Algorithmic Decision Theory 2024

Learning multiple multicriteria models from heterogenous preferences

Vincent Auriau, Khaled Belahcène, Emmanuel Malherbe, Vincent Mousseau







Contents

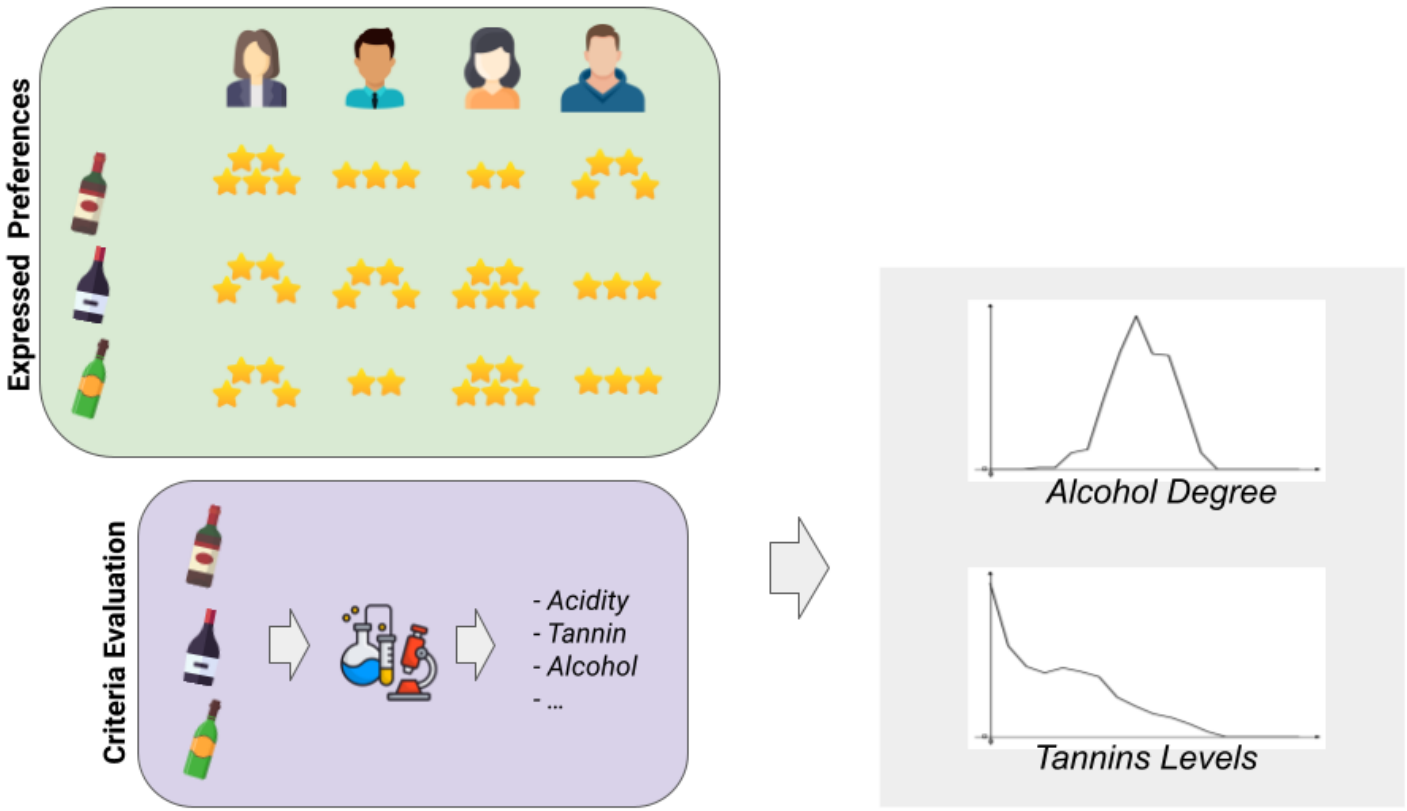
Preference Learning and Segmentation
- Motivations
- Definition

