



Associative Retrieval Over Different Knowledge Organization Systems



Peter Scheir

Mathias Lux

Stefanie Lindstaedt

{pscheir, mlux, slind}@know-center.at

<http://www.know-center.at>

Road map

1. Goal
2. Associative retrieval
3. Associative networks
4. Spreading activation
5. Prototype
6. Future work

Goal

We want to

- 🌐 Build an information system helping people to learn while working
- 🌐 Find information related to current situation

We start with

- 🌐 Heterogeneous data sources
- 🌐 Different knowledge organisation systems (user, work, resources, ...)



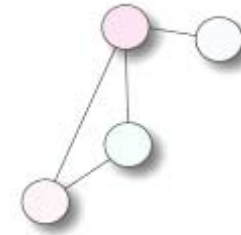
Associative retrieval

- Retrieve information which is *associated* with information that is known to be relevant
- Associations between information items (e.g. documents) can be:
 - ◆ Static (thesaurus relationships between index terms, or statistical similarity)
 - ◆ Dynamic (interaction with the user)



Associative networks

1/2

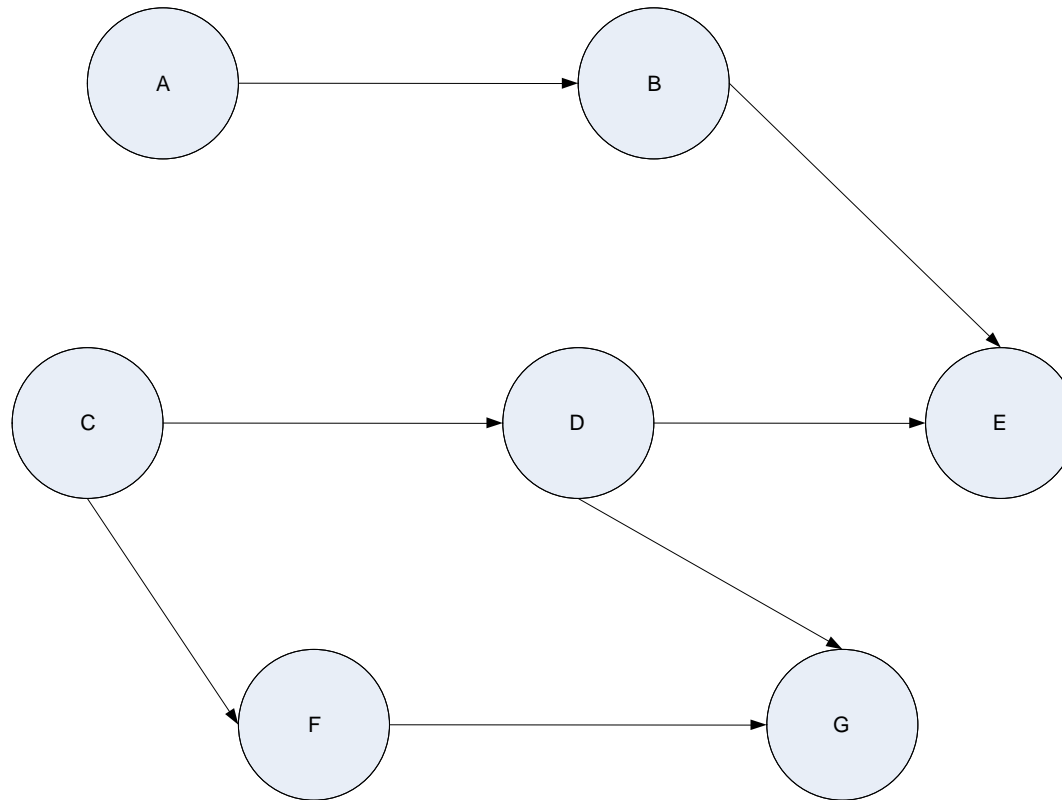


- Network of information items
- Information items represented by nodes
- Associations between information expressed as edges
- Edges can be weighted to represent the strength of the association (e.g. based on statistical measures)
- Edges can be labelled and / or directed

