**Index**

**Abstract 1**

**Graph-based Inference Algorithms**

1. ***Path Finding* 1**
2. ***Centrality Detection*  2**
3. ***Community Detection*  5**

**Ontology-based Inference Algorithms**

1. ***taxonomic inference*  6**
2. ***rule-based inference.* 6**

***Citation*  7**

**Abstract:**

In knowledge graph, mainly two types of inference algorithm–graph algorithms and ontology-based algorithms–are used to retrieve information and generate new conclusion. In this document, we will discuss these two classes of algorithms in details and their advantage/disadvantage.

Graph-based Inference Algorithms

**Path Finding** - Used when 2 or more nodes in the graph satisfy certain properties.

- A\* Pathing Finding

* Advantage: direct search solution to path finding. 1 source to 1 destination.
* Disadvtage: does not iterate over all possible path therefore output is not necessarily the shortest.

1) One source and One Destination-

→ Use A\* Search Algorithm (For Unweighted as well as Weighted Graphs)

2) One Source, All Destination –

→ Use BFS (For Unweighted Graphs)

→ Use Dijkstra (For Weighted Graphs without negative weights)

→ Use Bellman Ford (For Weighted Graphs with negative weights)

3) Between every pair of nodes-

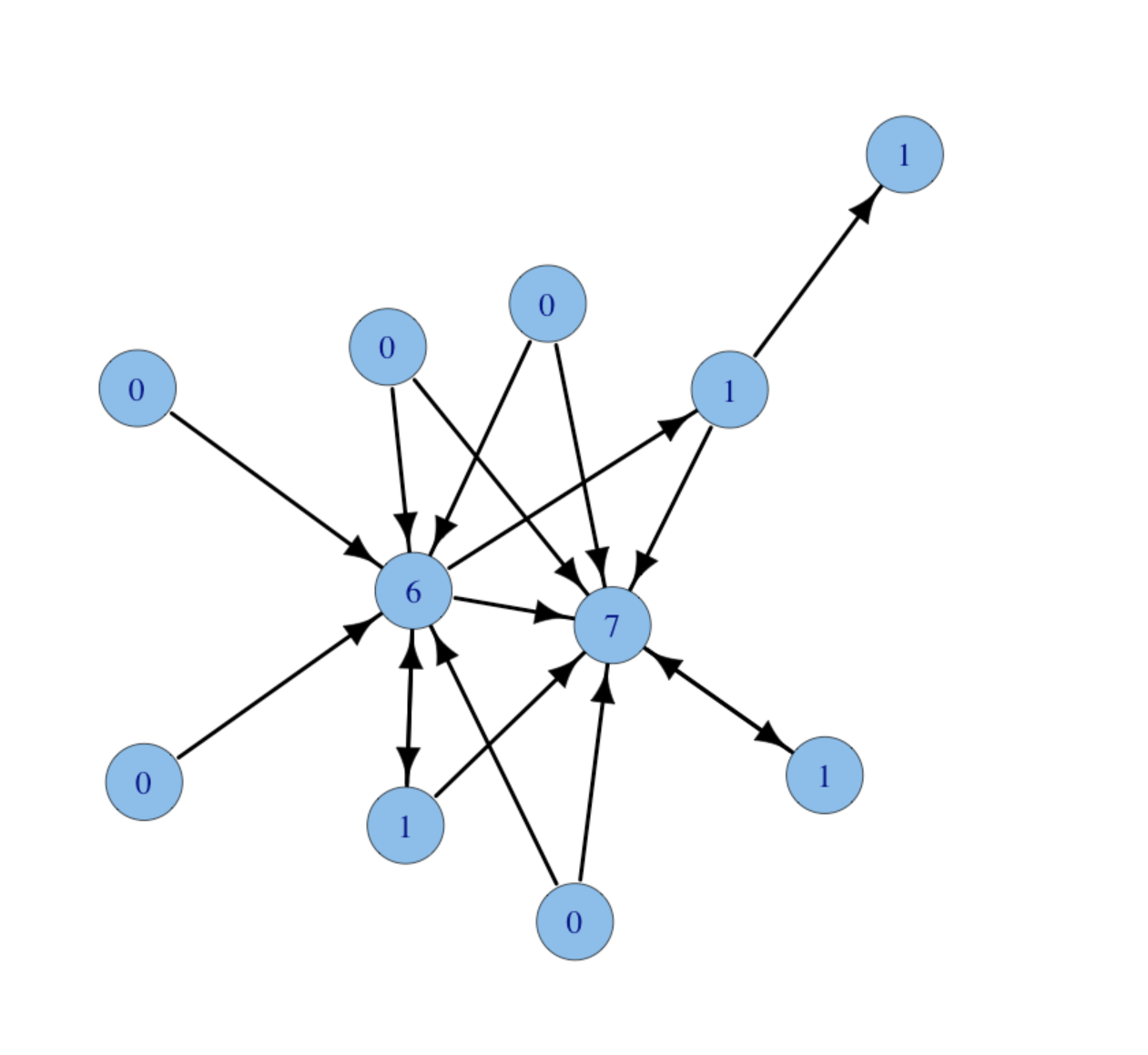
→ Floyd-Warshall

→ Johnson’s Algorithm

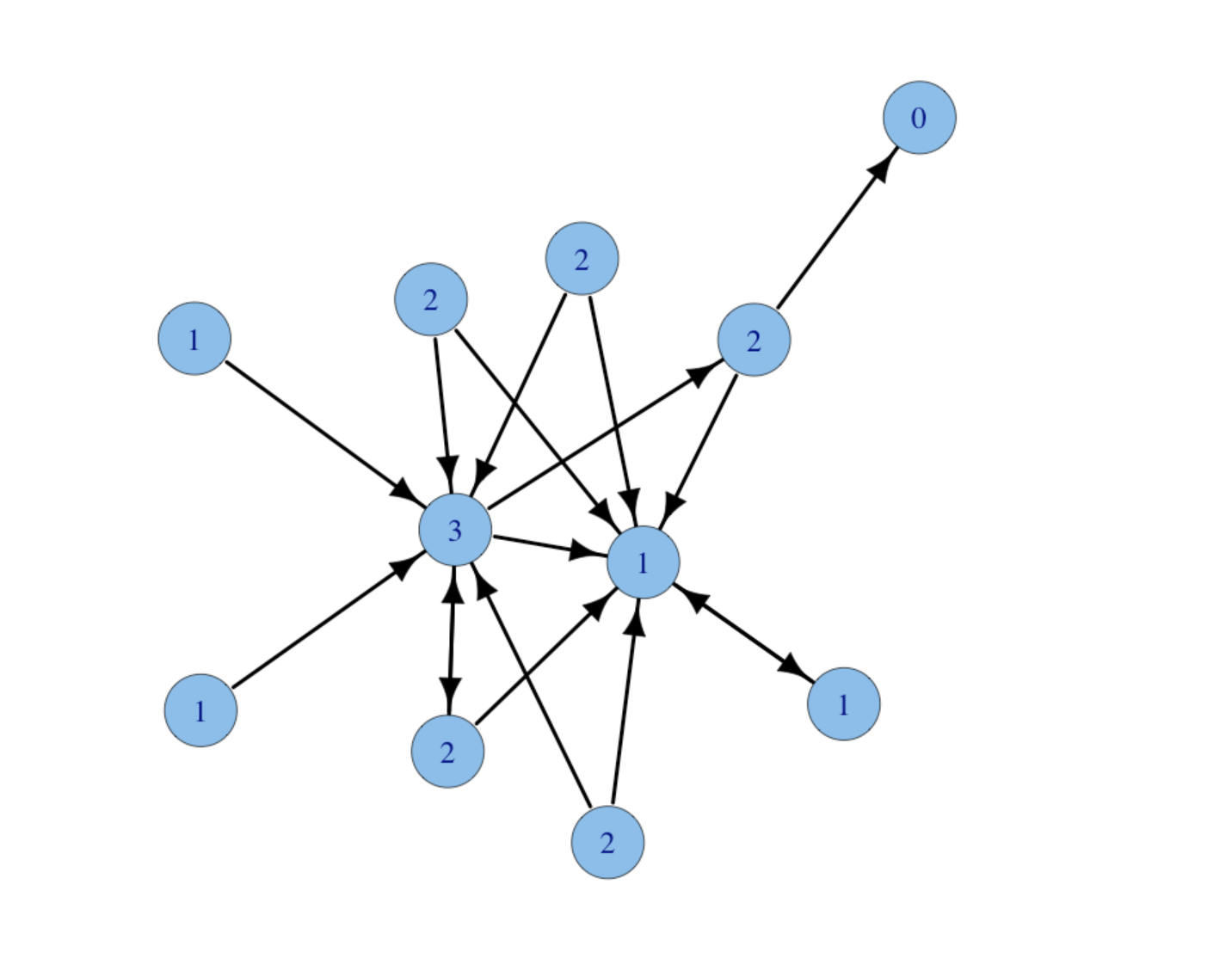
**Centrality Detection** - Used to understand importance of nodes

