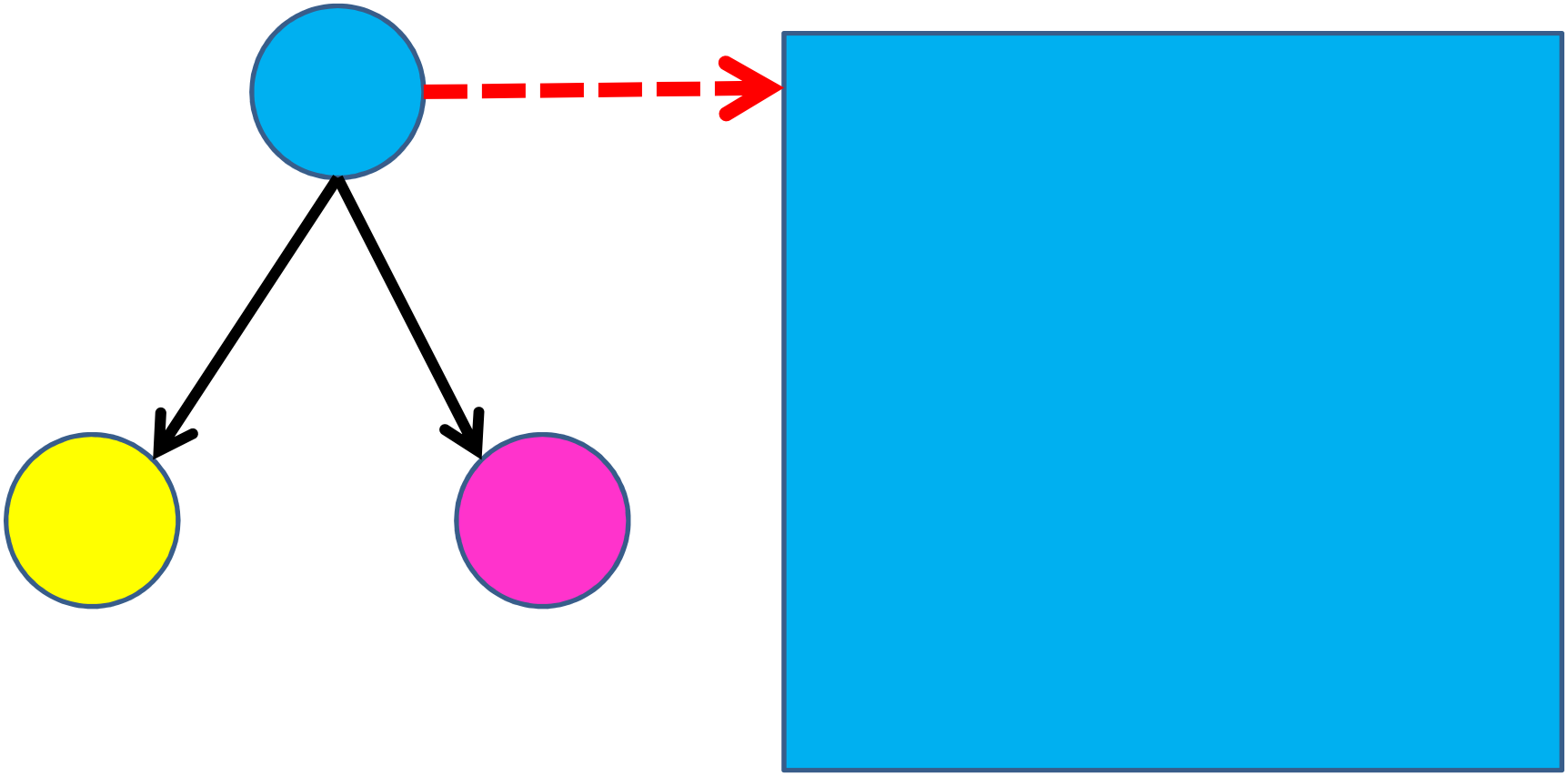
C-type opsins

TREEMAP DRAWINGS (Johnson and Shneiderman, 1990)

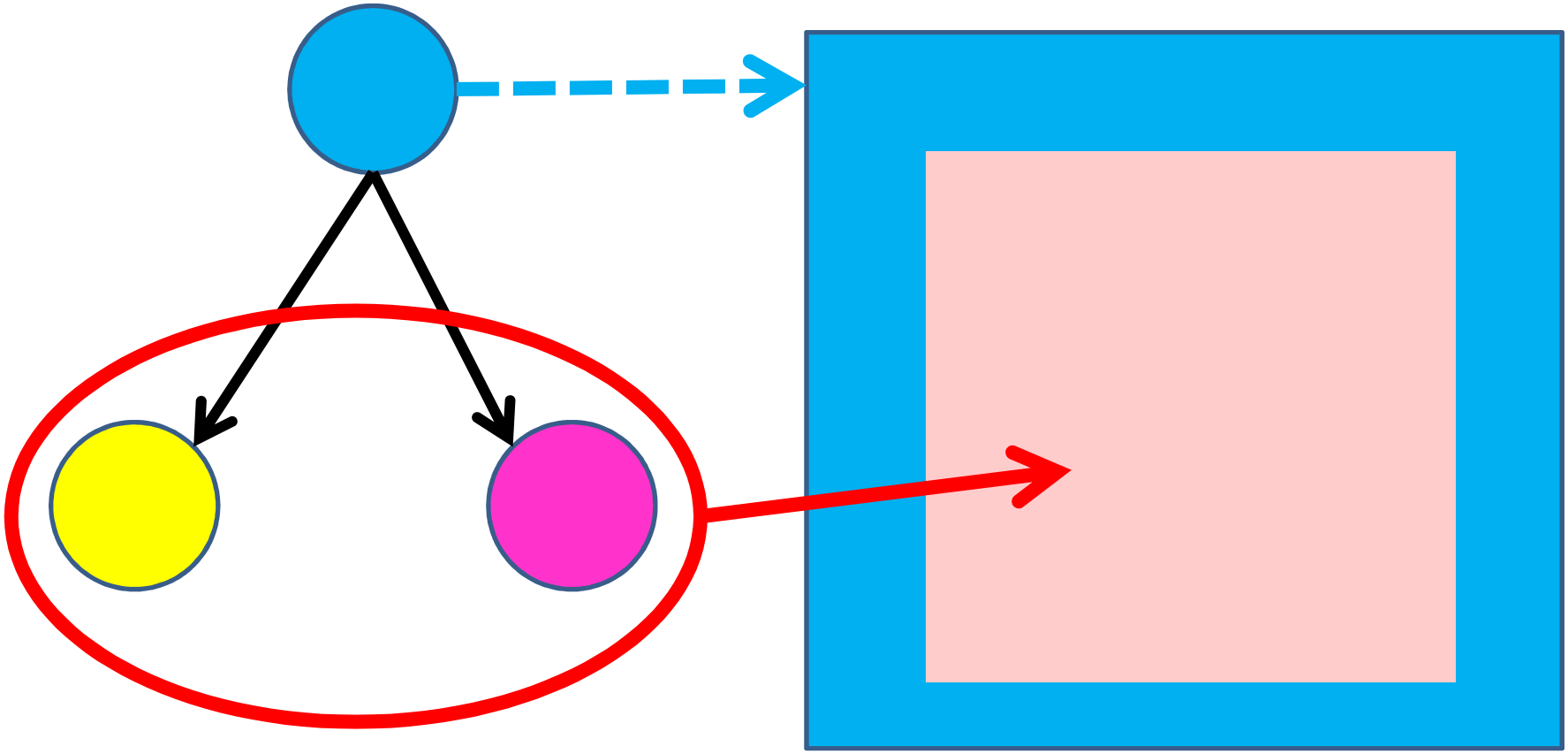


- A space filling technique for visualizing large hierarchical data sets
- Display trees as a set of nested rectangles