Associative networks

2/2

Simple network with un-weighted edges



Spreading activation

1/5

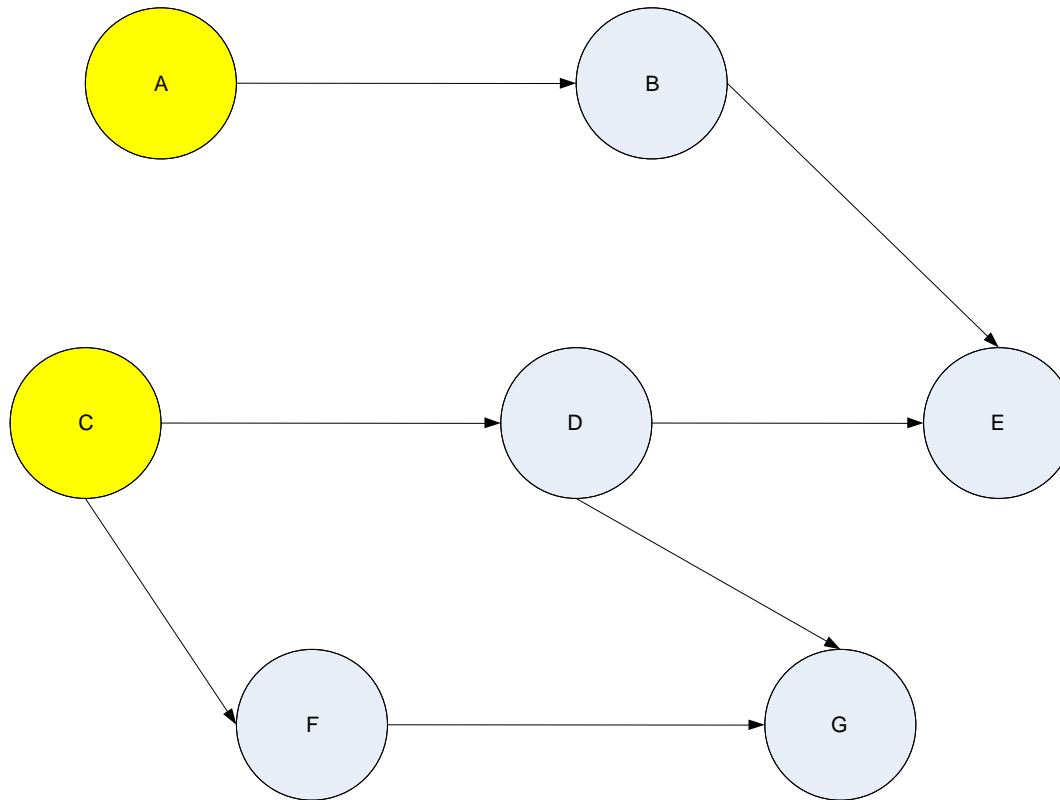
- 🌐 Technique for processing (searching in) associative networks
- 🌐 Based on the supposed mechanism of human mind
- 🌐 Initial nodes are activated
- 🌐 Energy flows over the links to the associated nodes