**Degree Centrality**

**-** simple centrality measure that counts how many neighbors a node has.

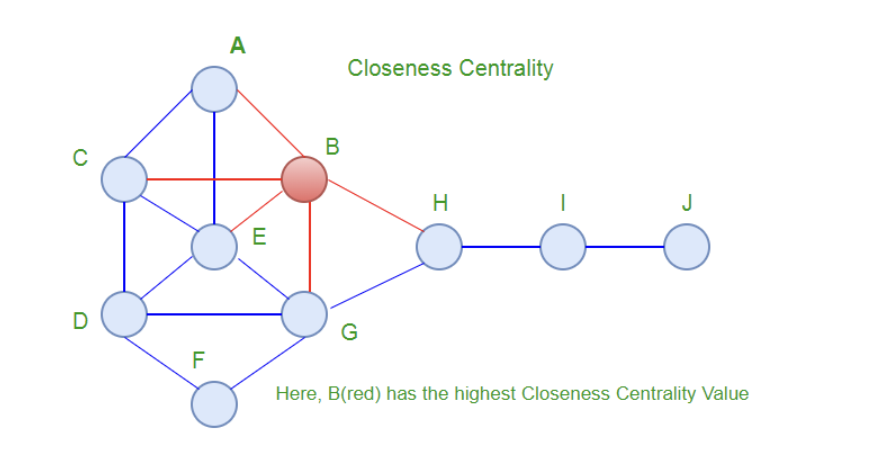
Example: A [graph](https://www.sci.unich.it/~francesc/teaching/network/img/dgraph.gml) with nodes labelled with their in-degree centrality follows:

Same graph with nodes labelled with their out-degree centrality:

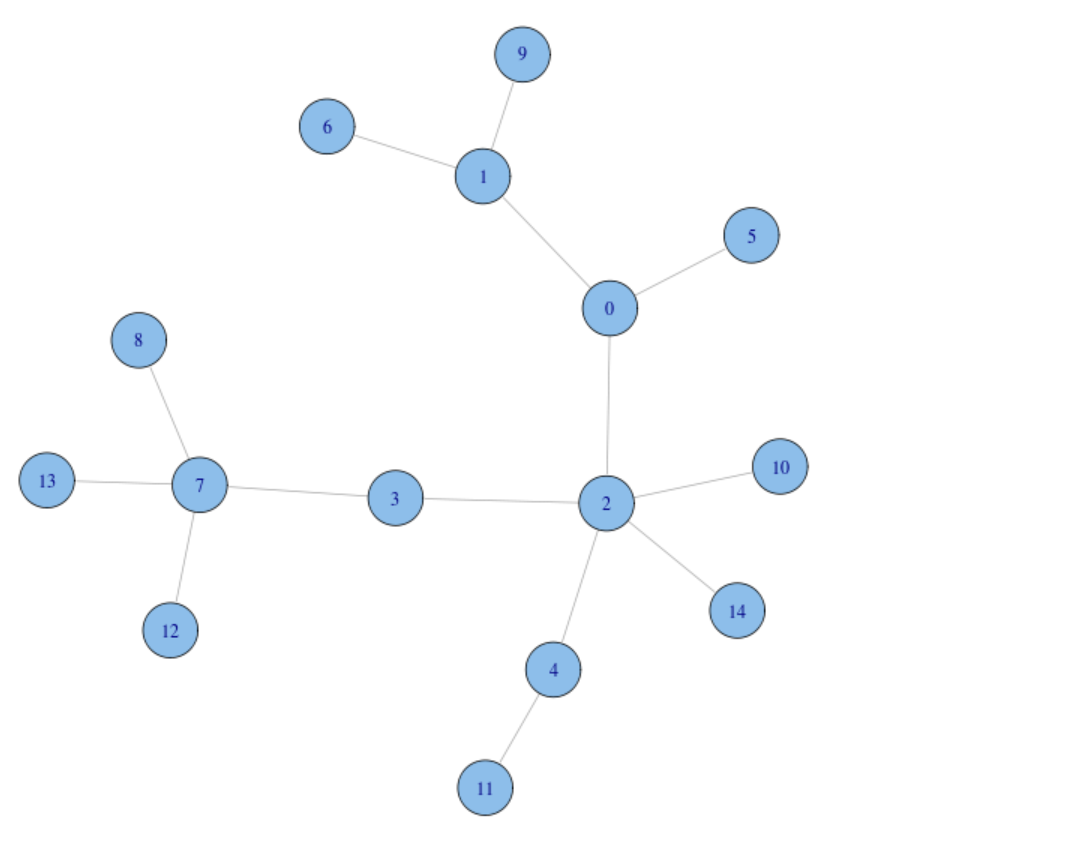
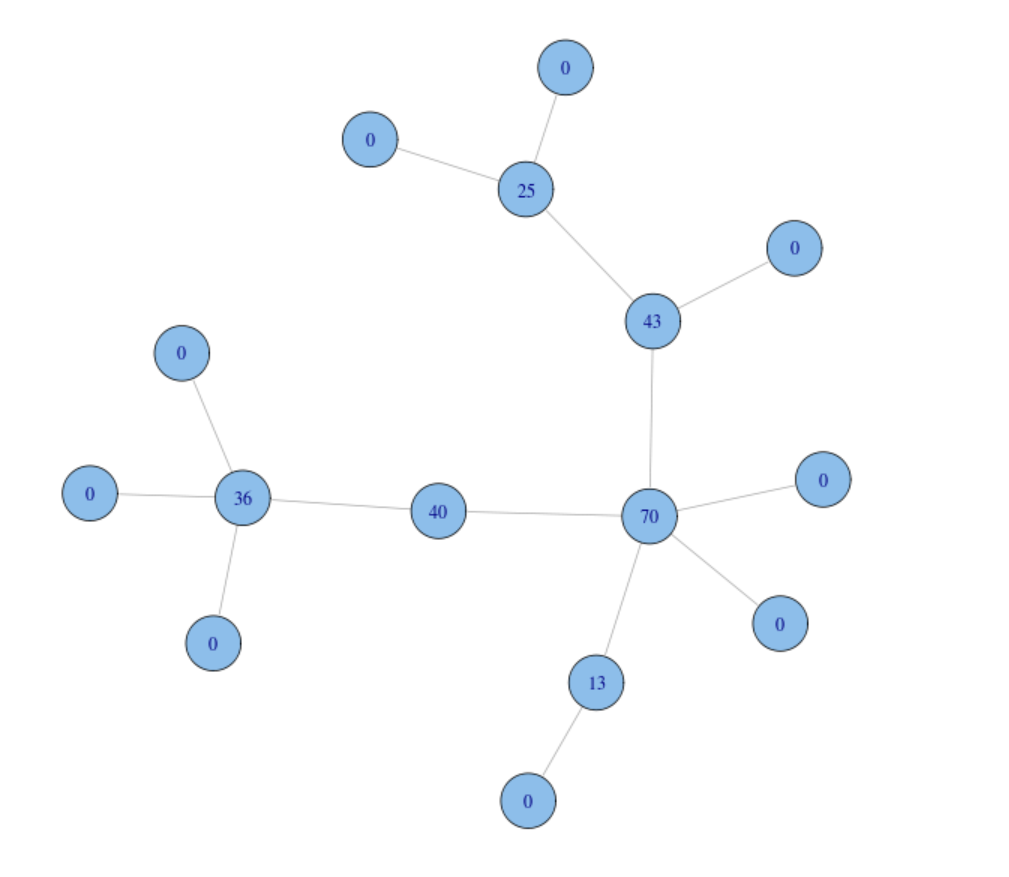