Algorithm to learn from heterogeneous preferences
- Notations
- The UTA Model
- Proposed MILO formulation for PLS
- Proposed tractable heuristic for PLS

Numerical results
- Synthetic experiments
- Real-World experiments

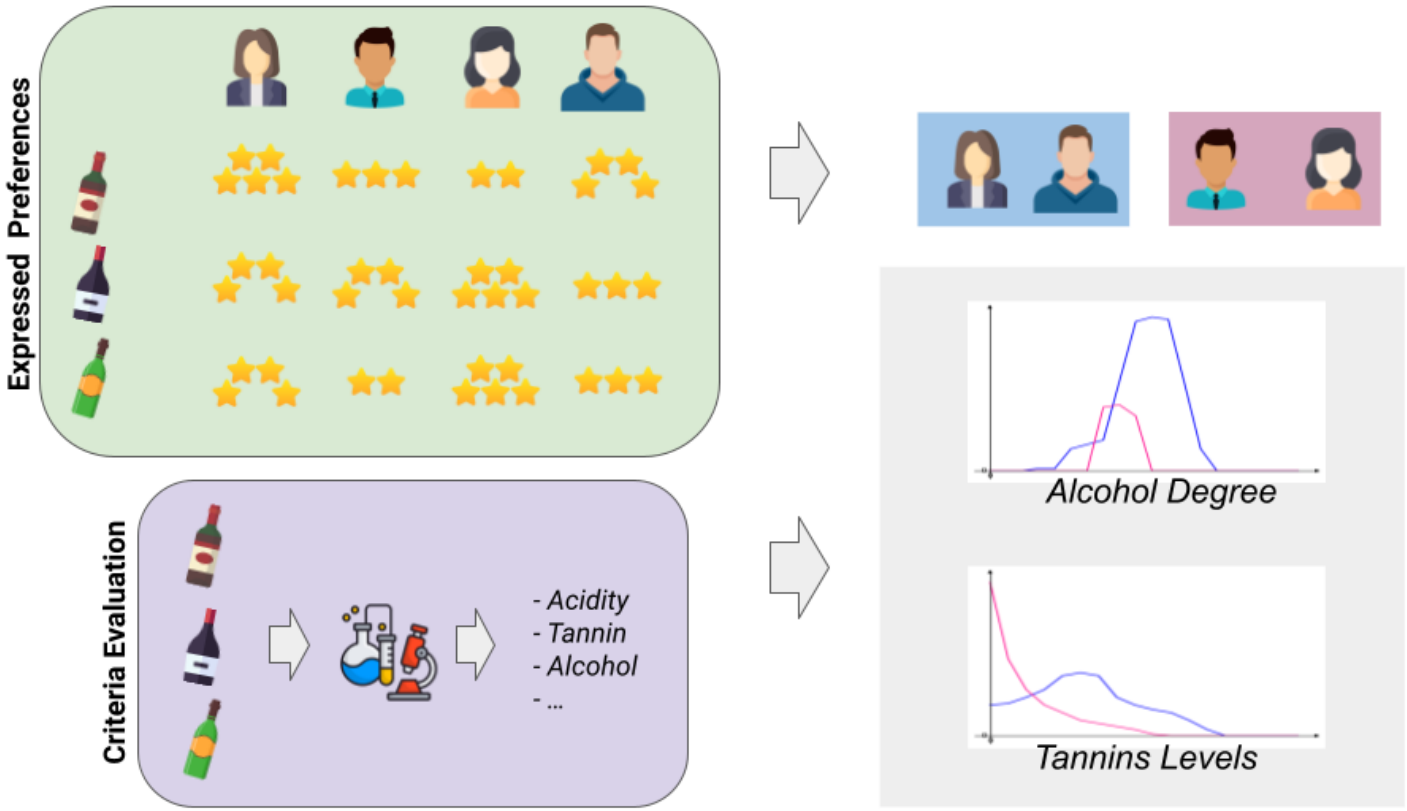