7

Spreading activation

2/5

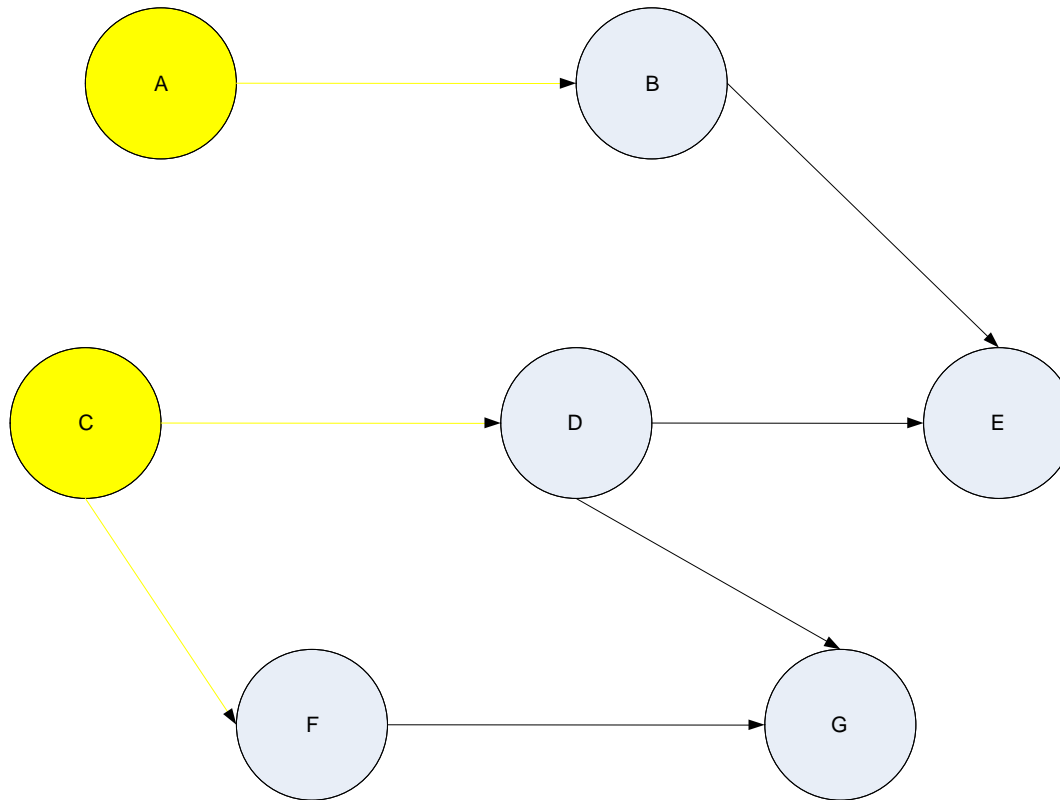
Simple network with un-weighted edges



Spreading activation

2/5

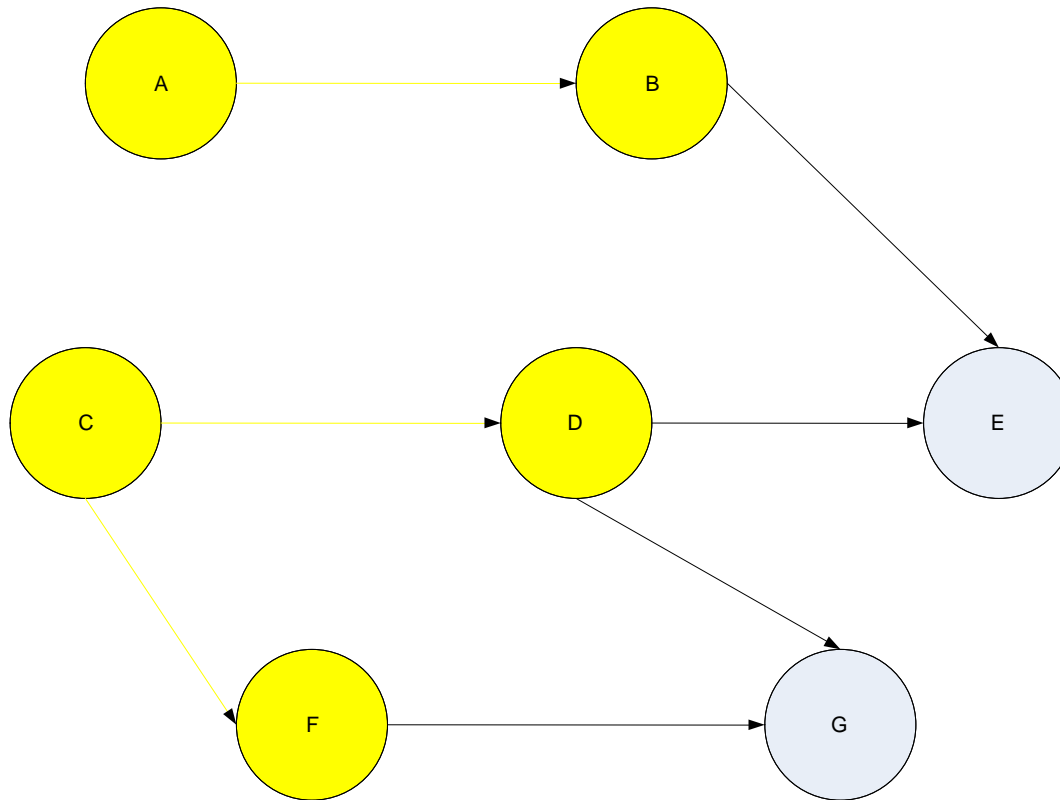
Simple network with un-weighted edges



