# Conditionally Acyclic CO-Networks for Efficient Preferential Optimization


Pierre-François Gimenez [1] Jérôme Mengin [2]

Contact: pierre-francois.gimenez@centralesupelec.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 $opt(u)$ the preferred extension of $u$?

## State of the art

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

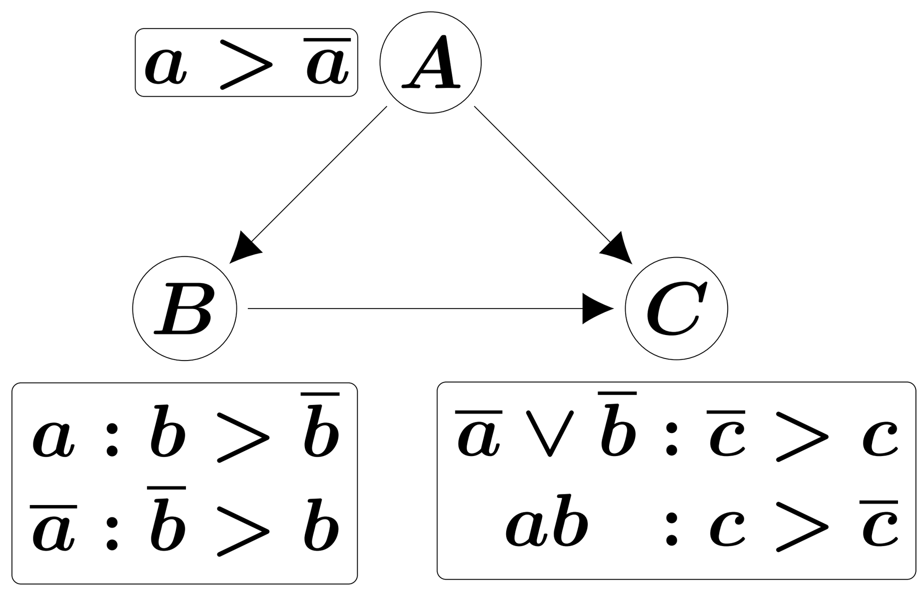


Figure: A CP-net

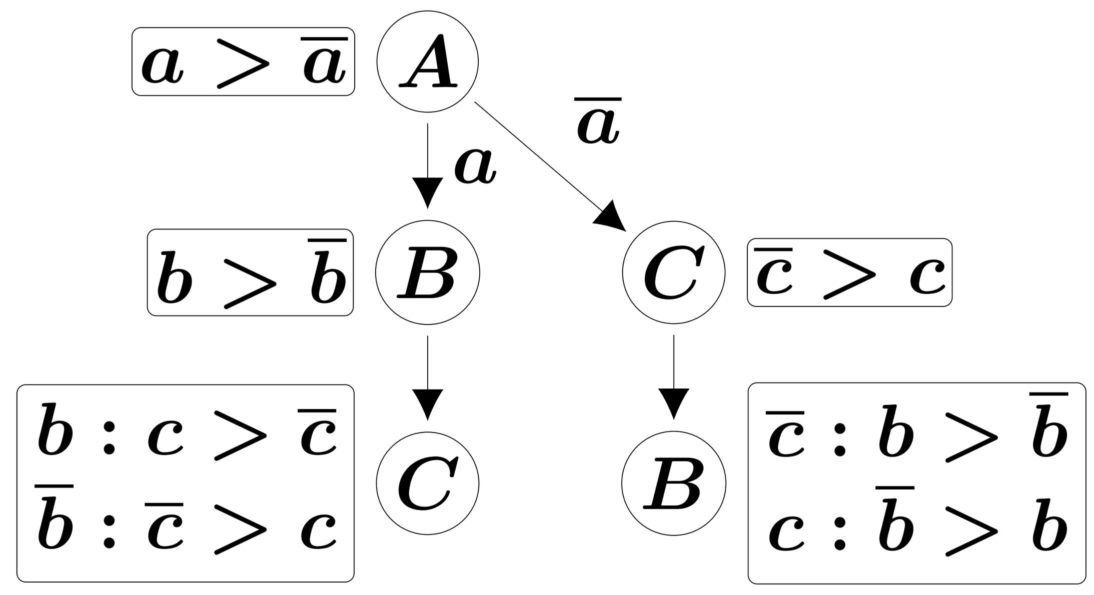


Figure: A LP-tree

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

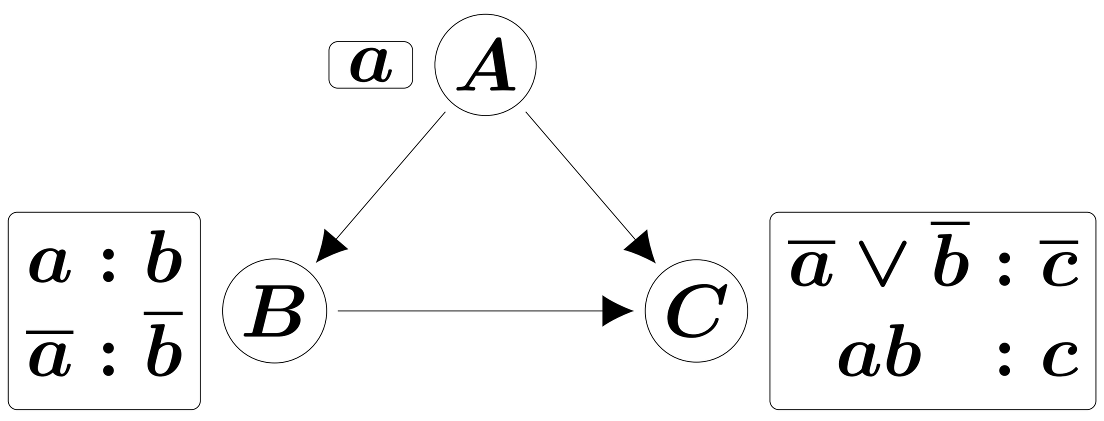


Figure: A CO-net

Conditionally acyclic CO-nets are:

- 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

## Optimization as decompression

- The preferential optimization task transforms a partially defined vector into a fully defined vector
- We propose to introduce the inverse function to compress vectors: $code(o)$ is the smallest vector such that $opt(code(o)) = o$
- Functions $opt$ (for decompression) and $code$ (for compression) can be computed quickly with variants of the Forward Sweep algorithm
- CO-nets could be learned by minimising the description length of $D$: $L(D) = \min_{H\in\mathcal{H}} \left(L(H) + L(D|H)\right)$, where the size of a model $H$ is:

$$L(H) = L_{\mathbb{N}}(|\mathcal{X}|) + \sum_{N\in\mathcal{X}} \left( L_{\mathbb{N}}(|Pa(N)|) + \log_2 \binom{|\mathcal{X}|-1}{|Pa(N)|} + |Pa(N)| \log_2 |N| \right)$$

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

$$L(D \mid H) = \sum_{o\in D} \left( L_{\mathbb{N}}(|code(o,H)|) + \log_2 \binom{|\mathcal{X}|}{|code(o,H)|} + \sum_{x\in code(o,H)} \log_2(|X|-1) \right)$$

## 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

| Dataset | LZMA | PPMd | bzip2 | DEFLATE | CO-net |
|---|---|---|---|---|---|
| Small | 95.80% | 97.90% | 97.46% | 94.50% | 97.03% |
| Medium | 96.04% | 97.98% | 97.71% | 94.82% | 97.12% |
| Big | 96.40% | 97.93% | 97.64% | 94.90% | 97.67% |

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