Dendrogramix:
a Hybrid Tree-Matrix Visualization Technique
to Support Interactive Exploration of Dendrograms

Renaud Blanch (IIHM) <blanch@imag.fr>
Rémy Dautriche (IIHM) <dautriche@imag.fr>
Gilles Bisson (AMA) <bisson@imag.fr>

Université Grenoble Alpes, CNRS
Laboratoire d’Informatique de Grenoble

april 14, 2015
<PacificVis 2015>
Hierarchical agglomerative clustering

Classification
A greedy algorithm that recursively builds a hierarchy of clusters:
- one cluster per item;
- the two most similar clusters are agglomerated until there is a single cluster left.
Hierarchical agglomerative clustering (cont.)

**Inputs**

- a set of **items**
- a **metric** to measure the similarity of items
- a **linkage**, i.e. a generalization of this metric to groups of items
Example

**Items**
Points to classify according to their Euclidean distance.

- E
- C
- D
- B
- A

blanch@imag.fr, dau triche@imag.fr, bisson@imag.fr
Visual representation

Dendrogram

A dendrogram is the canonical representation of the clusters’ hierarchy.

B D E C A
Hierarchical agglomerative clustering

Limitation

Items similarities are lost, no explanation for the clustering. Heatmaps are often juxtaposed to dendrograms but:

- at the expense of screen space; and
- making items comparison difficult for distant items.
Dendrogramix builds a visualization by merging:

- the **hierarchy of clusters**; and
- the **similarity matrix**
**Design**

**Similarities**
encoded as dot sizes.

```
A B C D E
A . . . . .
B . . . . |
C . . . . .
D . . . . .
E . . . . .
```

blanch@imag.fr, dautriche@imag.fr, bisson@imag.fr
Matrix reordering using optimal leaf order [BarJoseph et al., 2003].

B D E C A
B ●●●●
D ●●●
E ●●●●●
C ●●●●
A ●●●●
Design (cont.)

Clusters encoding

using enclosure and luminance.
Dendrogramix: visualisation

Design (cont.)

Dendrogramix

top half of the matrix, rotated.

B  D  E  C  A
Whole process

Using the tree of clusters and the similarity matrix as inputs, the Dendrogramix is built: Using the tree of clusters and the similarity matrix as inputs, the Dendrogramix is built: Using the tree of clusters and the similarity matrix as inputs, the Dendrogramix is built: Using the tree of clusters and the similarity matrix as inputs, the Dendrogramix is built: Using the tree of clusters and the similarity matrix as inputs, the Dendrogramix is built:
**Dendrogramix**

**Vis Community**

*People* are represented as the vector of their *coauthors*. Similarities are computed using the *cosine similarity* and *single linkage*.
Items Comparison

Single item highlight

blanch@imag.fr, dautriche@imag.fr, bisson@imag.fr

UGA, CNRS
Items Comparison (cont.)

Pairwise items comparison

Dendrogramix: interactions

blanch@imag.fr, dauvrich@imag.fr, bisson@imag.fr

blanch@imag.fr, dauvrich@imag.fr, bisson@imag.fr
Partition Selection

Partition selection

Dendrogramix: interactions

blanch@imag.fr, dautriche@imag.fr, bisson@imag.fr
Cluster Aggregation

Cluster Labeling

blanch@imag.fr, dau triche@imag.fr, bisson@imag.fr
Cluster Aggregation (cont.)

Cluster folding

blanch@imag.fr, dautriche@imag.fr, bisson@imag.fr

UGA, CNRS
Tree Reordering

Reordering by direct manipulation
Demo

time for a demo!

demo

blanch@imag.fr, dautriche@imag.fr, bisson@imag.fr

UGA, CNRS
Implementation

Portable open source code available at:

Software used:

- Python
- OpenGL/GLUT
- scipy

blanch@imag.fr, dautriche@imag.fr, bisson@imag.fr
Insight-based evaluation [North et al., 2011]

- **data set**: 189 permanent researchers of the lab, characterized by the 2688 co-authors of their 3245 publications over the last 4 years.

- **participants**: 6 senior researchers from the lab (deputy director or group heads).

- **two visualizations**: classical dendrogram vs. Dendrogramix.
Evaluation (cont.)

Dendrogram

blanch@imag.fr, dautriche@imag.fr, bisson@imag.fr
Evaluation (cont.)

Dendrogramix
## Insights and Value

<table>
<thead>
<tr>
<th>Category</th>
<th>CD</th>
<th>DX</th>
<th>common</th>
</tr>
</thead>
<tbody>
<tr>
<td>overview</td>
<td>6 6</td>
<td>7 7</td>
<td>5</td>
</tr>
<tr>
<td>patterns</td>
<td>3 8</td>
<td>14 45</td>
<td>3</td>
</tr>
<tr>
<td>groups</td>
<td>14 34</td>
<td>17 43</td>
<td>14</td>
</tr>
<tr>
<td>details</td>
<td>13 19</td>
<td>27 67</td>
<td>13</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td><strong>36</strong></td>
<td><strong>67</strong></td>
<td><strong>162</strong></td>
</tr>
<tr>
<td>per insight</td>
<td>1.86</td>
<td>2.49</td>
<td></td>
</tr>
<tr>
<td>per minute</td>
<td>0.30</td>
<td>0.56</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Dendrogramix

blanch@imag.fr, dautriche@imag.fr, bisson@imag.fr

UGA, CNRS
Contributions

Dendrogramix is an alternative to dendrograms,

- showing **items similarities**; and
- allowing **interactive exploration**.
Future work

Plenty of possibilities, e.g. combination with heatmaps.
Thank you for your attention!

<http://iihm.imag.fr/blanch/projects/dendrogramix/>
References