Spreading activation

2/5

Simple network with un-weighted edges

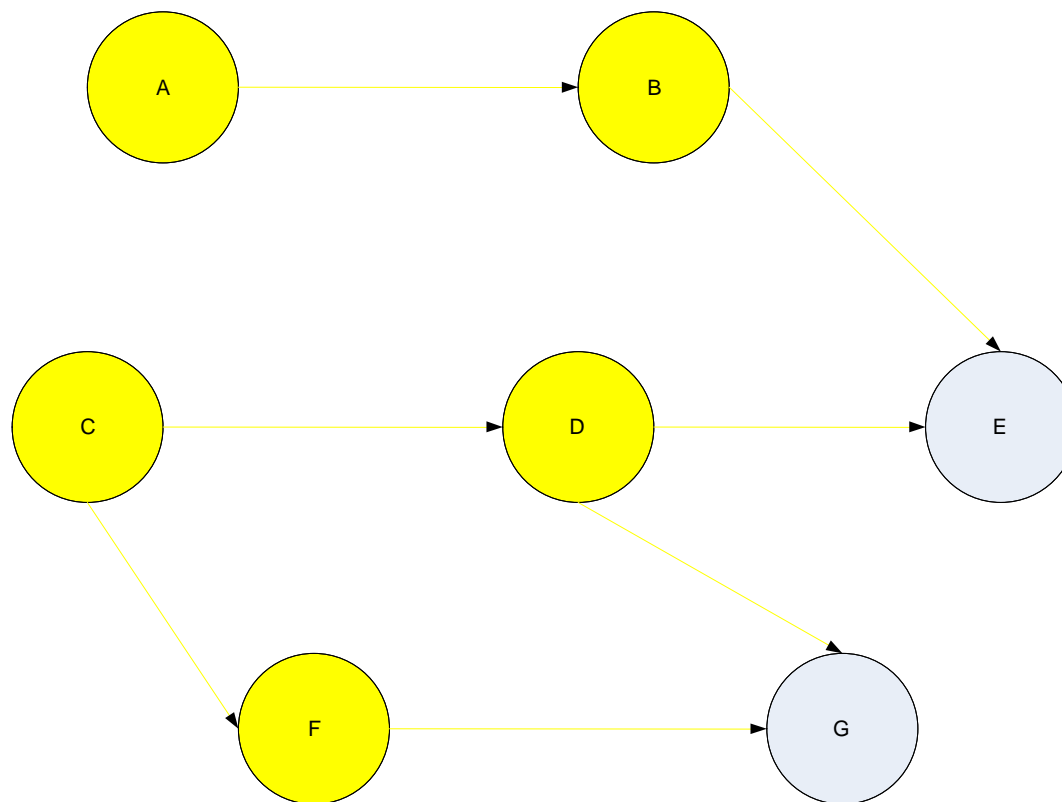


10

Spreading activation

2/5

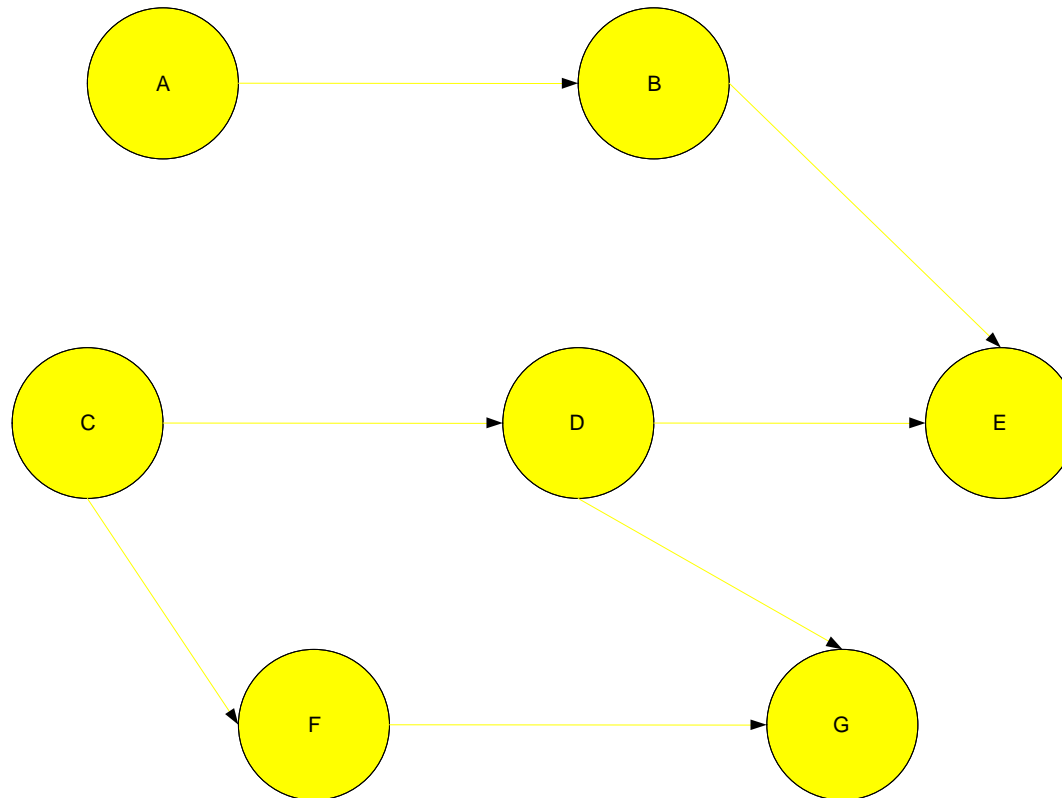
Simple network with un-weighted edges