Conclusions & Future Work

A usual preference learning setup



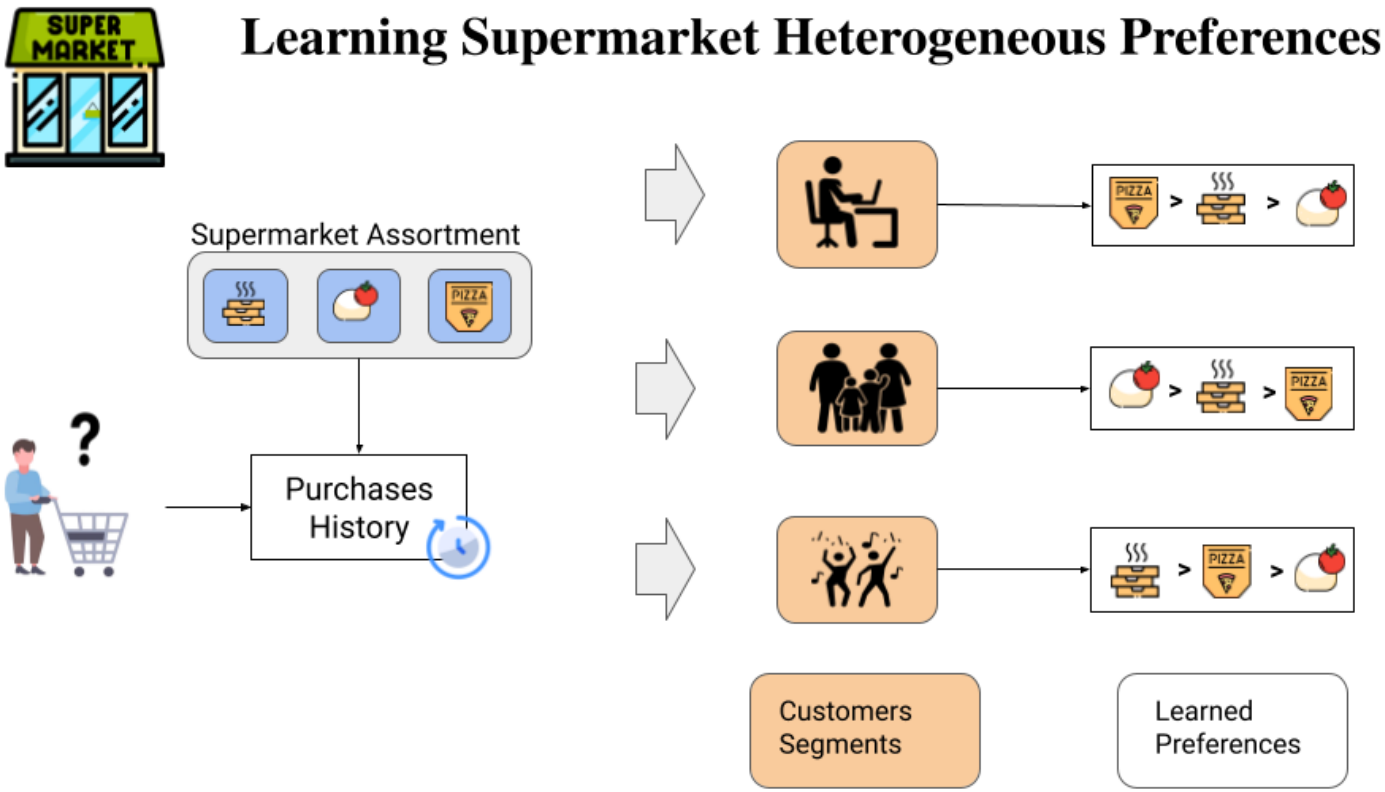
Learning Wine Preferences

Multiple decision makers with heterogeneous preferences



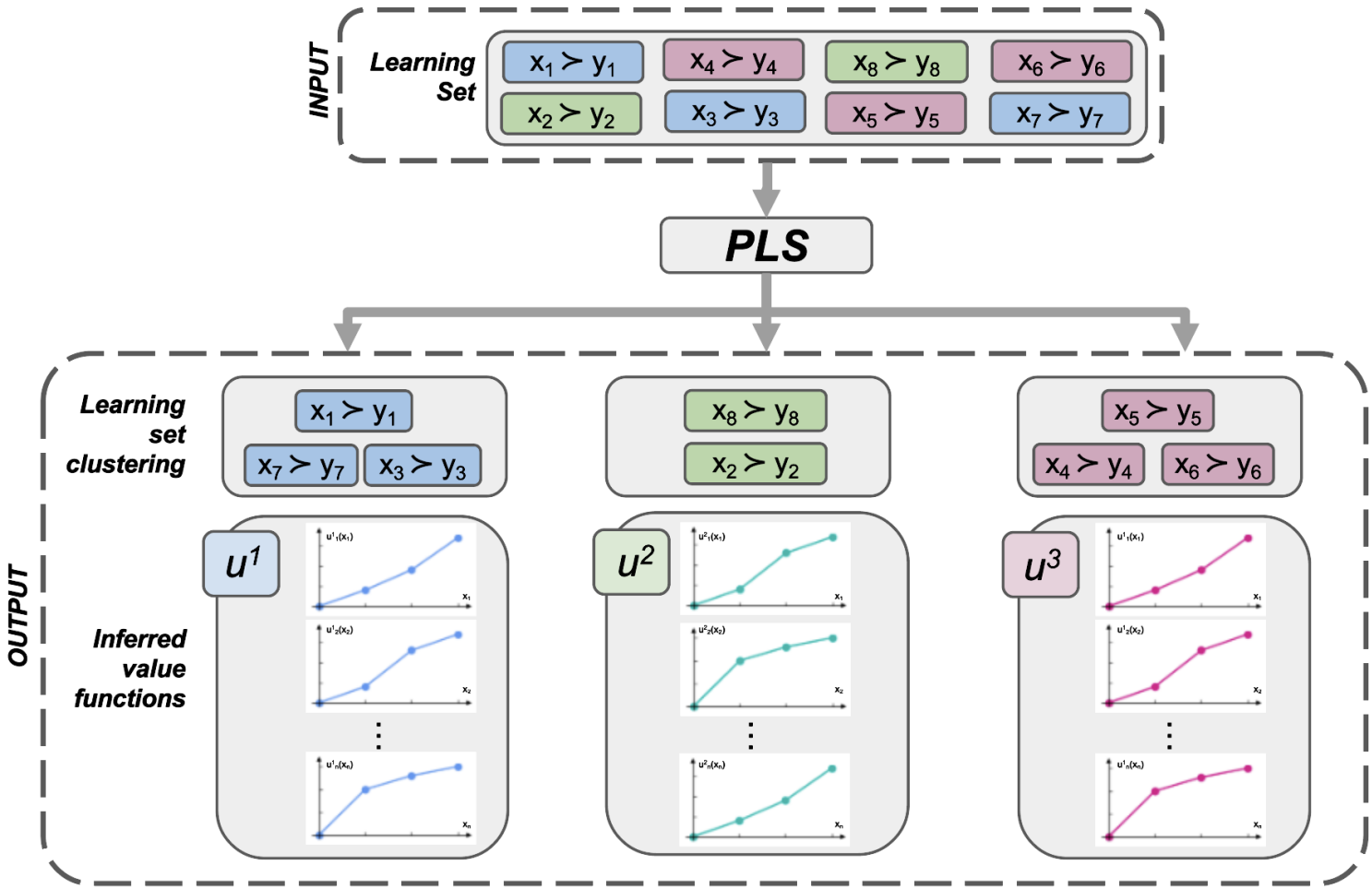
Learning Wine Heterogeneous Preferences

Multiple decision makers with heterogeneous preferences



Preference Learning and Segmentation (PLS)

- Partitioning preference statements
- Learning a preference model for each partition