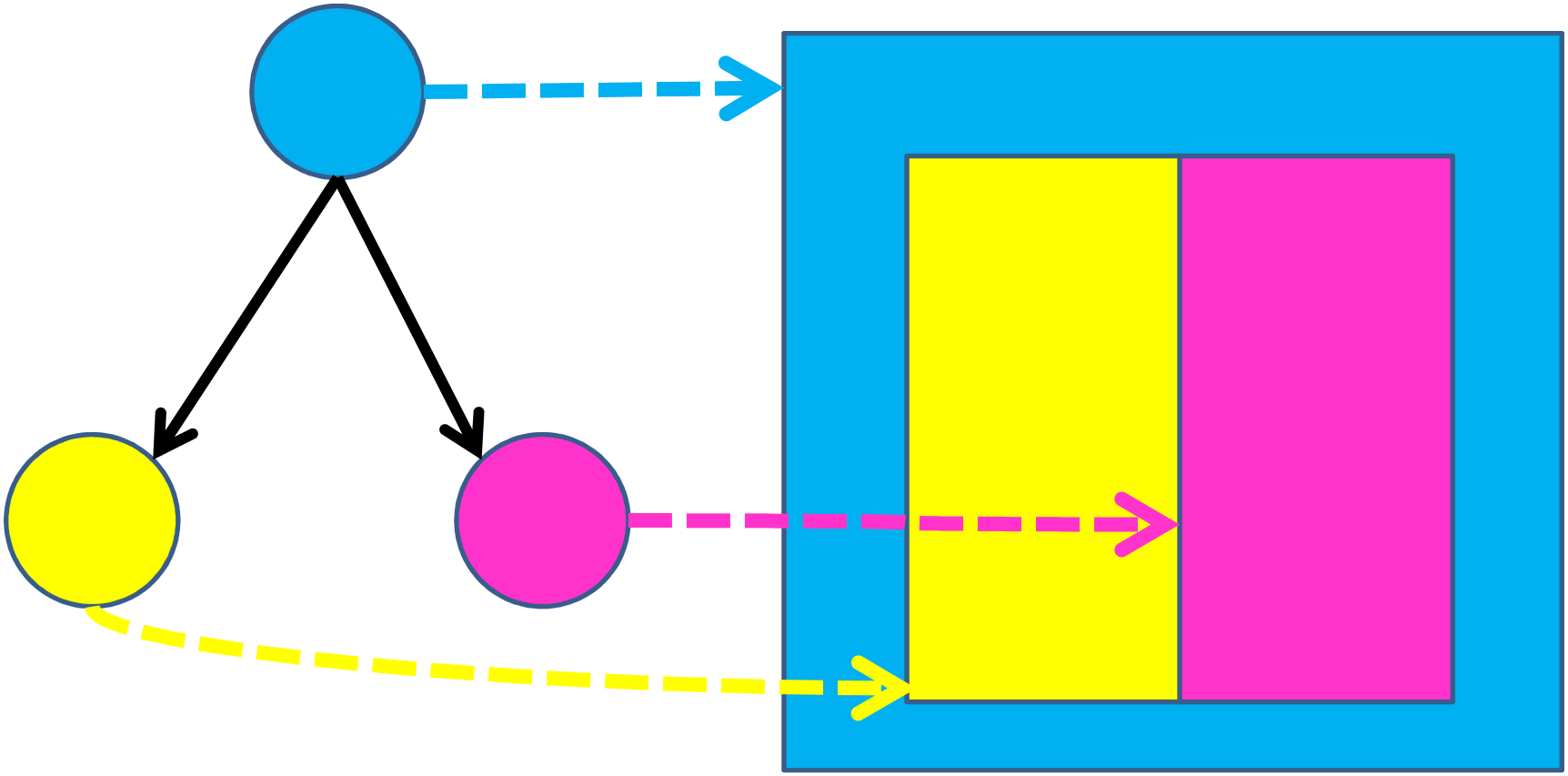
TREEMAP DRAWINGS



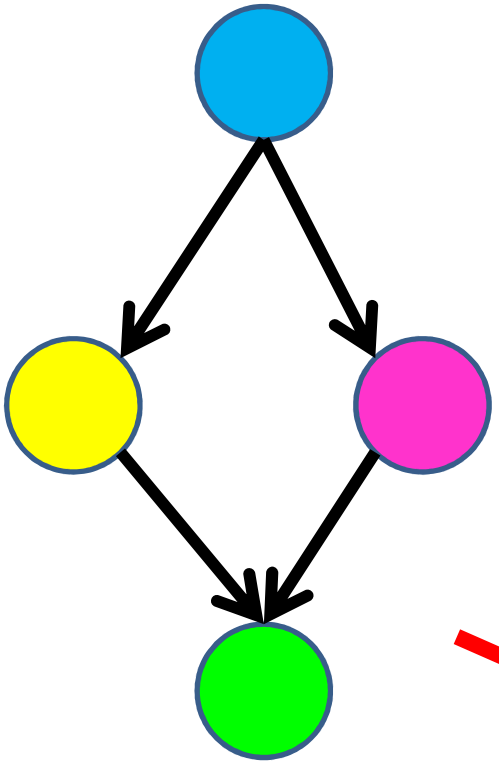
TREEMAP DRAWINGS



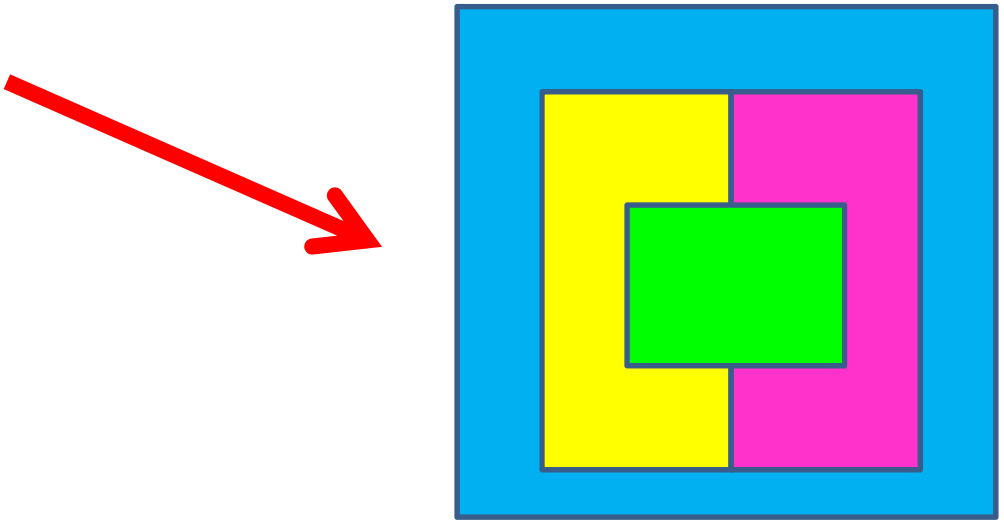
TREEMAP DRAWINGS



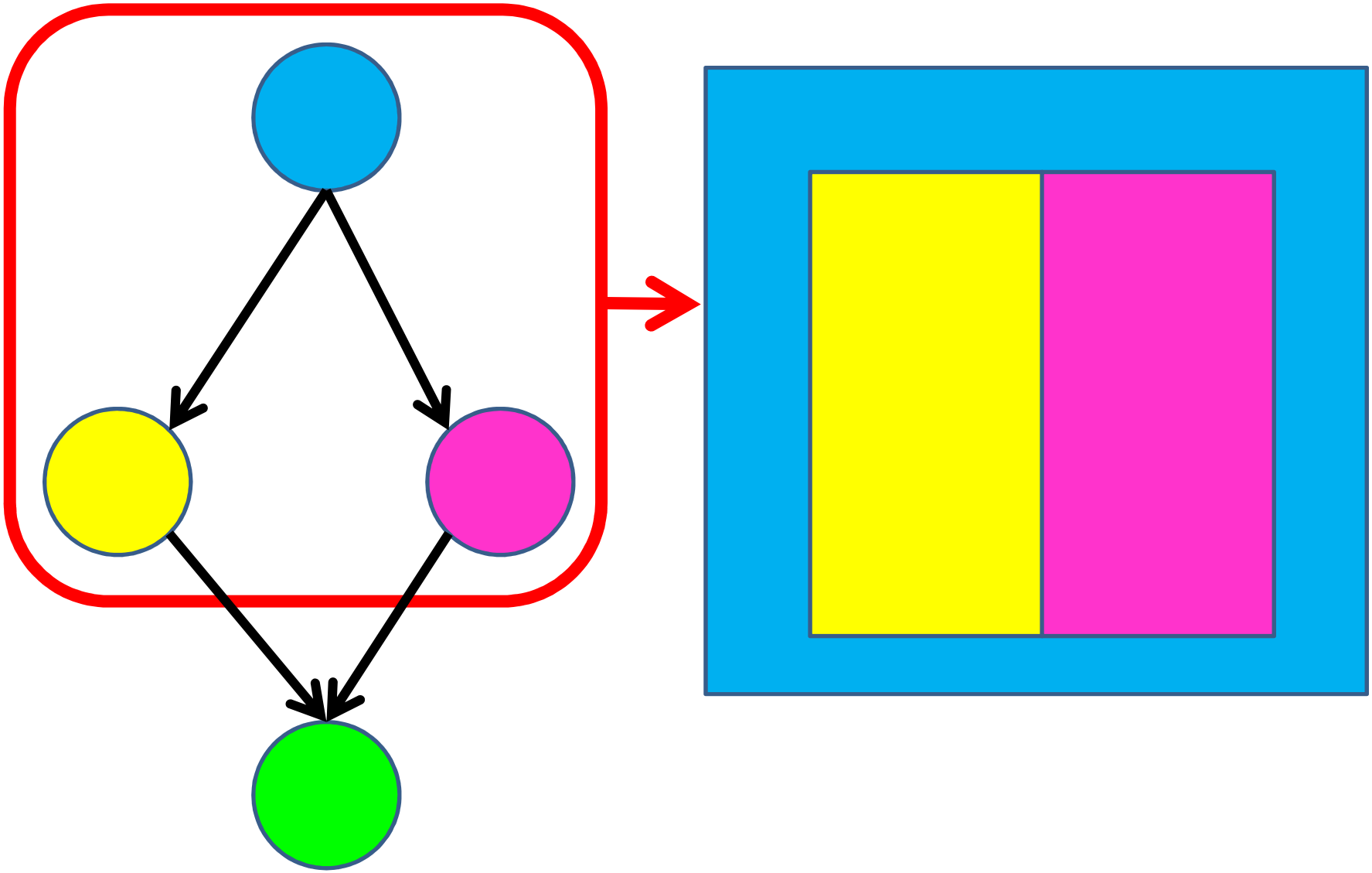
DAGMAP DRAWINGS (Tsiaras, Triantafilou, and Tollis, 2007)



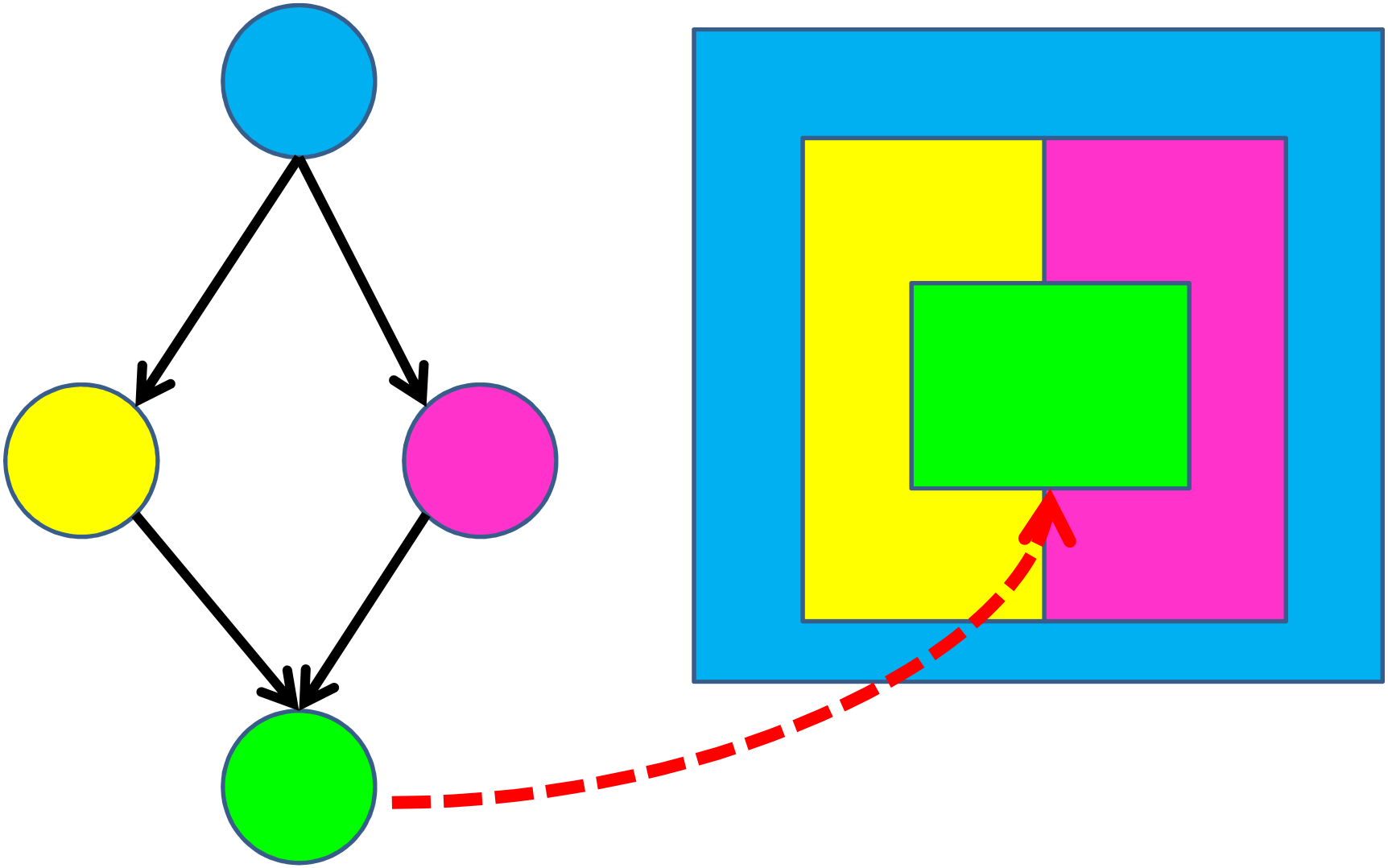
- **An extension of treemaps for visualizing Directed Acyclic Graphs (DAGs).**
- **It is not always possible to visualize a DAG with a DAGmap without having node duplications.**