Spreading activation

2/5

Simple network with un-weighted edges

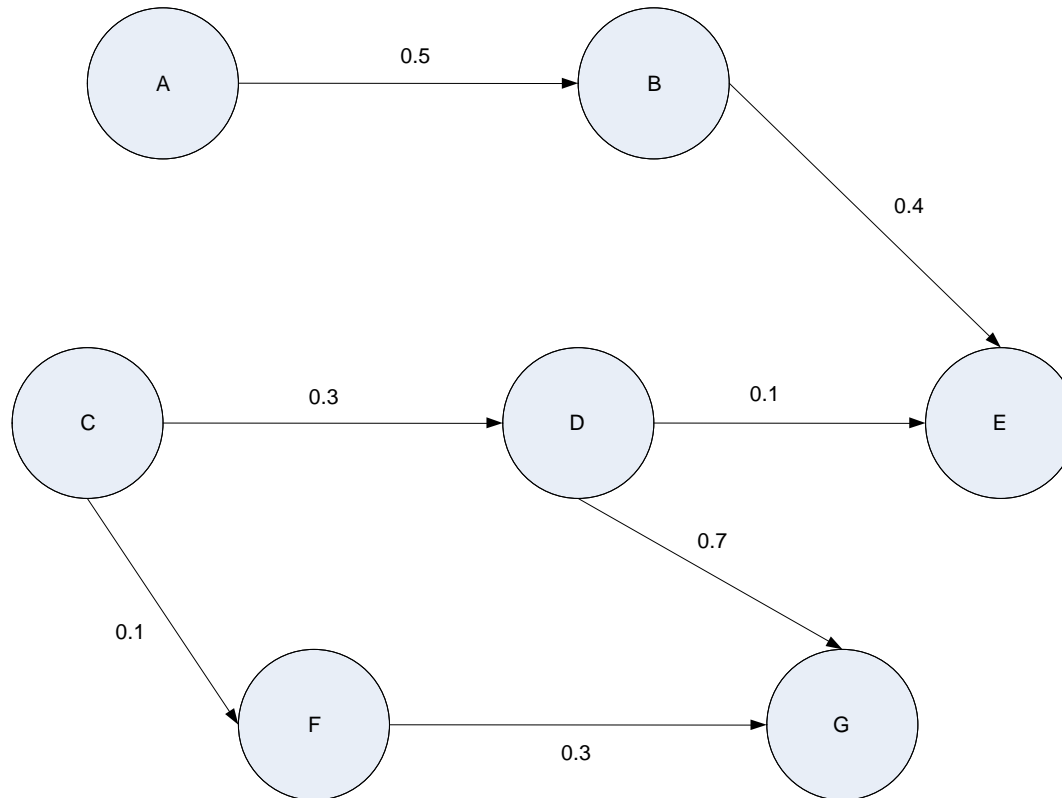


12

Spreading activation

3/5

Network with weighted edges



13

Spreading activation

4/5

