# Learning Conditional Preference Networks: an Approach Based on the Minimum Description Length Principle

*Pierre-François Gimenez, Jérôme Mengin* CentraleSupélec, IRIT
Contact: pierre-francois.gimenez@centralesupelec.fr
IJCAI'24

## Context: recommendation in e-commerce

- Highly customizable items (e.g., cars, computers, travel, kitchens) form a huge combinatorial space
- Classical recommendation algorithms are not scalable enough to be usable
- To help users find the product they prefer, we need to modelize their preference over this combinatorial space using a preference model class

Contribution: how to learn CP-nets preferences from past sales

## How do we guess what the user *likely* prefers?

Answer: Use *Sales History D* (a multiset of items sold in the past)
The higher an outcome is ranked in the user's preference, the greater the probability that they ends up with it.

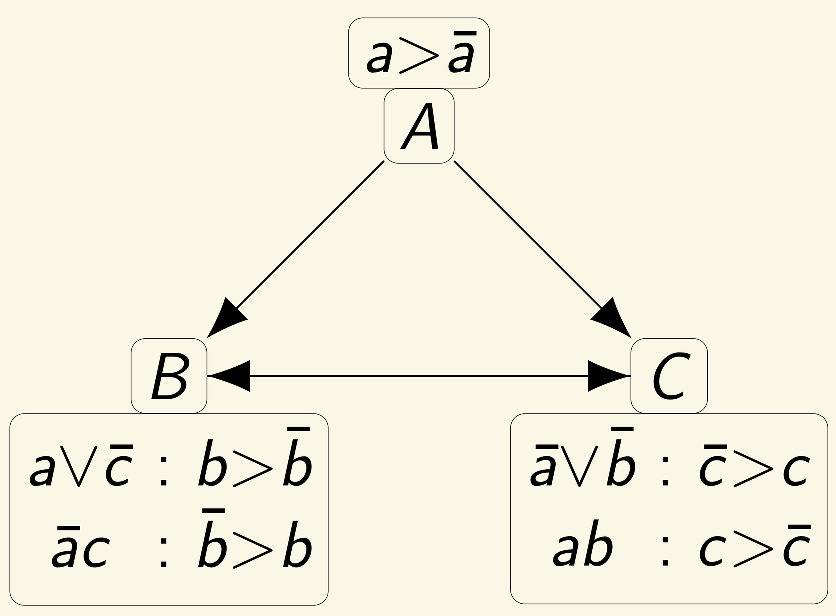
Induce a preference model that "explains" $D$

| Model class | Recommendation query complexity | Expressiveness | Learnable from $D$ |
|---|---|---|---|
| Conditional Lexicographic Preferences | P | Low | Yes |
| Bayesian Networks | NP-hard | Maximum | Yes |
| **Acyclic CP-nets** | **P** | **High** | **No*** |

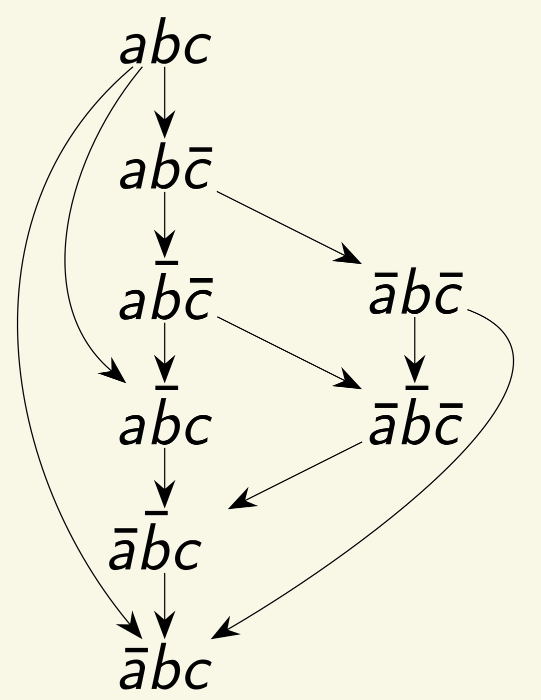
*until this article

## Conditional Preference Network (CP-net)

A CP-net = a directed graph + local preference tables




A CP-net $\phi_0$




Its relation order $\succ_{\phi_0}$

Polynomial computation of $o = \mathrm{opt}(u, \succ_\phi)$, the most preferred alternative $o$ that expands a partially defined alternative $u$

