# Extract Directed Network

Description

This algorithm accepts tabular data and constructs a directed network from entities in a given source column to entities in a given target column.

For instance, the table:

People	Friends
Albert	Bob, Cathy
Bob	Bob, Cathy
Cathy, Zeus	Albert

Produces the following directed network (when 'People' is the source column, 'Friends' is the target column, and the separator is ",")
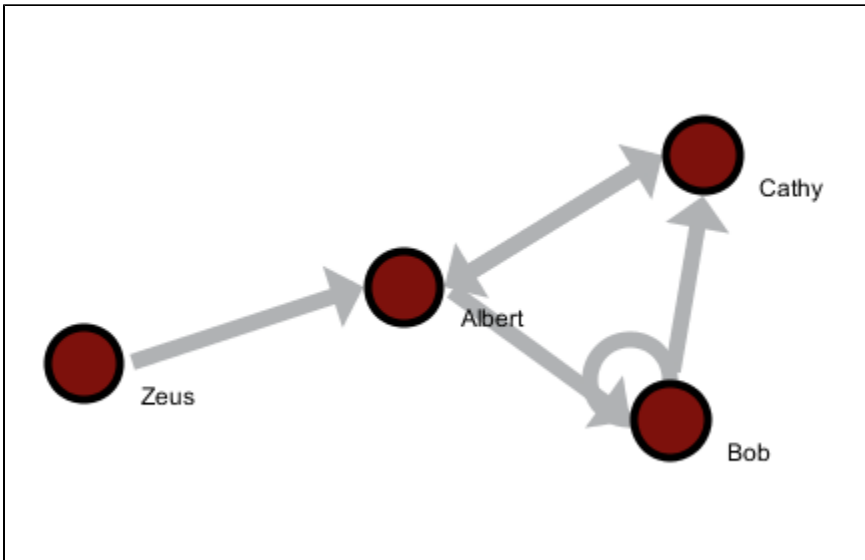
One node is created for each entity in either column. If the same entity occurs in different places (for instance 'Cathy' is mentioned in three places in our original table), duplicate nodes are not created – there will be only one "Cathy" node. If your cells are multi-valued (as in "Bob, Cathy") you can define a separator (in this example ",") to create two or more nodes ("Bob" and "Cathy") rather than one node "Bob, Cathy". This applies to both the source and target columns. Note also that you can have a node reference itself, as we do with "Bob" above.

For each row in the table, a directed edge is formed from each entity in the source column to each entity in the target column.

Sometimes you may want your resulting network to have certain attributes on the nodes and edges. See Update Network by Merging Nodes for details.

Usage Hints

The properties files for this algorithm are very similar to the properties files used in Update Network by Merging Nodes. For many pipelines you will need to create similar .properties files for both. Try creating one and adapting it for use in the other.

You can use this algorithm as a first step in producing a bibliographic coupling or co-citation network, using the Extract Reference Co-Occurrence (Bibliographic Coupling) Network or Extract Document Co-Citation Network algorithms. If you do, make sure to create your directed network in such a way that the arrows point in the direction the information flows. In the case of a documentation citation network this would mean the arrows point from the paper that is cited to the paper that is doing the citing. This can seem counter-intuitive, since the arrows will be pointing in the opposite direction of the act of citation (i.e. NOT from the citer to the citee). See Network Analysis (With Whom?) for a helpful diagram.

See Also