Calculation of activation with weighted edges

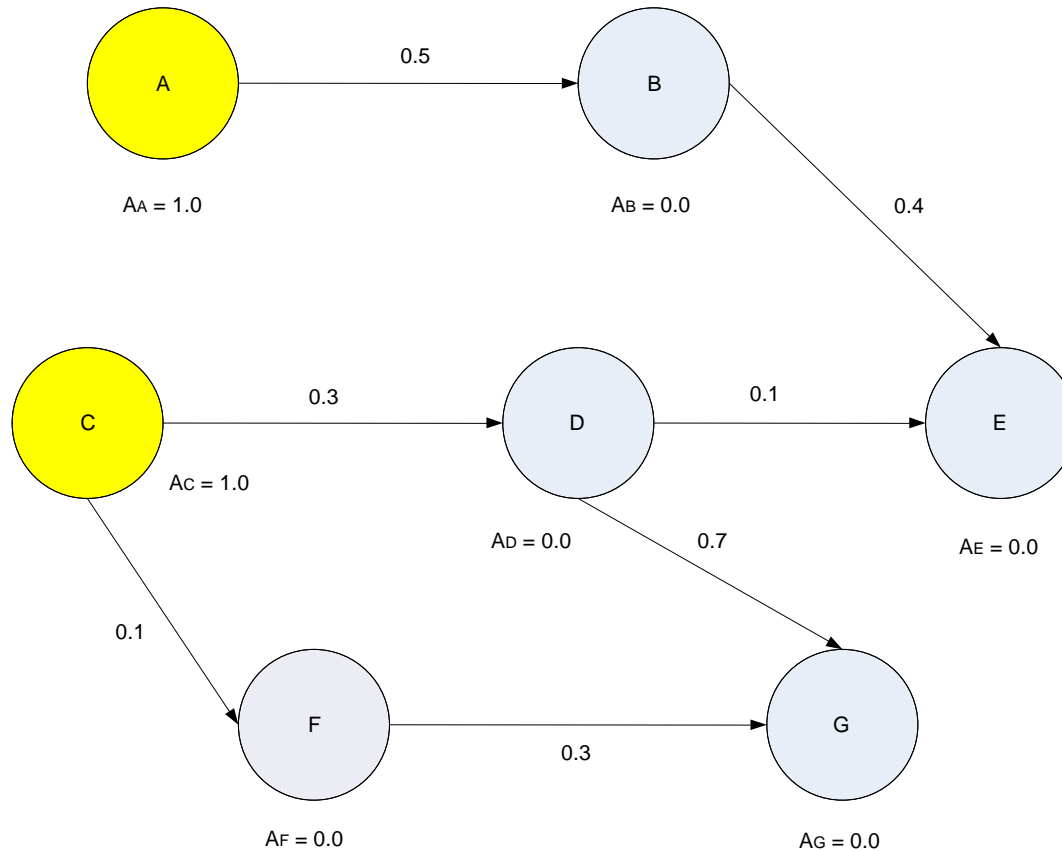
$$I_j = \sum_{i=1}^n O_i w_{ij}$$

- Input to a node is the sum of outputs from connected nodes relative to the weight of the edge connecting the nodes

