A Survey of
Hybrid Representations of Concept Lattices in Conceptual Knowledge Processing

Peter Eklund[1]   Jean Villerd[2]

[1]School of Information Systems and Technology
University of Wollongong
Australia

[2]Loria - INRIA Nancy Grand Est
France

ICFCA 2010

Outline

- Drawing line diagrams
  - geometrical heuristic
  - additive line diagrams
  - layer approach
  - hybrid layer + additive approach
  - force-directed approach

- Line diagrams for data exploration
  - Toscana systems and nested-line diagrams
  - handling numerical attributes through an *overview + details* approach
  - highlighting conceptual similarity
  - faceted browsing

Drawing line diagrams

Lattices in data analysis are more than just mathematical structures : they carry meaning. Therefore, drawings of such lattices should not only reflect the mathematical structure but also give a meaningful presentation for the data. (Wille 1989)

line diagrams constraints

- preservation of the partial order
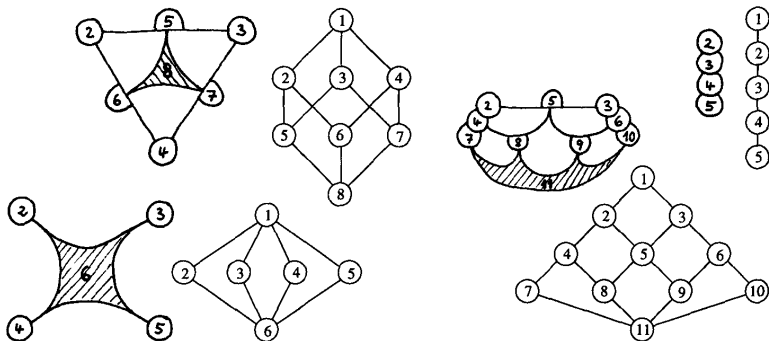- provide insights into the structure

aesthetic constraints

- minimize edge crossings
- maximize parallel lines
- no two vertices at the same point
- an edge should not cross another vertex

Geometrical heuristic [Stumme & Wille 1995]

$|U|$	shape
1	point
2	line
3	triangle
...	...
n	$(n-1)$-simplex

- each concept has a geometric representation w.r.t. is upper cover U

- identify meaningful substructures

Geometrical heuristic [Stumme & Wille 1995]

$|U|$	shape
1	point
2	line
3	triangle
...	...
n	$(n-1)$-simplex

- each concept has a geometric representation w.r.t. is upper cover U

- identify meaningful substructures

Additive line diagrams [Ganter & Wille 1999]

additive	attribute-additive
representation set	
X	irreducible attributes
representation function	
$L \to \mathcal{P}(X)$	$rep(c) = Int(c)$
$c_1 \leq c_2 \Rightarrow rep(c_2) \subseteq rep(c_1)$	
grid projection	
$X \to \mathbb{R}^2$	$\vec{a} = (-1,1), \vec{b} = (0,1), \vec{c} = (1,1)$
position function	
$pos(c) = \sum_{x \in rep(c)} \vec{x}$	$pos(\bot) = \vec{a} + \vec{b} + \vec{c} = (0,3)$
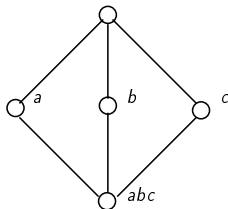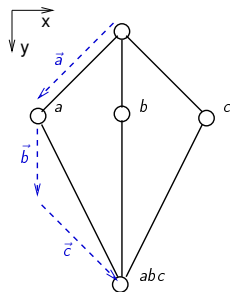
Additive line diagrams [Ganter & Wille 1999]

- in attribute-additive diagrams, some concepts may be **distended** (distance to parents large compared to average distance)

- occurs in non-distributive lattices

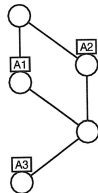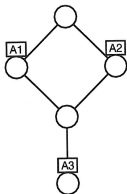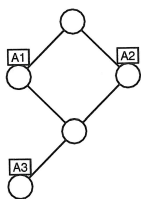irregularities may provide insights [Becker 2005]
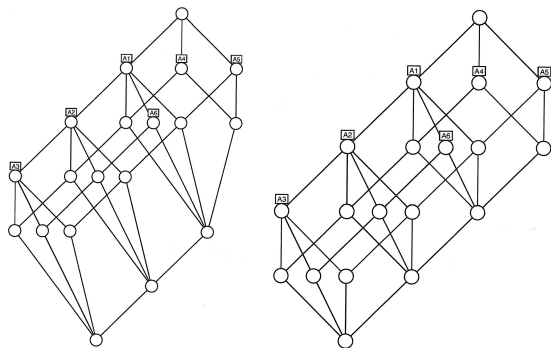
$A_1 A_4 A_5 \rightarrow A_6$ appears clearer on the left

Additive line diagrams [Ganter & Wille 1999]