Notation and Setting

Preference Learning and Segmentation
- Motivations
- Definition

Algorithm to learn from heterogeneous preferences
- Notations
- The UTA Model
- Proposed MILO formulation for PLS
- Proposed tractable heuristic for PLS

Numerical results
- Synthetic experiments
- Real-World experiments

Conclusions & Future Work

Problem setting

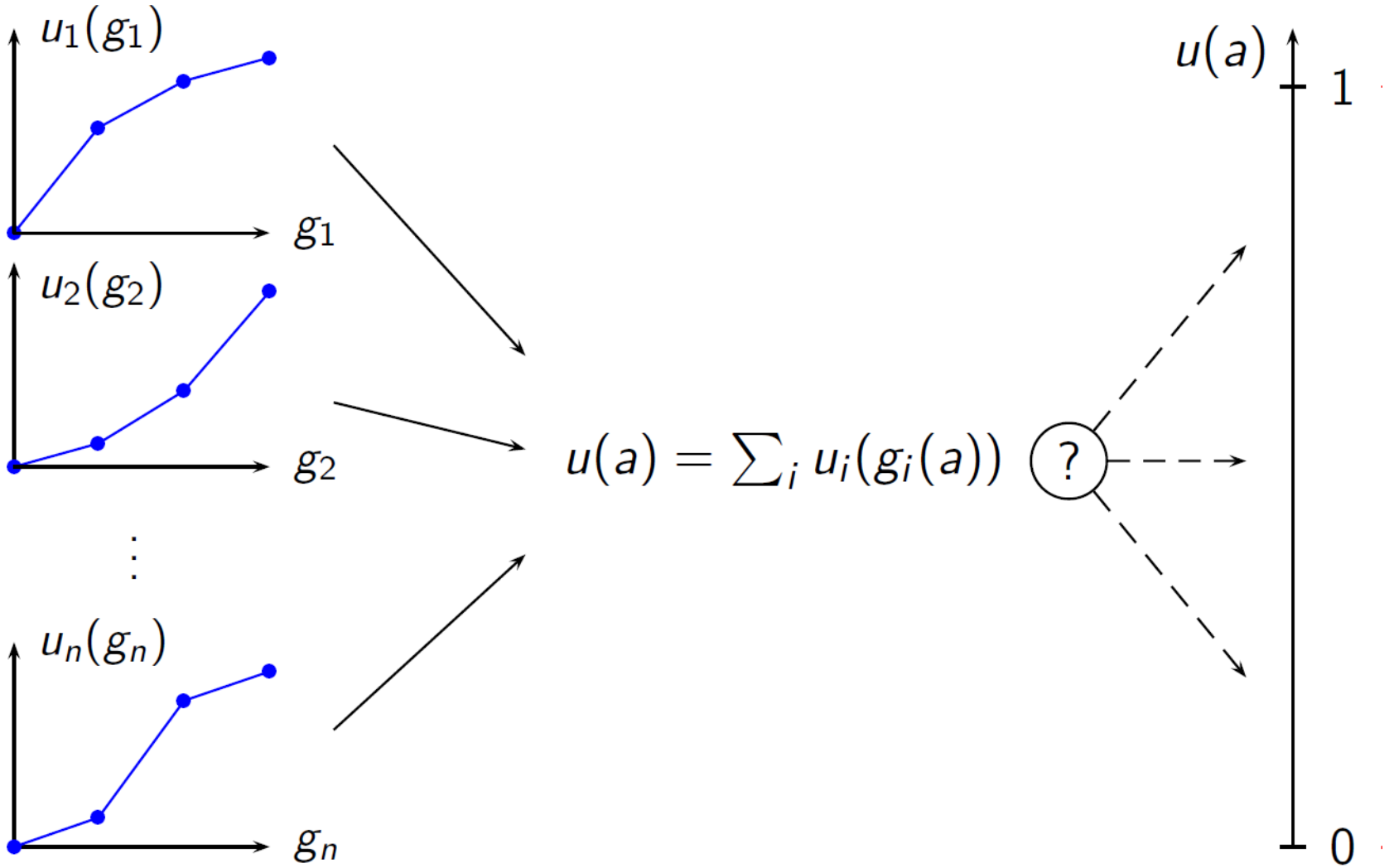
Observed Preferences

- Alternatives described on n criteria $i \in [\![1, ..., n]\!]$
- Learning Set : P preference pairs $(x^{(j)}, y^{(j)}) : x^{(j)} \succ y^{(j)}$
- $(x_1^{(j)}, ..., x_n^{(j)}) \succ (y_1^{(j)}, ..., y_n^{(j)}), j \in [\![1, P]\!]$

Expected result

- Segmentation of the learning set into K partitions
- A preference model for each partition : $\succ_k$
- $x^{(j)} \succ_k y^{(j)}$, at least for one k, for each pair in the learning set

The UTA Model [Jacquet-Lagrèze & Siskos, 1982]



The UTA Model [Jacquet-Lagrèze & Siskos]

Decision Variables

- u_i piece-wise linear with breaking points values $u_i(x_i^l)$
- $\sigma^{(j)} \geq 0$ comparison error

Constraints

- $u_i(x_i^0) = 0, \forall i \in [\![1, n]\!]$
- $\sum_{i=1}^{n} u_i(x_i^L) = 1$
- $u_i(x_i^{l+1}) \geq u_i(x_i^l), \forall i \in [\![1, n]\!], \forall l \in [\![0, L-1]\!]$
- $\sum_{i=1}^{n} u_i(x_i^{(j)}) - \sum_{i=1}^{n} u_i(y_i^{(j)}) + \sigma^{(j)} \geq 0, \forall j \in [\![1, P]\!]$

Objective

- $\min \sum_{j=1}^{n} \sigma^{(j)}$

Extension to K Clusters - MILO formulation

- Problem becomes NP-hard *See proof in the paper from K-colorability*

Additional Decision Variables

- $z_k^{(j)} \in \{0,1\}$, $z_k^{(j)} = 1$ if the pair $(x^{(j)}, y^{(j)})$ belongs to cluster k