Spreading activation

5/5

Network with weighted edges

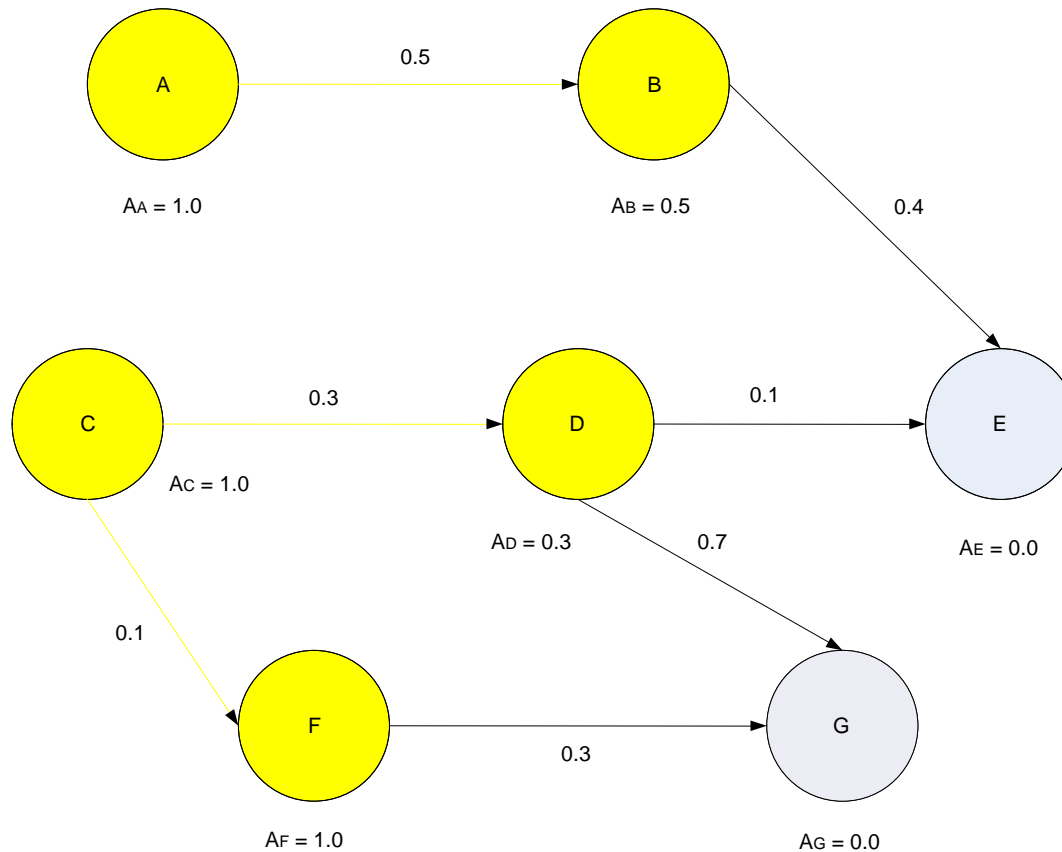


15

Spreading activation

5/5

Network with weighted edges

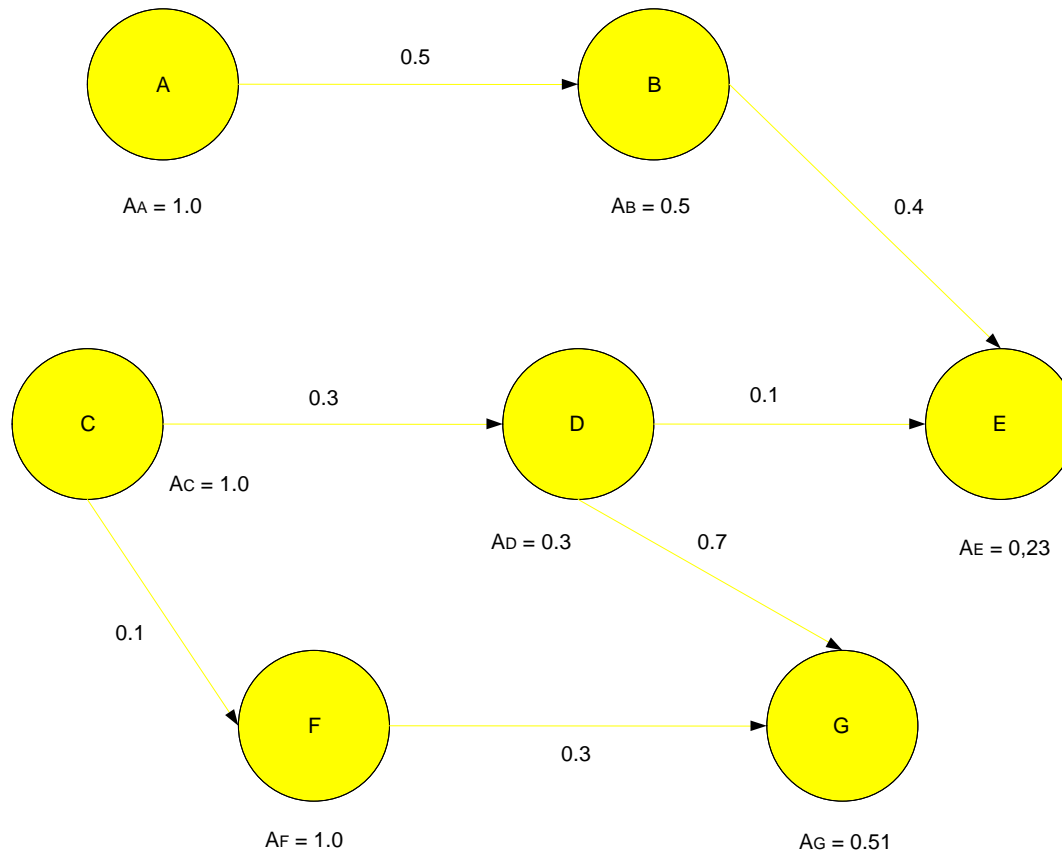


16

Spreading activation

5/5

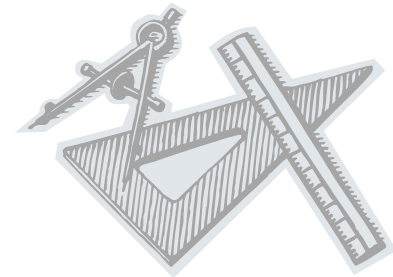
Network with weighted edges



Prototype

1/5

- Mapping of *wine* and *food* ontologies into an associative network
- Concepts and concept instances mapped to nodes (information items)
- Concept hierarchy and relations between concepts (object properties) mapped to edges
- Edges weighted via:
 - ◆ Relation type to value mapping
 - ◆ Weight mapping techniques
- Process via spreading activation



18

Prototype

2/5

Relation type to value mapping (e.g. Alani et al. 2002)

- Static mapping of relation type to edge weight
 - ◆ `has_super_class` = 1.0
 - ◆ `is_super_class_of` = 1.0
 - ◆ `has_instance` = 1.0
 - ◆ `is_instance_of` = 1.0

Prototype

3/5

Weight mapping techniques (Rocha et al. 2004)

- Cluster measure (similar to tf in IR)
 - Concepts that share more relations to other concepts are more similar

$$W(C_j, C_k) = \frac{\sum_{i=1}^n n_{ijk}}{\sum_{i=1}^n n_{ij}}$$

20

Prototype

4/5

Weight mapping techniques (Rocha et al. 2004)