- in attribute-additive diagrams, some concepts may be **distended** (distance to parents large compared to average distance)

- occurs in non-distributive lattices

irregularities may provide insights [Becker 2005]

best diagram ? depends on which interpretation is wanted

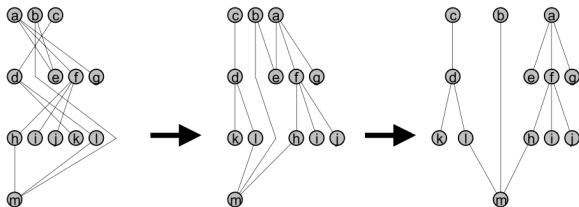# Layered approach [Cole 2001, Yevtushenko 2004]

[Sugiyama *et al.* 1981] method for drawing **directed acyclic graphs**

1. layer nodes
   - proper layered digraph (no edge span > 1), add dummy nodes
   - raking strategy : minimize height, width, dummy nodes

2. minimize edge crossings
   - NP-complete problem [Garey & Johnson 1983]
   - many heuristics, e.g. layer-by-layer sweep
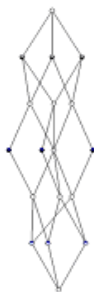

[Dalen & Spaans 2001]

Layered approach [Cole 2001, Yevtushenko 2004]

advantages
efficient for planar or "close to planar" line diagrams

drawbacks
does not emphasize regular structures and symmetrical displays

layered approach   additive approach

[Yevtushenko 2004]

Hybrid layer + additive approach [Cole 2001]

- **layer approach** : determine vertical positioning

- **additive approach** : determine horizontal positioning

1. **choose horizontal vectors** such that the hybrid diagram is satisfactory :
   - no two concepts on the same layer have the same x-coordinates
   - no edge crosses the coordinates of a concept

   formulated as a constraint satisfaction problem to produce a list of satisfactory diagrams

2. **satisfactory diagrams are partially order** w.r.t. quality functions
   - symmetry between siblings on the same layer
   - minimize lines (distinct edge vectors)
   - maximize chains

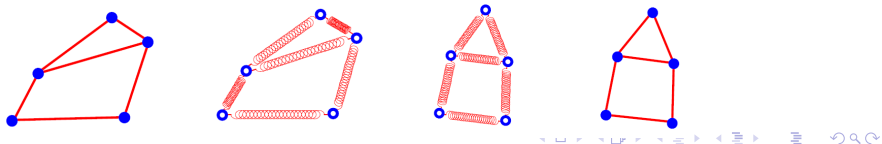# Force-directed approach [Cole 2000, Freese 2004]

[Eades 1984][Kamada & Kawai 1989][Fruchterman & Reingold 1991]

forces exerted on a vertex

- **attractive force** : edges act as springs $f_a(d) = -k_a d$
- **repulsive force** : nonneighboring nodes repel each other $f_r(d) = \frac{k_r}{d^2}$

heuristic

1. assign random positions to vertices
2. until an equilibrium configuration is reached, do
   - for each vertex, compute the sum of exerted forces
   - update each vertex position

Force-directed approach [Freese 2004]

1. **layering** : $layer(a) = height(a) - depth(a) + M$ (fixes z-coordinate)

2. **attractive force** : between comparable nodes (z-coordinate remains fixed)

3. **repulsive force** : between incomparable nodes (z-coordinate remains fixed)
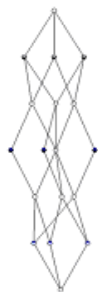
4. **projection** into the plane
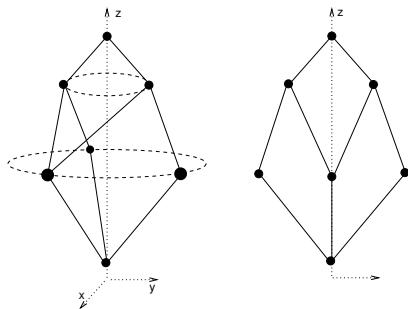
Force-directed approach [Freese 2004]

advantages
efficient with non-planar diagrams, reveals symmetries

drawbacks
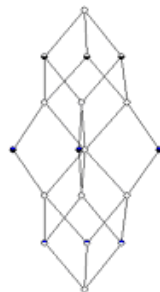does not maximize parallel lines

layered approach   additive approach   Freese
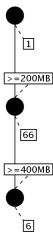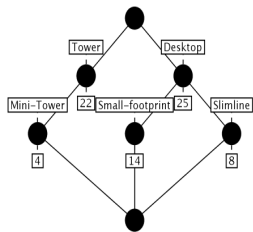
[Yevtushenko 2004]

other approaches

- multidimensional additive diagrams [Becker 2001]

- planar diagrams and forces [Zschladig 2005]
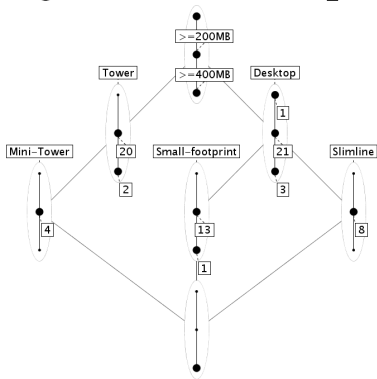
- conflict distance [Ganter 2004]

- . . .

tools

- Glad

- ConExp

- Galicia

- Toscana

- . . .

