Graphium

UNIVERSIDAD SIMÓN BOLÍVAR

Chrysalis

Exploiting Graph Database Engines to Analyze RDF Graphs

Alejandro Flores
María-Esther Vidal
Guillermo Palma

Graph Order

URI's, BlankNodes & Literals

The number of vertices $|V|$ (i.e. *graph order*) for a particular RDF graph. Vertices are categorized as URIs, BlankNodes, or Literals.

Nodes (vertices)					**6706163**
1483969 URIs	+	**4565** Blank Nodes	+	**5217629** Literals	

Graph Size

Triples & Predicates

Returns the number of edges $|E|$ (i.e. *graph size*) for a particular RDF graph. Edges are also discriminated in terms of different predicates.

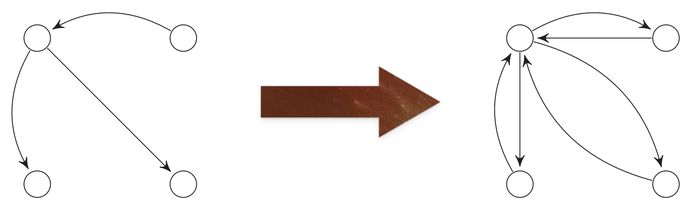
10010277 Edges (triples)	*with*	**40** Predicates

Reciprocity

Dyad & Arc based

Reciprocity measures the extend to which a triple that relates resources *A* and *B* is reciprocated by a another triple that relates *B* with *A* too.

0.000086718407 *Dyad*-based Reciprocity	**0.00017342177** *Arc*-based Reciprocity



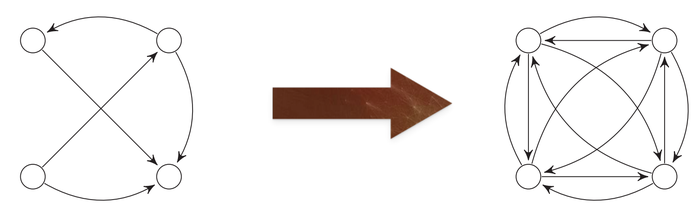
Reciprocal edges indicates stronger relationships between vertices.

Density

Measures how close is the graph to a *complete digraph* using the following formula:

$$D = \frac{|E|}{|V|(|V|-1)}$$

Density of the graph	**2.2258605e-7**



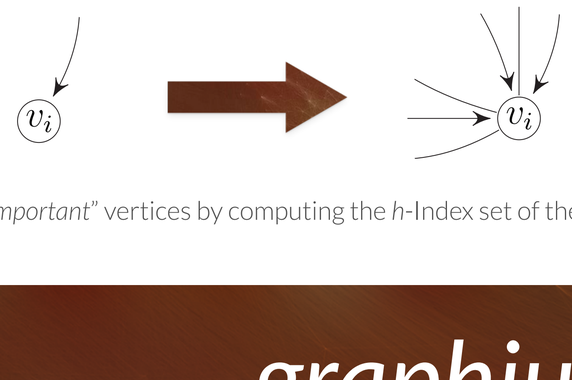
While the density augments, traverse the graph becomes more expensive.

h-Index

for in-coming & out-going edges

h is the maximum number, such that *h* vertices have each at least *h in-coming* (resp., *out-going*) neighbors.

452 *h*-Index (In-coming)	**52** *h*-Index (Out-going)

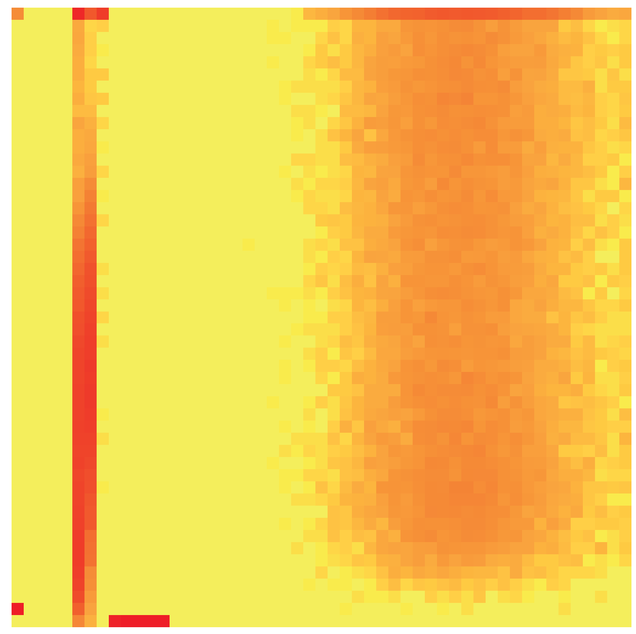


Detecting "*important*" vertices by computing the *h*-Index set of the graph.

Degree distribution

for in-coming & out-going edges

Visualization of the frequency of *in* & *out* degree values for every vertex in the RDF graph. This helps uncover hidden patterns in the structure of the graph, and is useful to explain low/high selectivity of queries that involve low/highly connected vertices of the graph.



graphium.ldc.usb.ve/***chrysalis***/