**Closeness centrality** (or closeness) of a node is a measure of centrality, calculated as the sum of the length of the shortest paths between the node and all other nodes in the graph.

Example:



**Betweenness centrality** measures the extent to which a vertex lies on paths between other vertices. Vertices with high betweenness may have considerable influence within a network by virtue of their control over information passing between others.



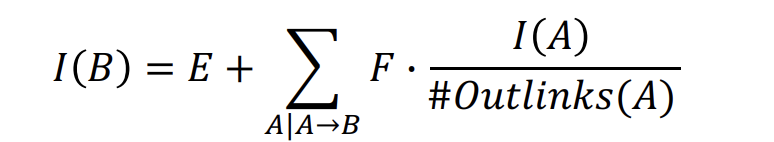
**Page Rank**

Assume that the importance of a page P given by I(P) come from two sources

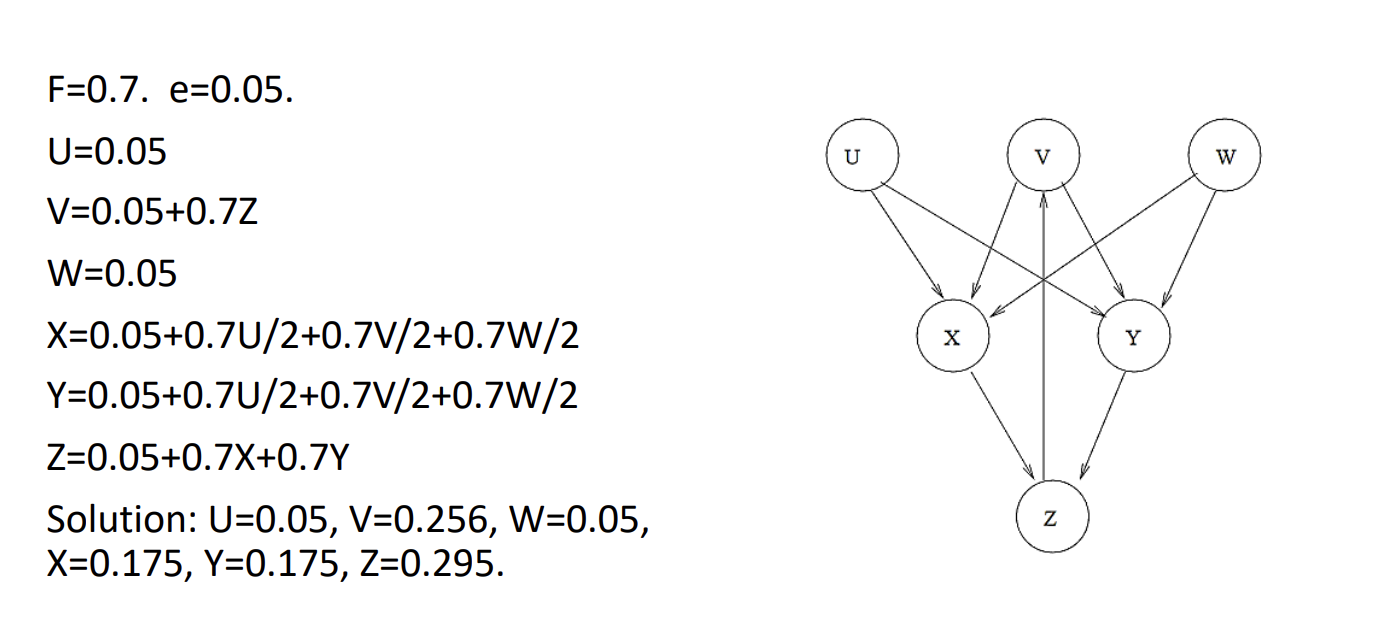
1) A page is important just by virtue of being a page. Call that E, some constant.