DAGMAP DRAWINGS

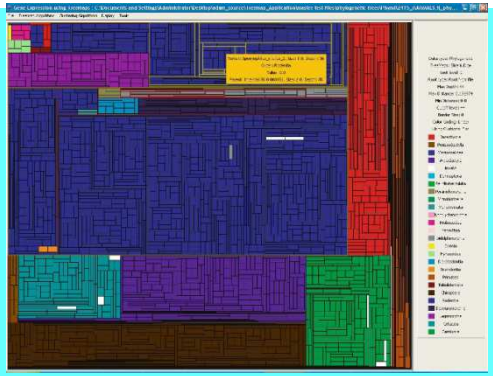


DAGMAP DRAWINGS

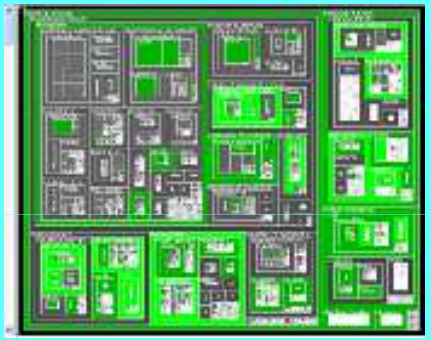


Treemaps have been used in bioinformatics to visualize:

1. Phylogenetic trees



2. Gene expression data



3. Gene ontologies

4. Encyclopedia of Life

Main Results

Drawing Galled Networks as DAGmaps, without node duplications, is NP-Complete.

<u>Linear time algorithms for:</u>

1. **Drawing Galled Trees as DAGmaps**
2. **Drawing Planar Galled Networks as DAGmaps**

Main Results

Drawing Galled Networks as DAGmaps, without node duplications, is NP-Complete.

Linear time algorithms for:

1. **Drawing Galled Trees as DAGmaps**
2. **Drawing planar galled networks as DAGmaps**

NOTE

- **Galled trees and galled networks have received much attention in recent years.**
- **They are important types of phylogenetic networks.**
- **A galled tree or network may suffice to accurately describe an evolutionary process when the number of recombination events is limited and most of them have occurred recently (Guseld, Eddhu, and Langley, 2004).**

DRAWINGS GALLED TREES AS DAGMAPS

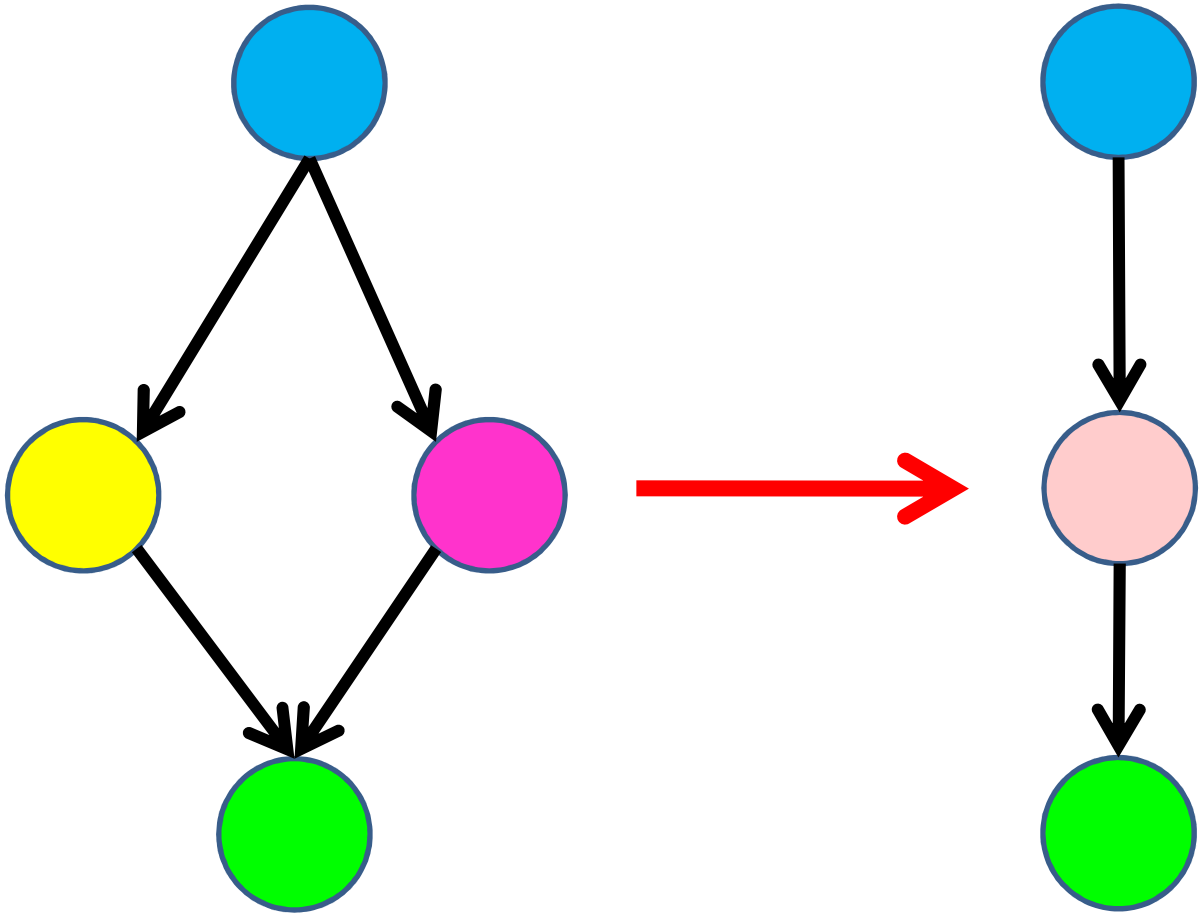
Algorithm

Input: A galled tree G.
Output: A DAGmap drawing of G.

1. **Transform the galled tree G into a tree T, by unifying the two chains of each gall.**
2. **Draw the treemap of T.**
3. **Split the rectangles, corresponding to the nodes of the unified chains of the galls, to obtain the initial parallel chains.**

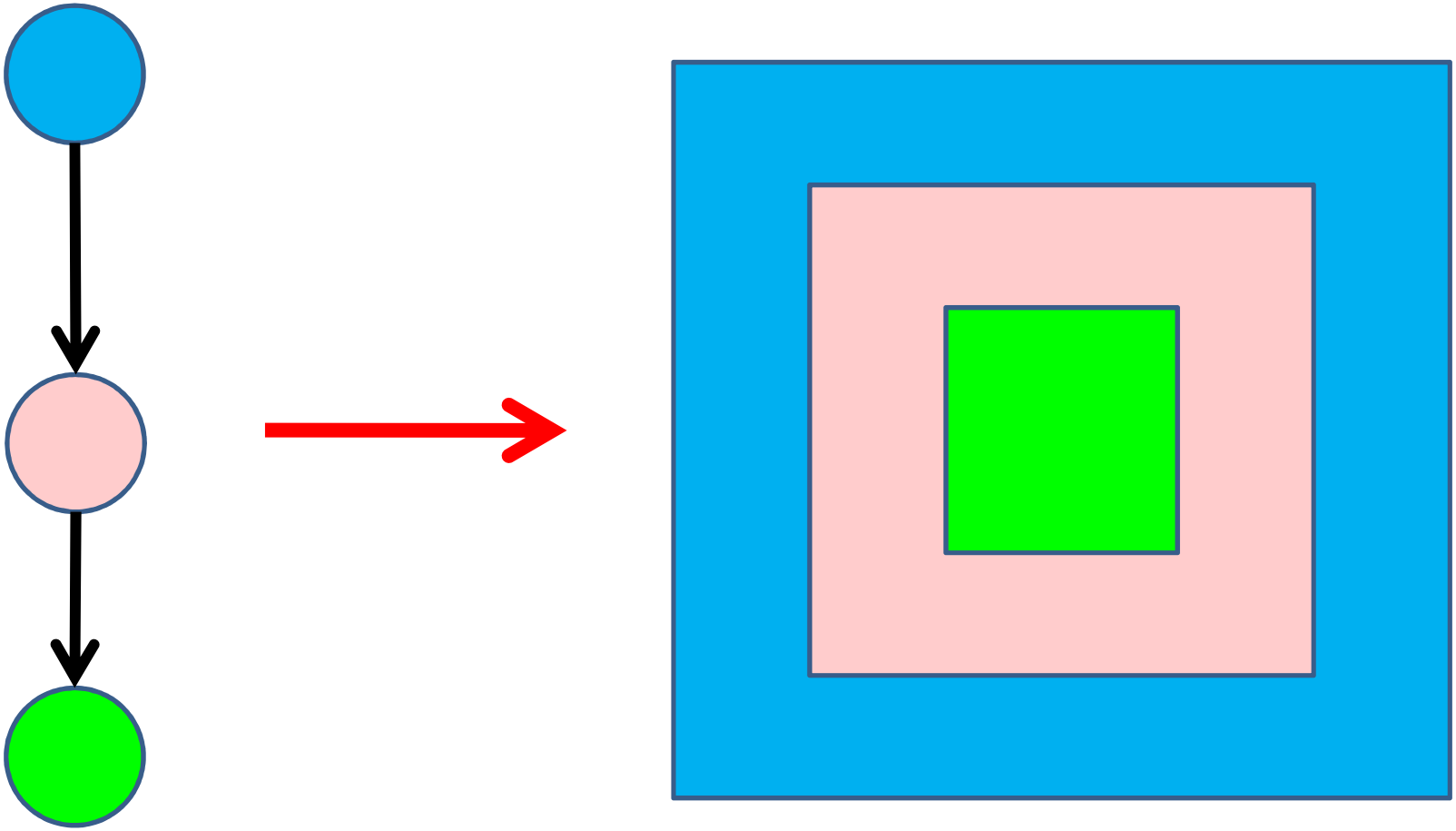
DRAWINGS GALLED TREES AS DAGMAPS

STEP 1: Transform the galled tree G into a tree T, by unifying the two chains of each gall.