## Minimum Description Length principle for learning CP-nets from sales history

MDL principle: choose model $\phi$ that maximises compression of $D$:

$$\min_\phi \left(L(\phi) + L(D|\phi)\right)$$

Lossless compression for alternative $o$: compress $o$ with $\mathrm{code}(o, \phi)$ = smallest partial instantiation $u$ s.t. $\mathrm{opt}(u, \succ_\phi) = o$. Uncompress with opt.
For instance: $\mathrm{code}(a\bar{b}\bar{c}, \phi_0) = \bar{b}$ because $a\bar{b}\bar{c}$ is the optimal alternative when $B = \bar{b}$.

$$L(\phi) = L_{\mathbb{N}}(n) + \sum_{N \in \mathcal{X}} L_{\mathbb{N}}(|Pa(N)|) + \log_2 \binom{n-1}{|Pa(N)|} + |\underline{Pa(N)}| \log_2 |\underline{N}|$$

$$L(D|\phi) = \sum_{o \in D} \left[ L_{\mathbb{N}}(|\,\mathrm{code}(o,\phi)|) + \log_2 \binom{n}{|\,\mathrm{code}(o,\phi)|} + \sum_{X \in \mathrm{code}(o,\phi)} \log_2(|\underline{X}| - 1) \right]$$

## Complexity of CP-net learning with MDL principle

We study an approximation of $L(\phi) + L(D|\phi)$, the Normalized Mean Code Length: $NMCL(\phi) = \frac{1}{n} E_p[|\mathrm{code}(\cdot, \phi)|]$

- **Sample complexity :** For the family of CP-nets with $n$ nodes and whose nodes have at most $k$ parents:

$$N(\delta, \epsilon) = O\left(\frac{d^{2k}}{\epsilon^2}\left(\ln \frac{1}{\delta} + k(\ln d + \ln(n+1))\right)\right)$$

- **Computational complexity :** Finding the acyclic CP-net that minimizes the empirical score over $D$ is NP-complete

## Learning algorithm

**Algorithm 1:** Hill climbing search for CP-net learning

**Data:** a dataset $D$, an initial CP-net $\phi'$

1 $score \leftarrow L(\phi') + L(D|\phi')$; $previous\_score \leftarrow +\infty$
2 **while** $score < previous\_score$ **do**
3 $\quad \phi \leftarrow \phi'$
4 $\quad neighbors \leftarrow transformations(\phi)$
5 $\quad$ remove non-acyclic graphs from *neighbors*
6 $\quad$ fit CPTs of *neighbors* from $D$
7 $\quad \phi' \leftarrow \arg\min_{\phi'' \in neighbors} L(\phi'') + L(D|\phi'')$
8 $\quad previous\_score \leftarrow score$
9 $\quad score \leftarrow L(\phi') + L(D|\phi')$
10 **return** $\phi$

line 4 Neighbors of current CP-net $\phi'$ obtained by adding, removing or reversing edges
line 6 For every neighbor $\phi'$, attribute $X$, $u \in \mathrm{Pa}(\phi', X)$: $>_u$= order of decreasing conditional frequency in $D$

## Experiments

**Protocol**: recommendation task
**Conclusion**

- Lower accuracy than BN, but much faster
- Clustering help with the limited expressivity

**Future works**

- Apply this framework to other preference model classes
- Investigate the connection between BN and CP-nets