2) If page A links to page B, then that makes B more important. Specifically, the amount of importance that B gains is a constant F times the importance of A divided by the number of outlinks in A. (If A points to thousands of pages, they’re probably not all great. If it’s pointing to one or two, then those may have been chosen carefully.)

We can get the fomula:

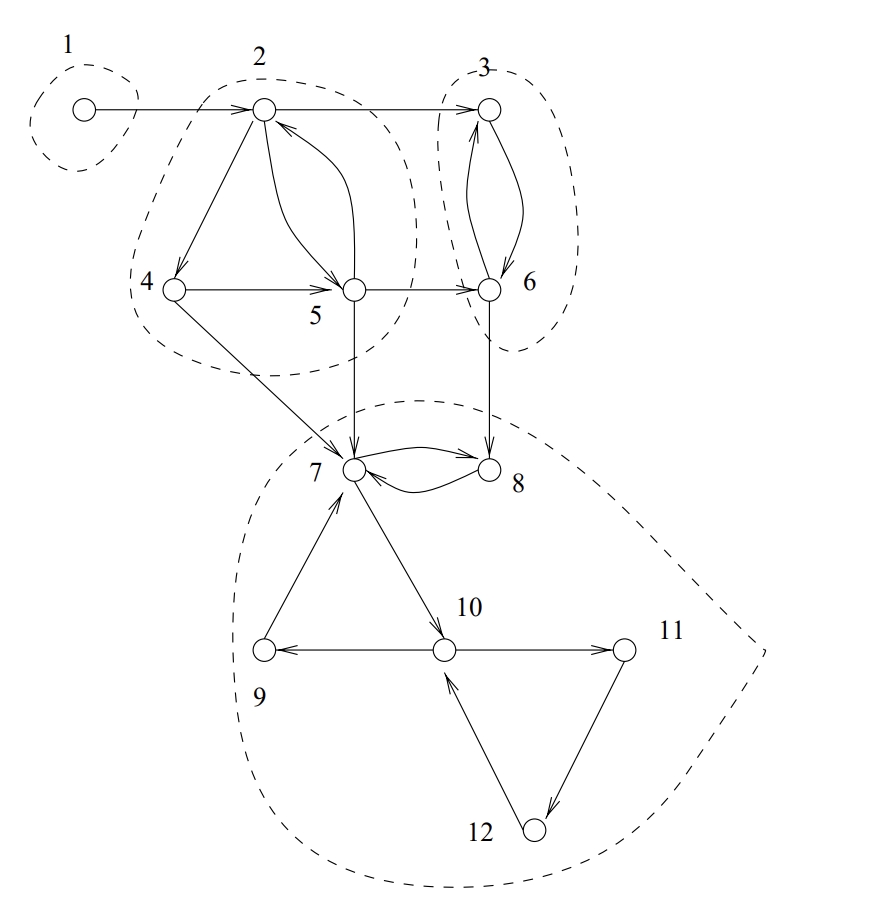


Example:



**Community Detection** - Identifying if a group of nodes shares common attributes and form community

**Connected Components/Strongly Connected Component**

* Advantage: powerful when used in inital analysis of the grap

# 

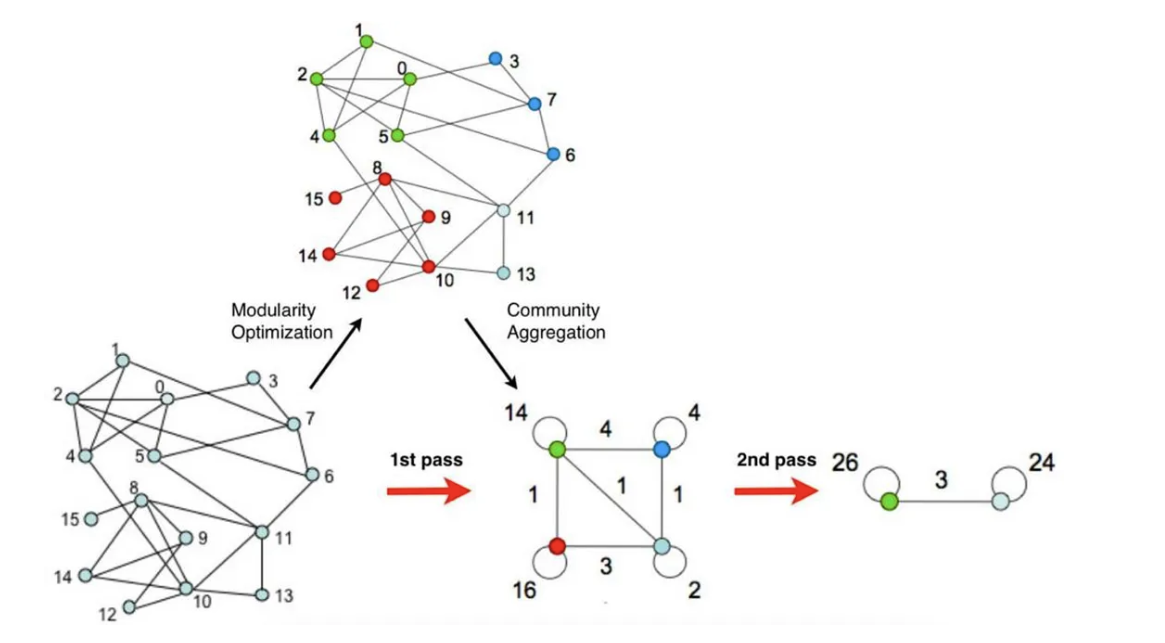
# 

# 

# 

**Louvain algorithm**

* unsupervised algorithm divided into two phrases:
  + Modularity Optimization and Community Aggregation
* Advantage: reveal emergent and potentially unanticipated communities



Example

# 

# **Ontology-based Inference Algorithms**

Ontology-based inference distinguishes a knowledge graph system from a general graph-based system. It has 2 the following categories:

* taxonomic inference

Advantage: organize knowledge into classes, which are unary relations

Good when answering these types of questions

* + Given two classes A and B, whether A is a subclass of B?
  + Given a class A and an instance I, whether I is an instance of I?
  + Given a ground relation atom determine whether it is true or false.
  + Given a relation atom, determine different values of variables for which it is true.

The first two inferences can be implemented as path finding algorithms on the graph defined by the classes and their instances. The last two inferences are equivalent to the view on the relation atom of interest.

* rule-based inference
  + bottom-up reasoning
    - Advantages: can proceed using traditional query processing methods after computation
  + top down query processing
    - Advantage: less computational space

**Citation**

Linear Algebra and Probability for Computer Science Applications , CRC Press, 2012.

https://iopscience.iop.org/article/10.1088/1742-5468/2008/10/P10008/pdf

https://people.eecs.berkeley.edu/~vazirani/s99cs170/notes/lec12.pdf

https://www.sci.unich.it/~francesc/teaching/network/degree.html

https://web.stanford.edu/~vinayc/kg/notes/What\_Are\_Some\_Inference\_Algorithms.html

http://theory.stanford.edu/~amitp/GameProgramming/

https://www.geeksforgeeks.org/a-search-algorithm/

https://www.geeksforgeeks.org/closeness-centrality-centrality-measure/