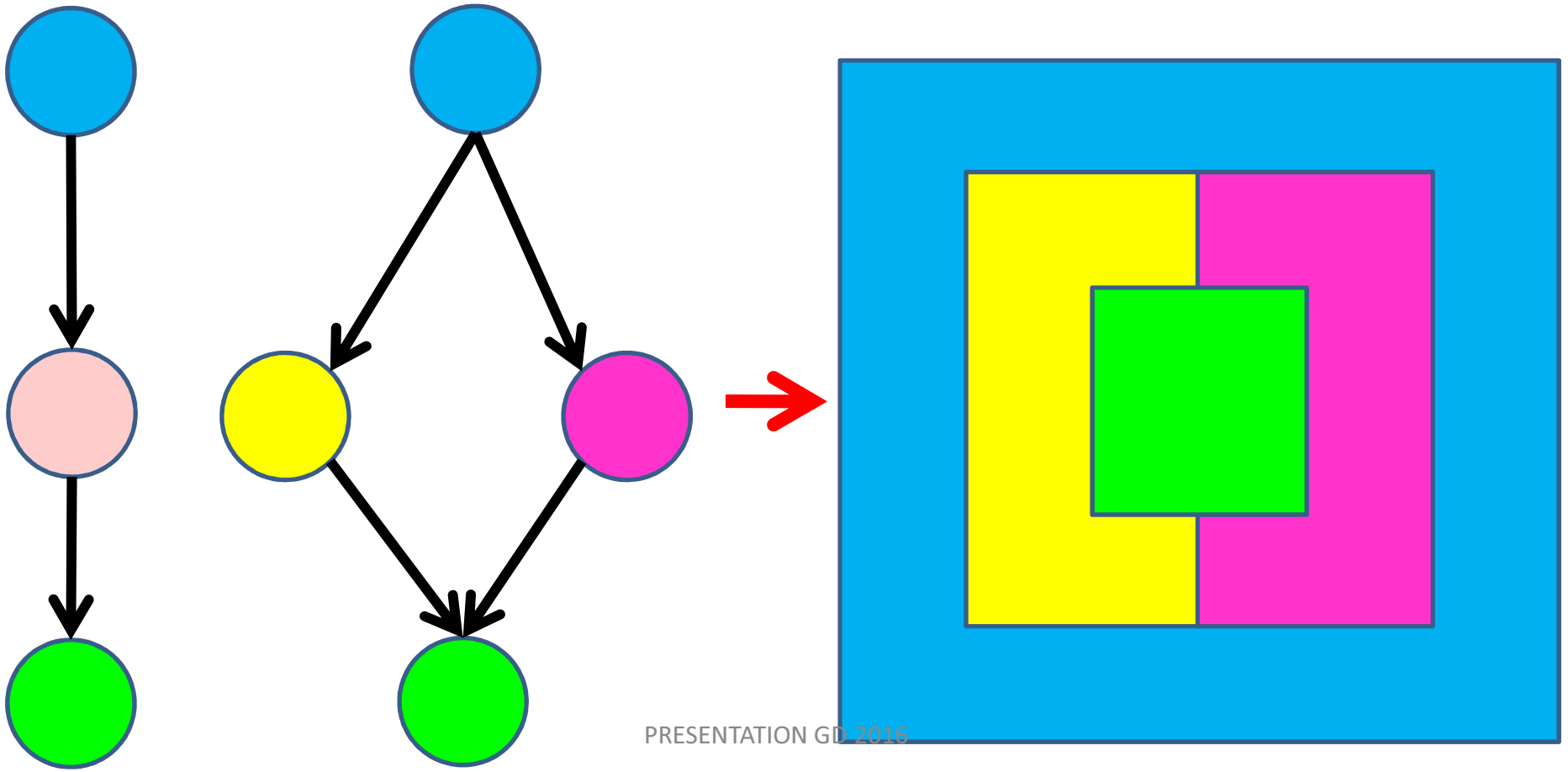
DRAWINGS GALLED TREES AS DAGMAPS

STEP 2: Draw the treemap of T.



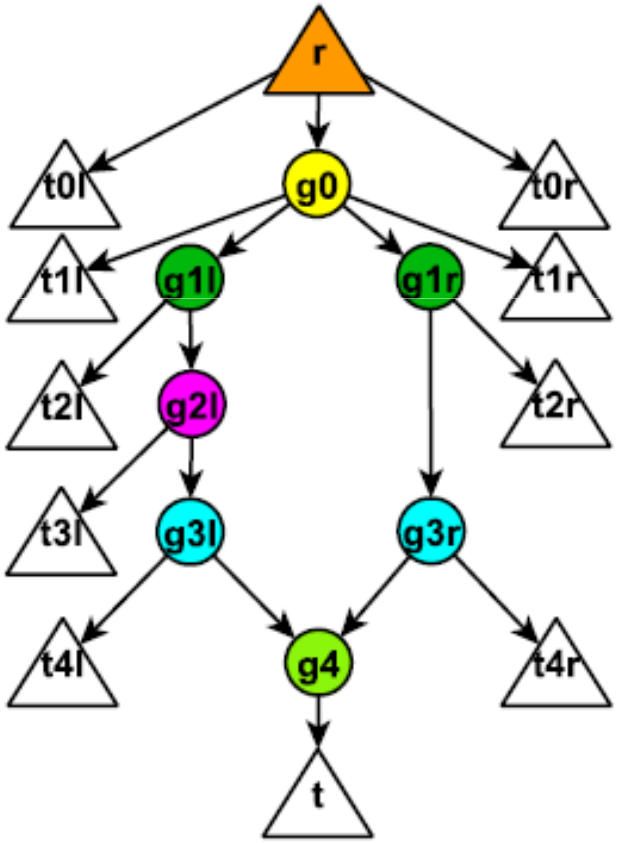
DRAWINGS GALLED TREES AS DAGMAPS

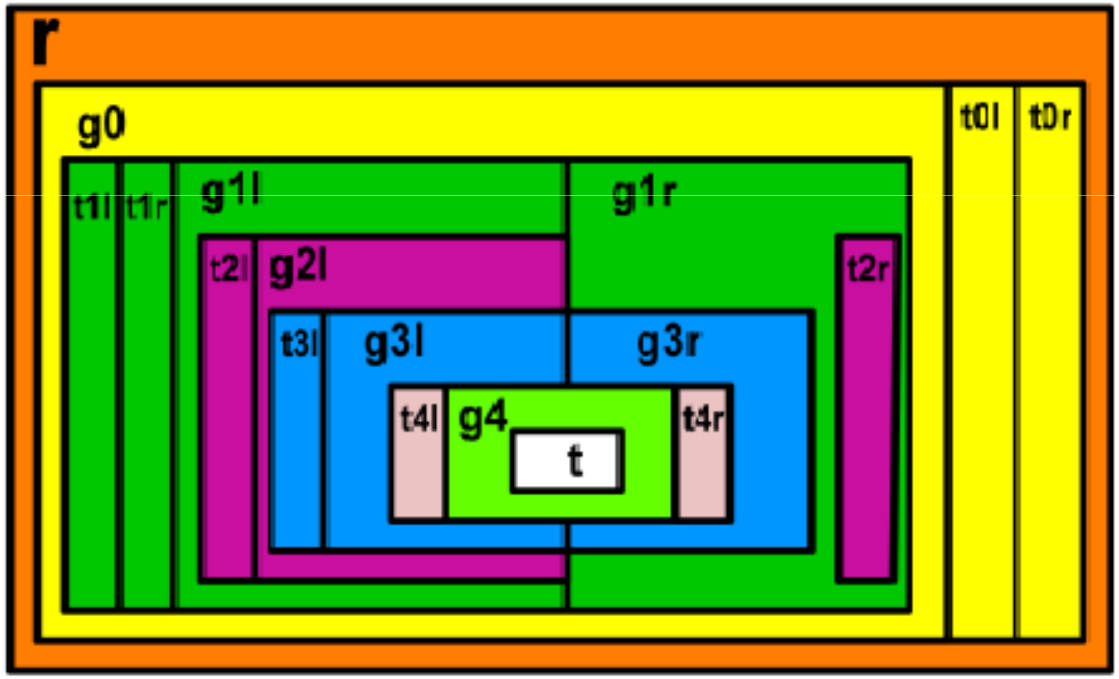
STEP 2: Split the rectangles, corresponding to the nodes of the unified chains of the galls, to obtain the initial parallel chains.



DRAWINGS GALLED TREES AS DAGMAPS

Example 2





DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

Algorithm

Input: A planar galled network G.

Output: A DAGmap drawing of G.

1. **Transform the galled network G into a galled tree GT.**

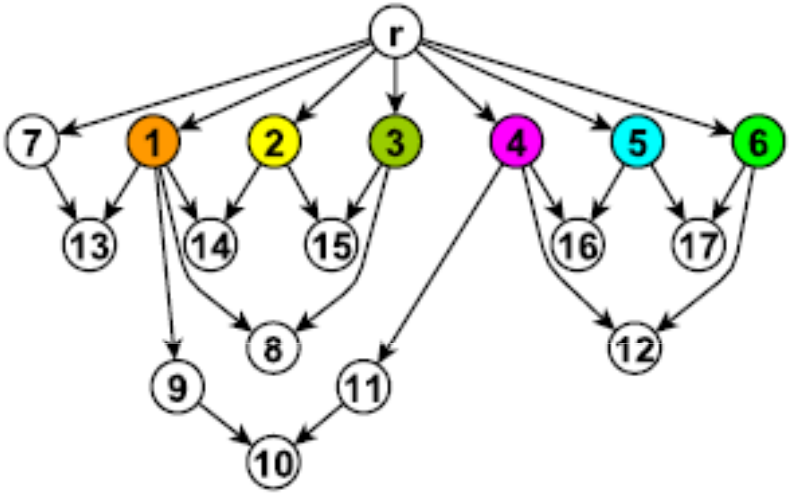
2. **Construct a planar embedding of GT.**

