Conditionally Acyclic CO-Networks for Efficient Preferential Optimization

Pierre-François Gimenez 1  Jérôme Mengin 2
Contact: pierre-francois.gimenez@centralesupélec.fr
1 CentraleSupélec, Univ. Rennes, IRISA, France
2 IRIT, Université Paul Sabatier, CNRS, France

Preferences representation

- Recommendations are ubiquitous: VOD streaming services, product configurators, e-commerce platforms... They help the user to navigate the space of items
- We focus on products described by vectors, such as kitchens, computers, cars, or vacations
- Classical recommendation algorithms cannot be applied because the number of bought/chosen items is negligible regarding the number of possible combinations
- A classical query is the preferential optimization of a partially defined object $u$: what is $\text{opt}(u)$ the preferred extension of $u$?

Two main ordinal model families have been proposed to model preferences in combinatorial domains, acyclic CP-nets and LP-trees:

![A CP-net](image)

$\begin{align*}
\alpha : a > b \\
\alpha : b > \bar{b} \\
\bar{b} : \bar{c} > b \\
ab : c > \bar{c} \\
\end{align*}$

![A LP-tree](image)

$\begin{align*}
\alpha : a > \bar{a} \\
\bar{a} : \bar{b} > b \\
\bar{c} : c > \bar{c} \\
\end{align*}$

However:

- LP-trees are not very succinct and the relative importance of attributes is not useful for preferential optimization
- acyclic CP-nets are not as expressive as LP-trees for optimization
- unsupervised preferences learning approach cannot be extended to CP-nets since they encode partial orders

Conditionally acyclic CO-nets

- A variant of CP-nets suited for optimization
- As expressive as LP-trees for optimization
- Even more succinct than CP-nets (especially with high-dimension variables)
- Fast preferential optimization with extended Forward Sweep algorithm

![A CO-net](image)

Conditionally acyclic CO-nets could be learned by minimising the description length of $D$:

$$L(D) = \min_{H \in \mathcal{H}} (L(H) + L(D | H)),$$

where the size of a model $H$ is:

$$L(H) = L_0(|X|) + \sum_{a \in X} L_0(|Pa(a,N)|) + \log_2 \left( \frac{|X| - 1}{|Pa(a,N)|} \right) + |Pa(N)| \log_2 |N|$$

and the size of the data $D$ given a model $H$ is:

$$L(D | H) = \sum_{a \in D} L_0(|\text{code}(a,H)|) + \log_2 \left( \frac{|X|}{|\text{code}(a,H)|} \right) + \sum_{x \in \text{code}(a,H)} \log_2(|X| - 1)$$

Optimization as decomposition

- The preferential optimization task transforms a partially defined vector into a fully defined vector
- We propose to introduce the inverse function to compress vectors: $\text{code}(o)$ is the smallest vector such that $\text{opt}(\text{code}(o)) = o$
- Functions opt (for decompression) and code (for compression) can be computed quickly with variants of the Forward Sweep algorithm

Experiment

- Evaluation of the compression rates of 3 sales histories of cars from Renault
- Datasets are in CSV format and weigh a few MB
- Data and code available at https://github.com/PFGimenez/co-net-ecai23

<table>
<thead>
<tr>
<th>Dataset</th>
<th>LZMA</th>
<th>PPMd</th>
<th>bzip2</th>
<th>DEFLATE</th>
<th>CO-net</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>95.80%</td>
<td>97.90%</td>
<td>97.46%</td>
<td>94.50%</td>
<td>97.03%</td>
</tr>
<tr>
<td>Medium</td>
<td>96.04%</td>
<td>97.98%</td>
<td>97.71%</td>
<td>94.82%</td>
<td>97.12%</td>
</tr>
<tr>
<td>Big</td>
<td>96.40%</td>
<td>97.93%</td>
<td>97.64%</td>
<td>94.90%</td>
<td>97.67%</td>
</tr>
</tbody>
</table>

Table: Space savings on the three Renault datasets

- Performances of CO-nets are similar to specialized compression algorithms
- CO-nets can effectively represent real-world preferences!

Conclusion

- CO-nets are models tailored for preferential optimization
- Experimental assessment of their representation of real-world preferences is promising
- This article paves the way toward unsupervised learning of CO-nets from sales histories with MDL