Conceptual knowledge retrieval [Wille 2006]

Conceptual knowledge retrieval is often a process in which humans search for something with they only vaguely imagine

Toscana [Vogt *et al.*]

- interactive retrieval and navigation
- conceptual scales as search structures
- attributes partitioning through nested-line diagrams

Overview + details navigation [Villerd *et al.* 2009]

- **overview** : line diagram built w.r.t. a set of binary or nominal attributes
- **detailed view** : projection of objets w.r.t. numerical attributes using Multi Dimensional Scaling (MDS), revealing proximities of objects
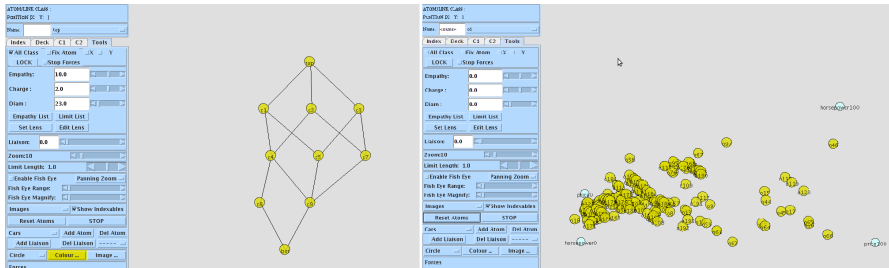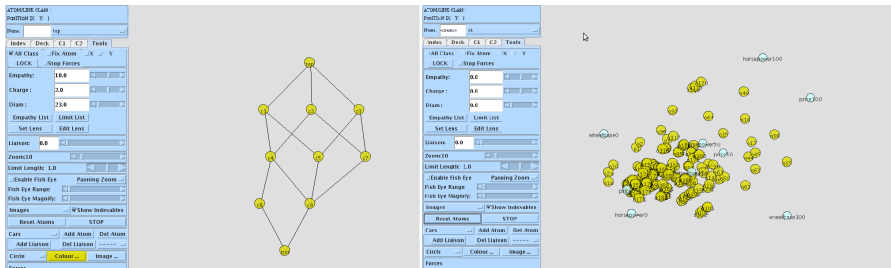

Force-directed MDS [Chalmers 1996] allows the user to

- dynamically select the subset of numerical attributes used to compute dissimilarities
- dynamically observe addition/removal of objects while navigating on the line diagram

Overview + details navigation [Villerd *et al.* 2009]

- **overview** : line diagram built w.r.t. a set of binary or nominal attributes
- **detailed view** : projection of objets w.r.t. numerical attributes using Multi Dimensional Scaling (MDS), revealing proximities of objects

Force-directed MDS [Chalmers 1996] allows the user to

- dynamically select the subset of numerical attributes used to compute dissimilarities
- dynamically observe addition/removal of objects while navigating on the line diagram
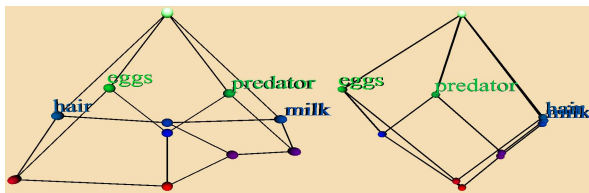
conceptual similarity

- 2 concepts are *conceptually similar* if their extents are nearly equal
- distance $\delta(c_i, c_j) = |Ext(c_i) \triangle Ext(c_j)|$

diagram improvement algorithm

- spring between each pair of concepts with $\delta(c_i, c_j)$ as natural length
- iteratively improves the diagram so that the actual distance $d(c_i, c_j)$ is proportional to $\delta(c_i, c_j)$

$$milk \xrightarrow{95.1\%} hair$$
$$hair \xrightarrow{90.6\%} milk$$

Faceted navigation [Eklund]

- the user sees the extent of a concept
- he progressively browses the collection by adding or removing proposed attributes (moving to an upper or lower concept)
- the font size is proportional to the number of retrieved objects

Virtual Museum of the Pacific

Conclusion and Perspectives

- hybrid representations that depart from the conventions of line diagram drawing have pratical value when used in an appropriate application context

- further experimentation in this direction will lead to innovations that enhance the application of Conceptual Knowledge Processing

- solutions for handling large line-diagrams

- scenarios that activate/deactivate specific forces in order to highlight specific parts of the lattice