3. **Draw the DAGmaps of the galls of GT.**

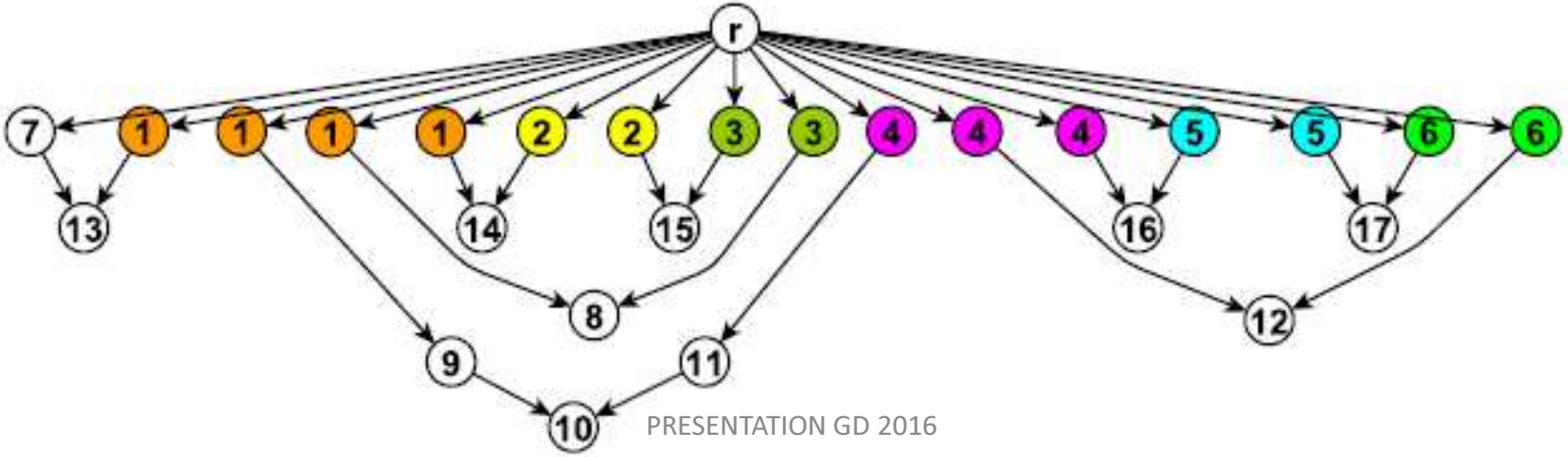
4. **Unify the split nodes and remove unused space.**

DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

STEP1: Transform the galled network G into a galled tree GT.

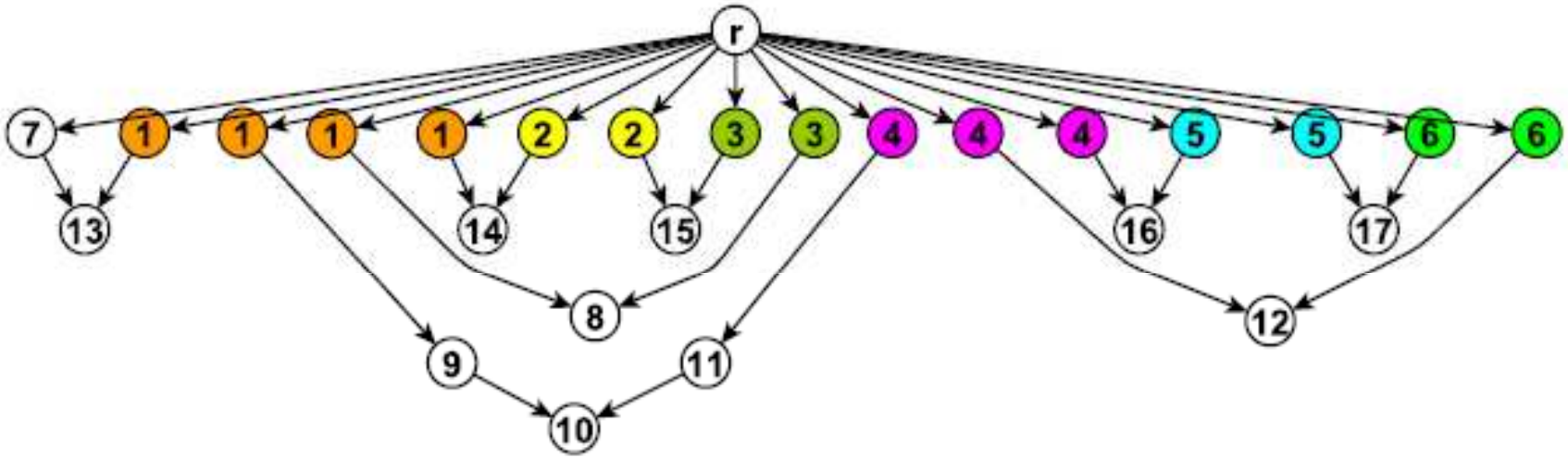


Split the nodes that belong to more than one galls.



DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

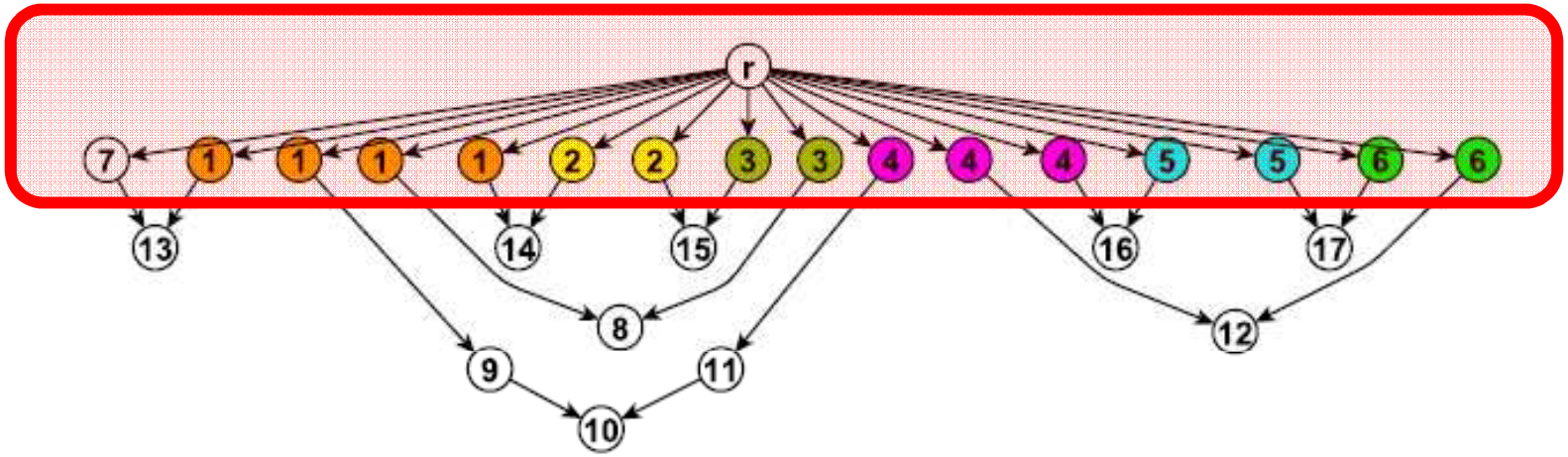
STEP 2: Construct a planar embedding of GT.



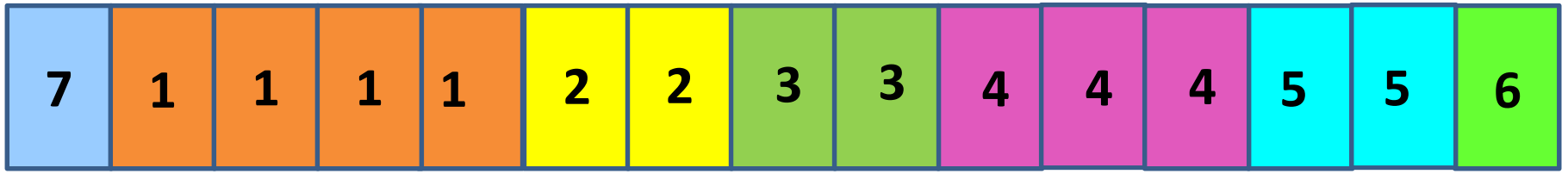
- **Each planar galled network is a single source upward planar DAG.**
- **Bertolazzi et al. (1998) have shown that a drawing of a single source upward planar DAG can be constructed in O(n) time.**
- **Thus, we can construct an upward planar drawing of a planar galled network in linear time.**

DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

STEP 3: Draw the DAGmaps of the galls of GT.