- Specificity measure (similar to idf in IR)
 - ◆ Number of relations with k as destination
 - ◆ High if few concepts of type j are related to concepts of type k

$$W(C_j, C_k) = \frac{1}{\sqrt{n_k}}$$

21

Prototype

5/5

Advantages (for us)

- Different knowledge organisations systems (e. g. ontologies) can be integrated into one network
 - ◆ This can done by means of:
 - Co-indexed documents
 - Human experts
 - Heuristics (as in ontology-mediation / -learning)
 - ◆ Weights of the relations between concepts of different KOSs may depend on the credibility of the source
- Different aspects to start search at the same time

22

(Current and) Future work

Second prototype (finished)

- 🌐 Associative network from terms and documents
- 🌐 Edges are weighted by term frequency / inverse document frequency

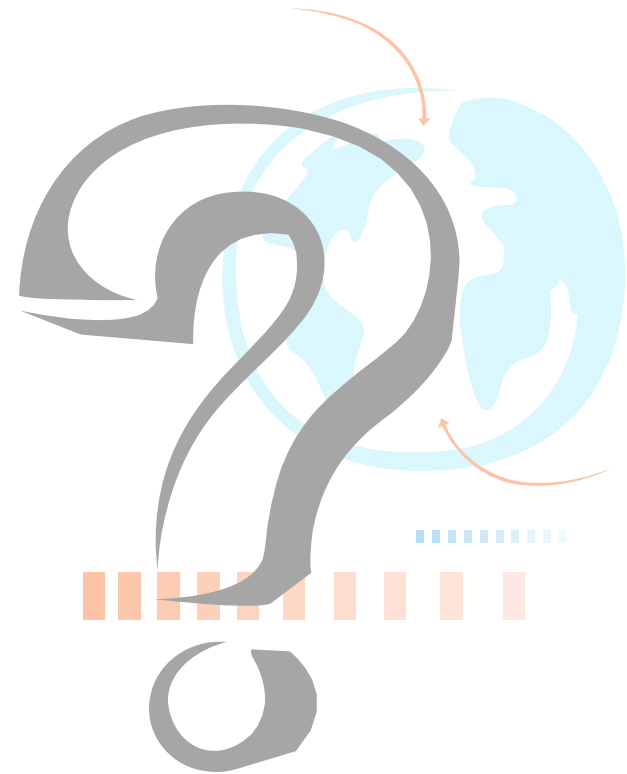
Integration of prototypes

- 🌐 Associative network containing
 - ◆ concepts (from KOSs)
 - ◆ terms (from documents)
 - ◆ documents
- 🌐 Find information related to given information
- 🌐 Ultimately: Find information related to the current situation

The end

Thank you for your attention!

Questions / comments?



Peter Scheir
pscheir@know-center.at