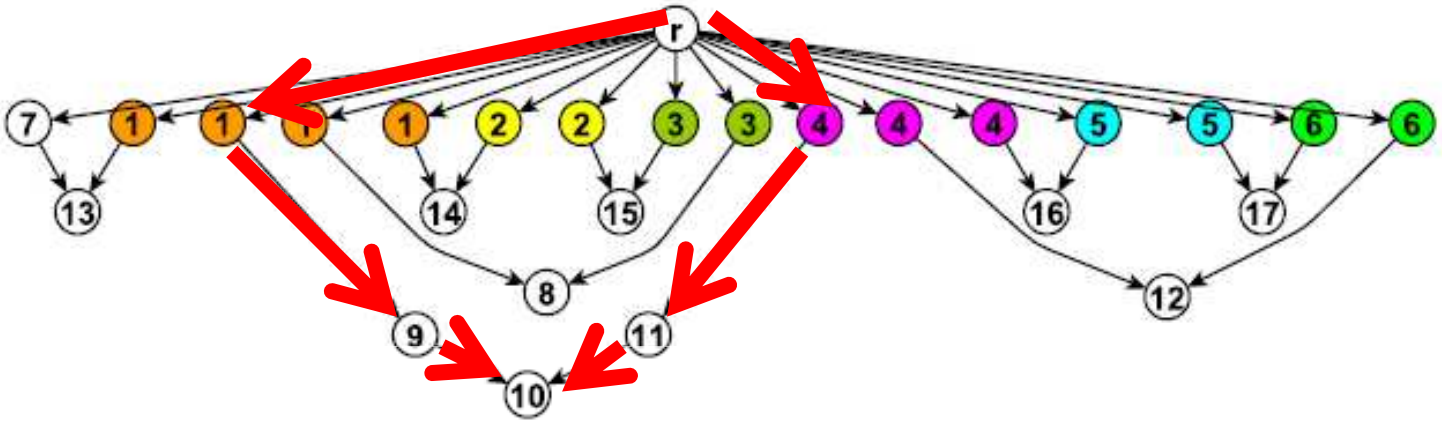


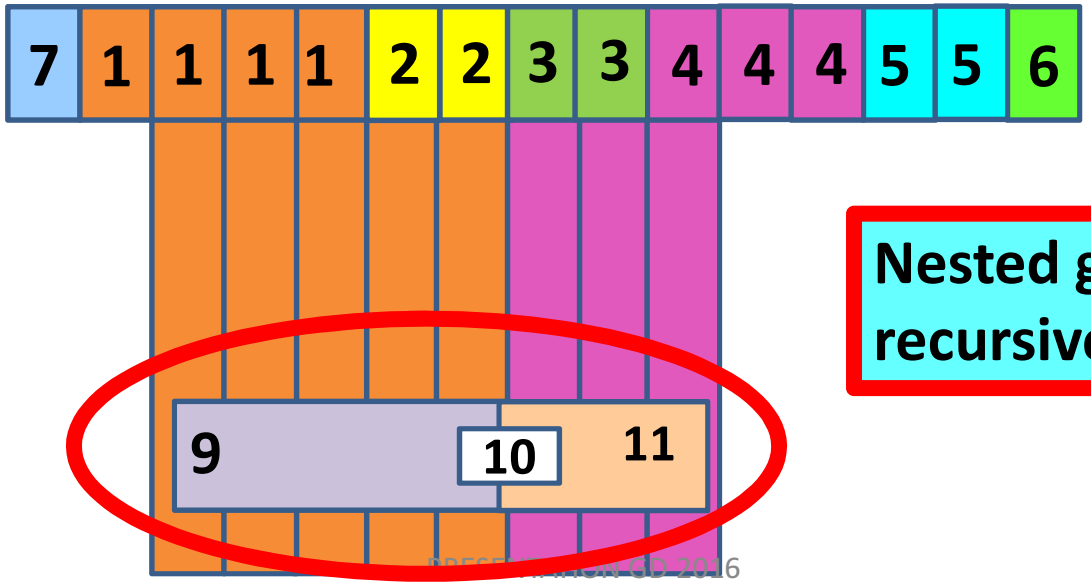
DRAW AS TREEMAP



DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

STEP 3: Draw the DAGmaps of the galls of GT.

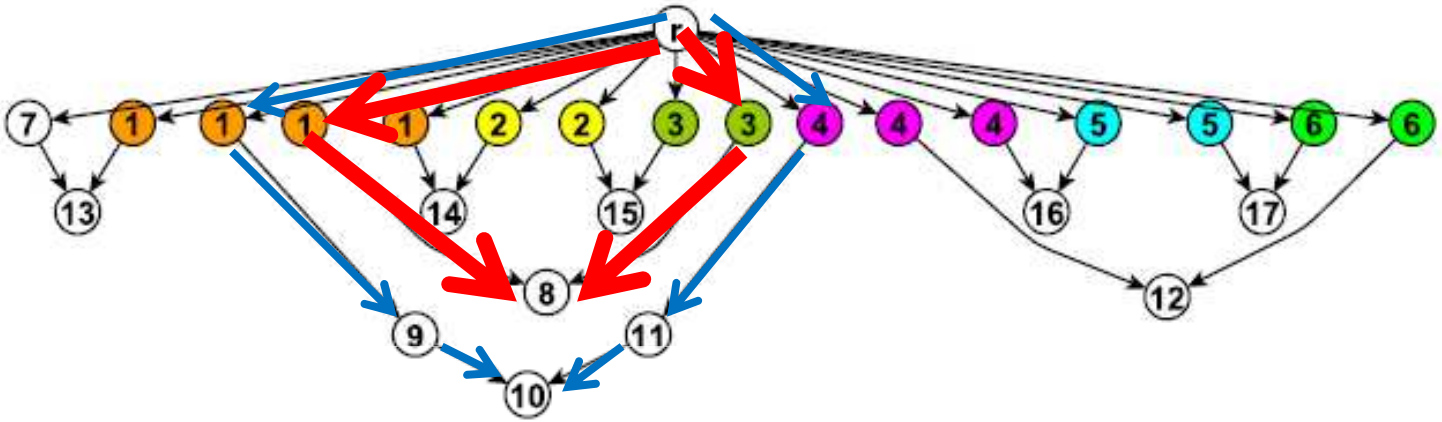


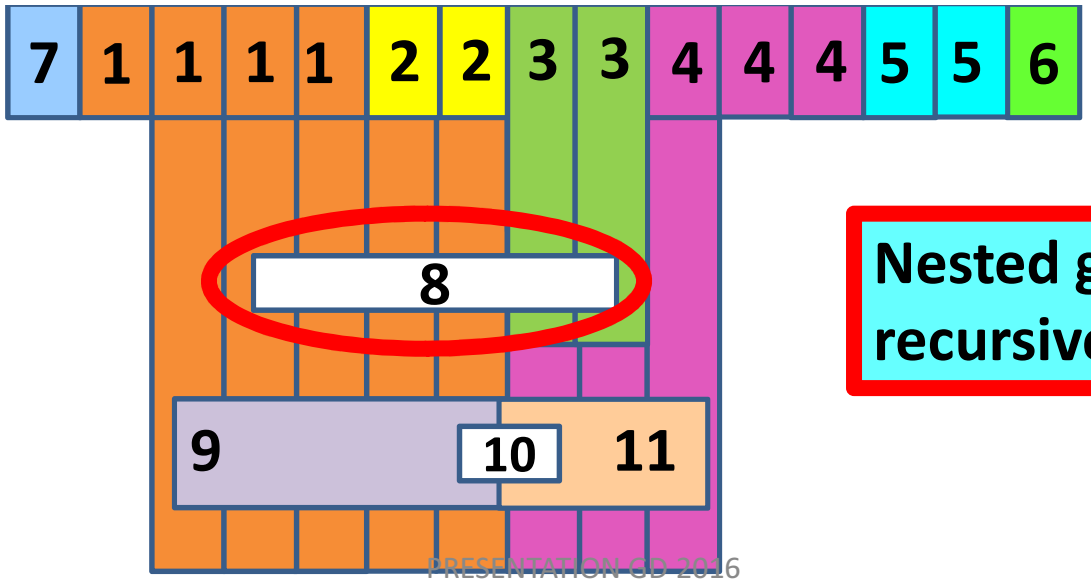


Nested galls are drawn recursively.

DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

STEP 3: Draw the DAGmaps of the galls of GT.

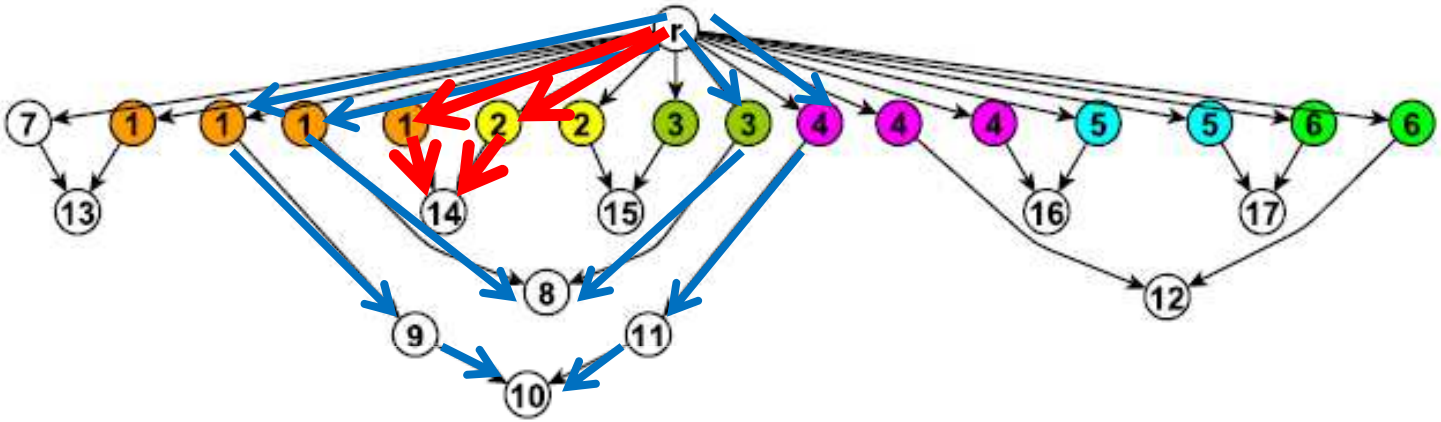


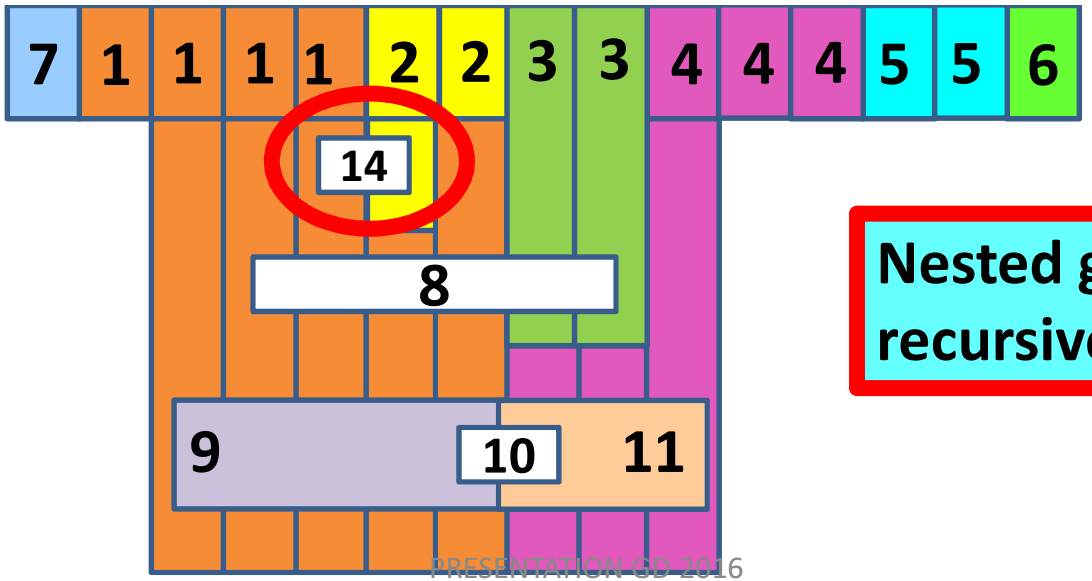


Nested galls are drawn recursively.

DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

STEP 3: Draw the DAGmaps of the galls of GT.

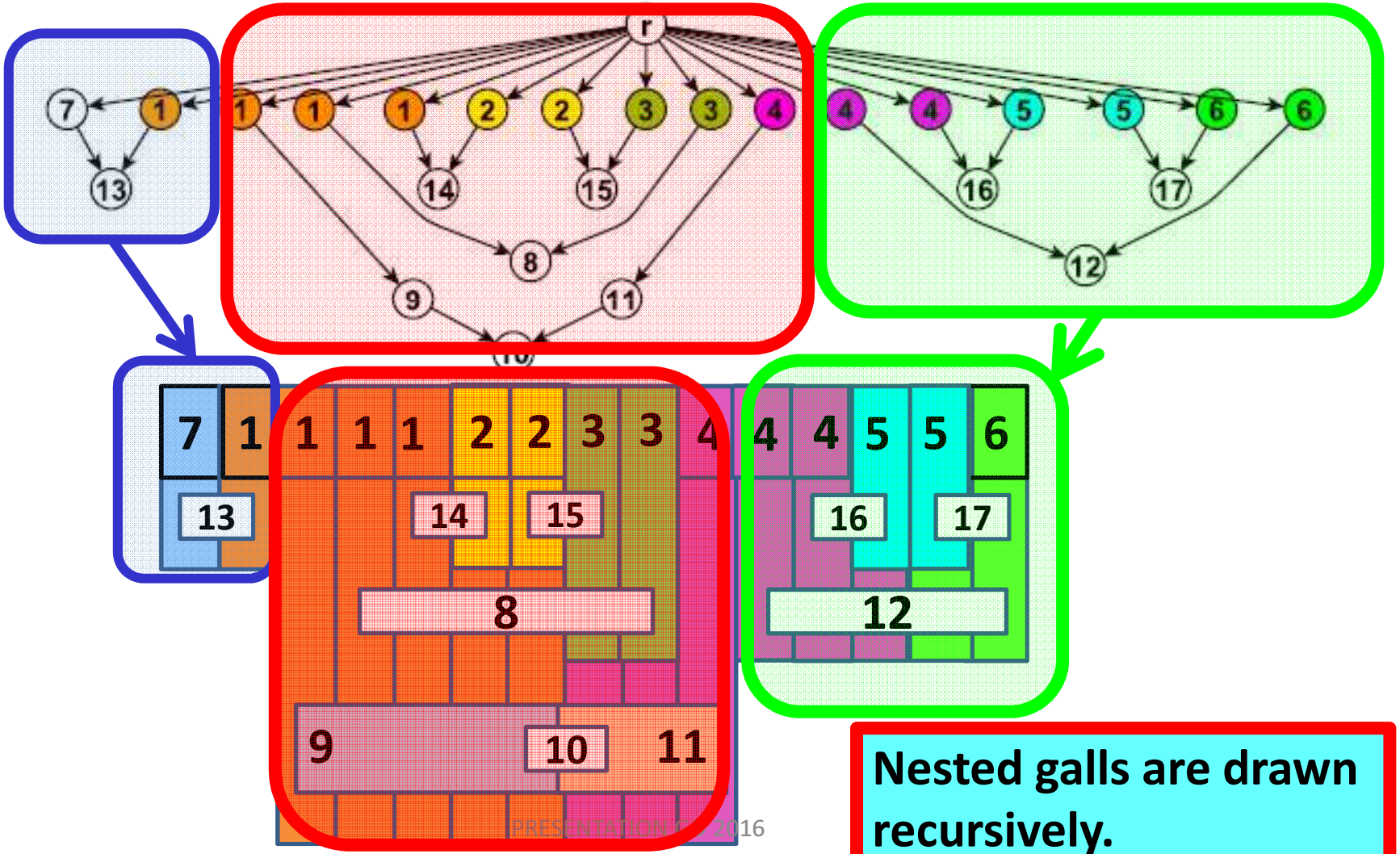




Nested galls are drawn recursively.

DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

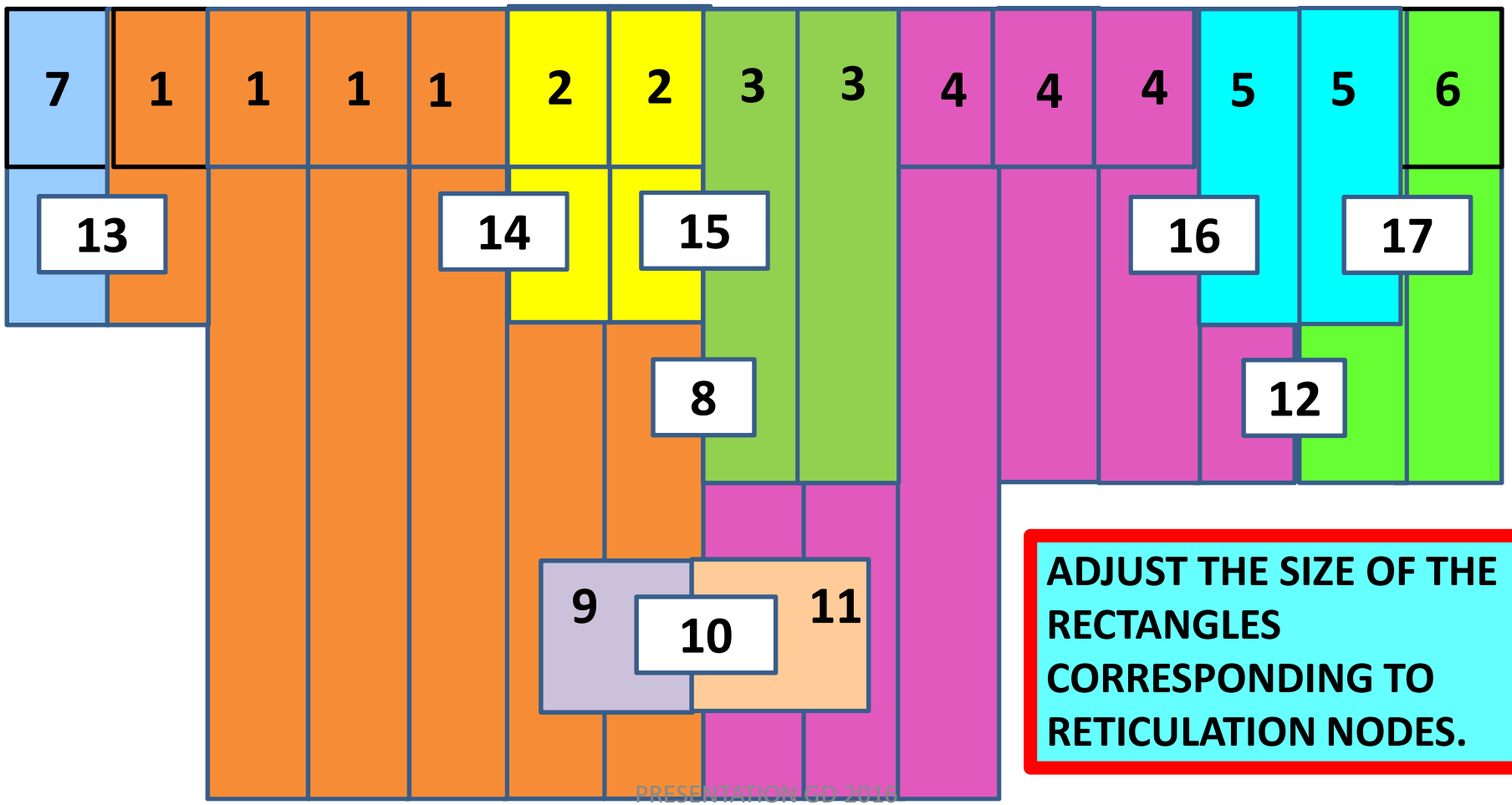
STEP 3: Draw the DAGmaps of the galls of GT.



Nested galls are drawn recursively.

DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

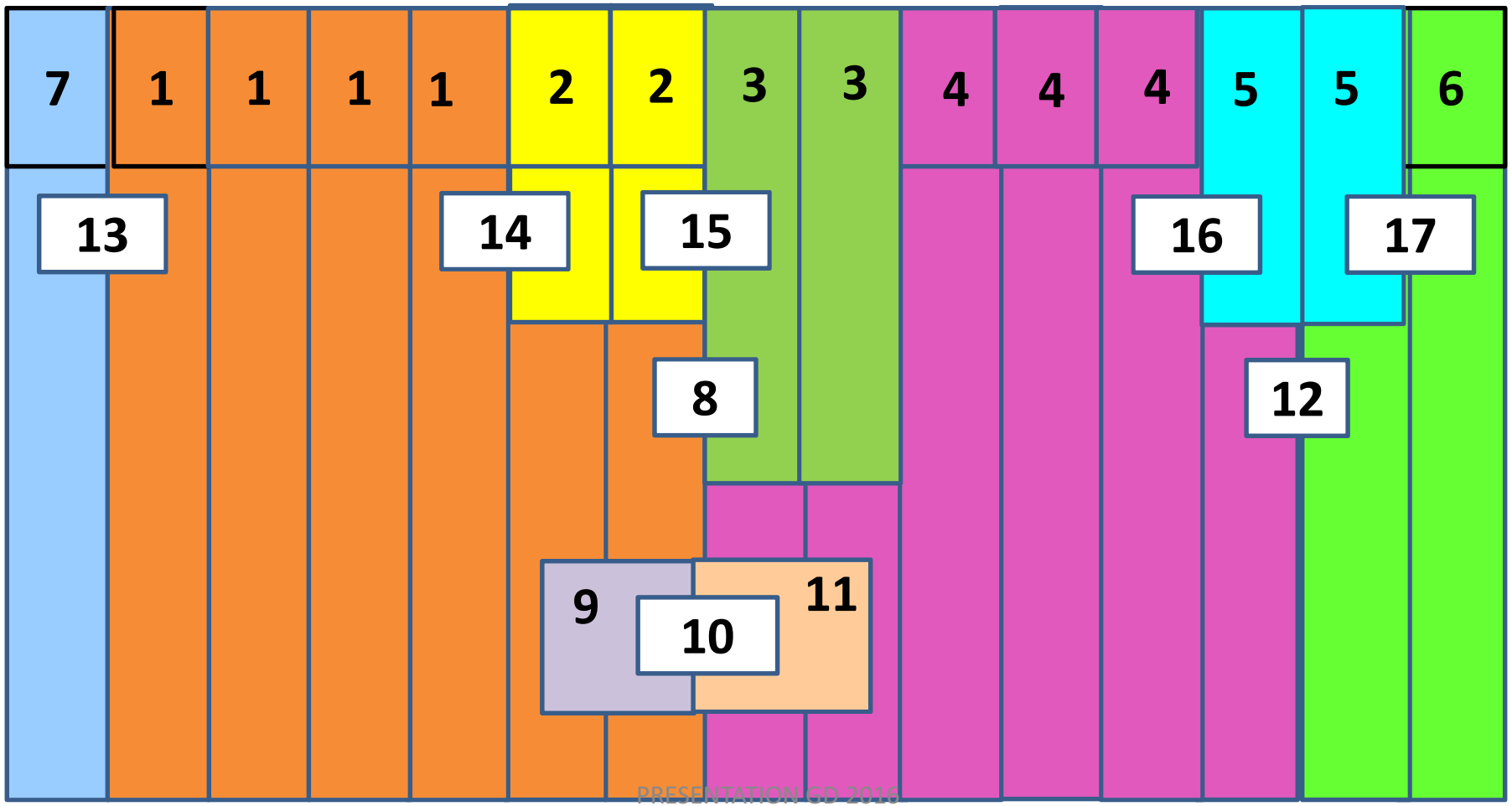
STEP 3: Draw the DAGmaps of the galls of GT.



ADJUST THE SIZE OF THE RECTANGLES CORRESPONDING TO RETICULATION NODES.

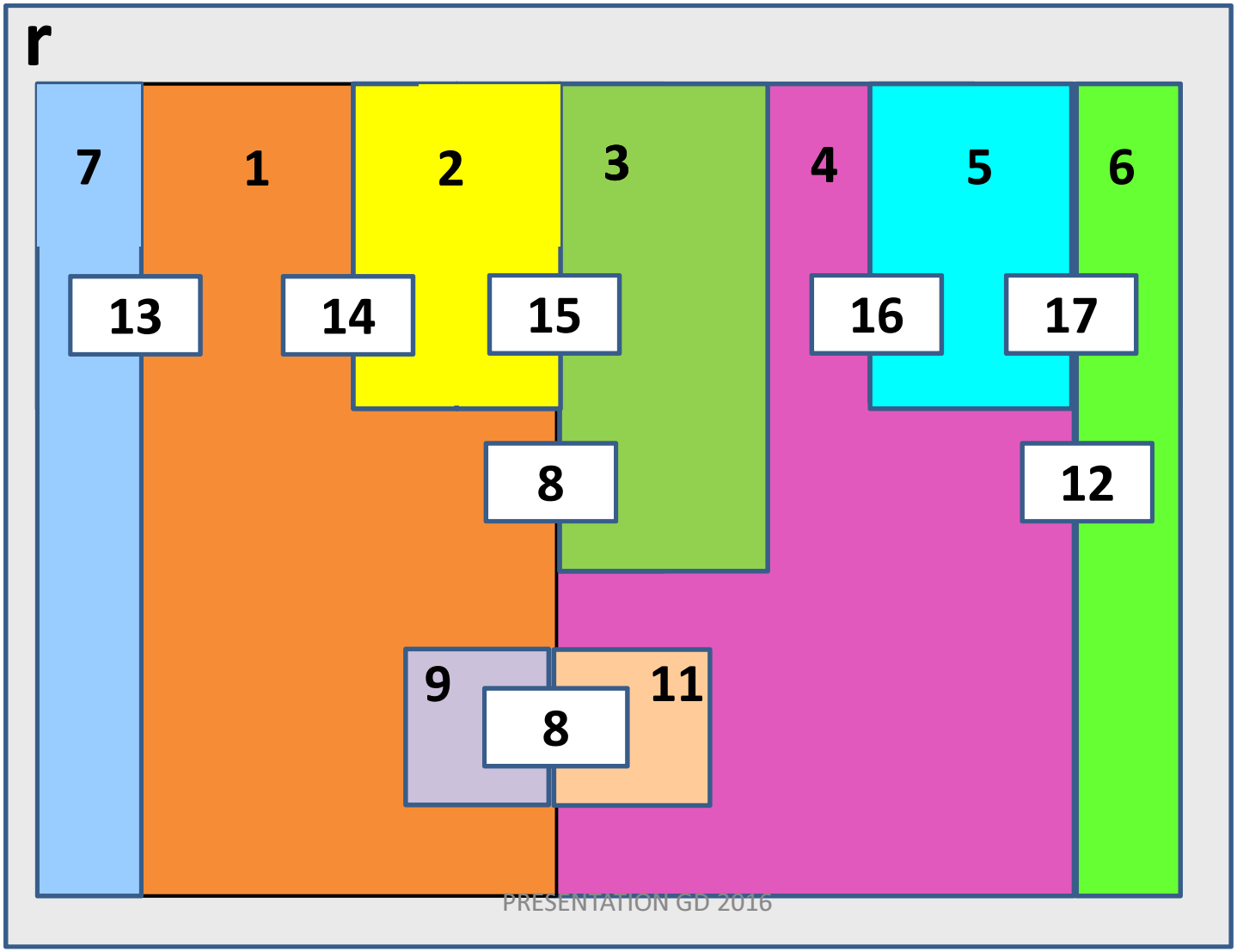
DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

STEP 3: Draw the DAGmaps of the galls of GT.



DRAWINGS PLANAR GALLED NETWORKS AS DAGMAPS

STEP 4: Unify the split nodes and remove unused space.



Future Work and Open Problems

Future Work and Open Problems

We have presented linear time algorithms for the visualization of two categories of phylogenetic networks (galled trees and planar galled networks) as DAGmap drawings.

Future Work and Open Problems

We have presented linear time algorithms for the visualization of two categories of phylogenetic networks (galled trees and planar galled networks) as DAGmap drawings.

Future Work

1. **Development of a visualization tool for processing phylogenetic networks and displaying them as DAGmaps.**

Future Work and Open Problems

We have presented linear time algorithms for the visualization of two categories of phylogenetic networks (galled trees and planar galled networks) as DAGmap drawings.

Future Work

1. Development of a visualization tool for processing phylogenetic networks and displaying them as DAGmaps.
2. **Devise DAGmap drawing algorithms for more categories of phylogenetic networks.**

Future Work and Open Problems

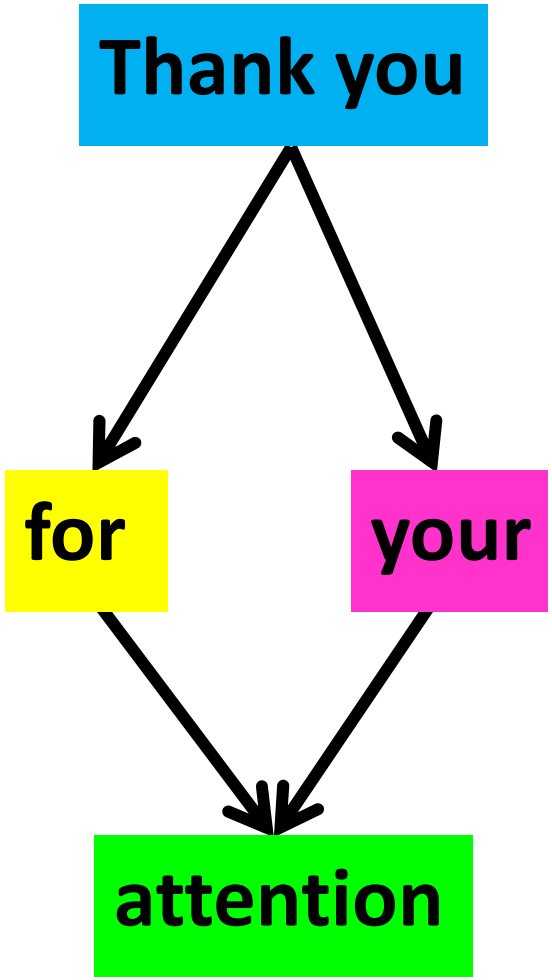
We have presented linear time algorithms for the visualization of two categories of phylogenetic networks (galled trees and planar galled networks) as DAGmap drawings.

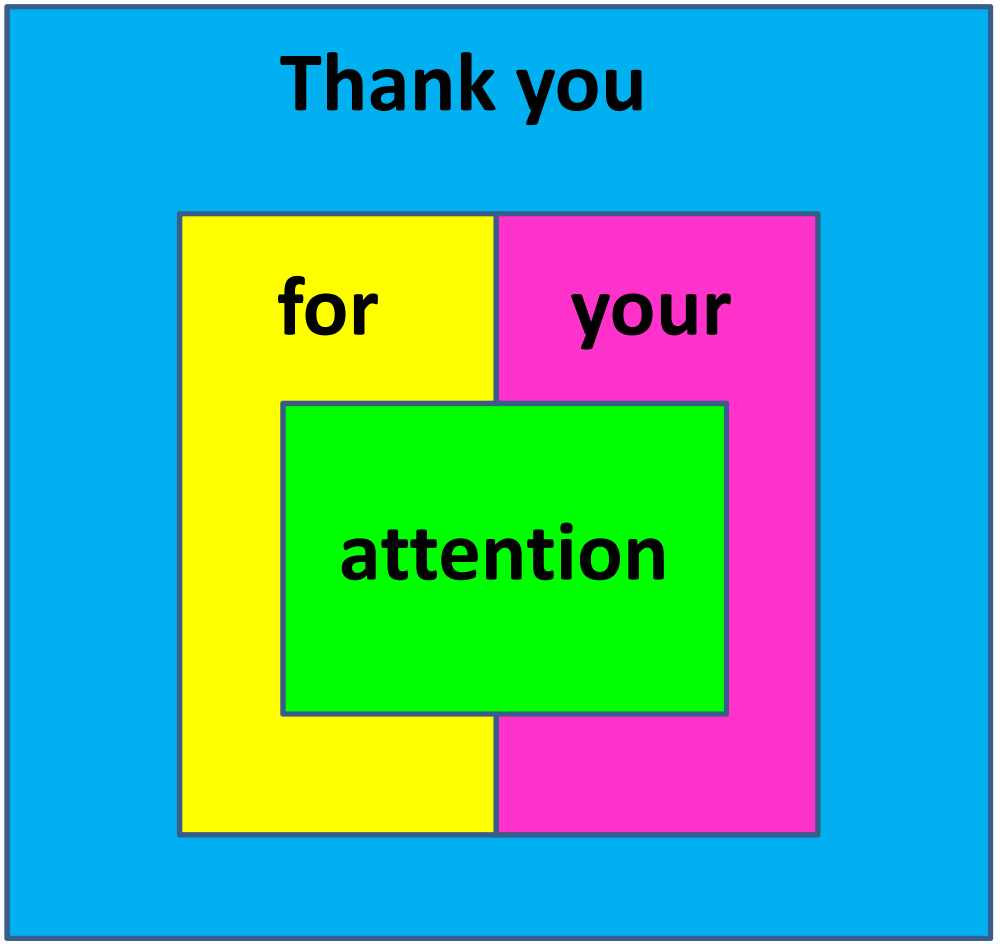
Future Work

1. Development of a visualization tool for processing phylogenetic networks and displaying them as DAGmaps.
2. Devise DAGmap drawing algorithms for more categories of phylogenetic networks.

Open Problems

- **Heuristics for reducing the number of node duplications for drawing non planar galled networks as DAGmaps.**


Thank you
for
your
attention


Thank you
for
your
attention