Constraints

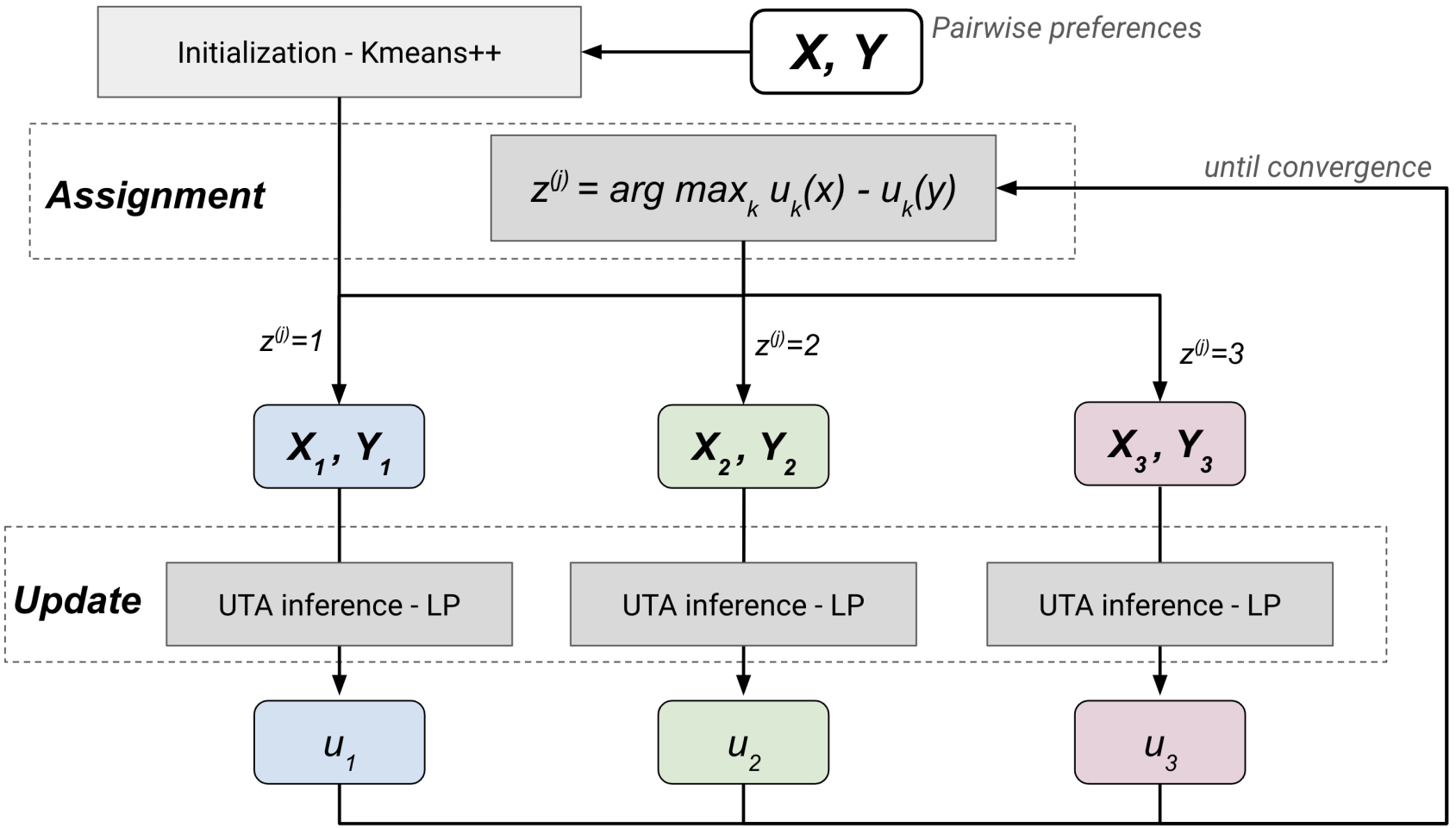
- $\sum_{k=1}^{K} z_k^{(j)} \geq 1, \forall j \in [\![1, k]\!]$
- $u_k(x^{(j)}) - u_k(y^{(j)}) + M(1 - z_k^{(j)}) + \sigma^{(j)} \geq 0$

$\forall j \in [\![1, P]\!], \forall k \in [\![1, K]\!]$

Objective

- $\min \sum_{j=1}^{n} \sigma^{(j)}$

Extension to K Clusters - Tractable Heuristic



Heuristic : Algorithm

for $j = 1$ to$P$ **do** {Random Initialization of Partition}
$z^{(j)} \leftarrow Random((1, K))$
end for
while Not Converged **do**
for $j = 1$ to$P$ **do** {"Assignment"}
$z^{(j)} \leftarrow \arg\max_{k \in [\![1,K]\!]} u_k(x^{(j)}) - u_k(y^{(j)})$ {Assign pair j to the cluster that best represents it}
end for
for $k = 1$ to$K$ **do** {"Models inference"}
$u_k \leftarrow UTA(\{(x^{(j)}, y^{(j)}) | z^{(j)} = k\})$ {Infer UTA model u_k}
end for
end while

- LP is scalable (while MIP is not)

Numerical Results

Preference Learning and Segmentation
- Motivations
- Definition

Algorithm to learn from heterogeneous preferences
- Notations
- The UTA Model
- Proposed MILO formulation for PLS
- Proposed tractable heuristic for PLS

Numerical results
- Synthetic experiments
- Real-World experiments

Conclusions & Future Work

Experimental Settings

Empirical Questions

- Do the models learn the heterogeneous preferences ?
- Are the algorithms scalable ?

Method

Data Generation

- Randomly draw K additive piecewise linear functions
- Randomly draw $\frac{P}{K}$ pairs $x^{(j)} \succ_k y^{(j)}$ for each cluster k
- Vary $K \in \{2, 3, 4\}$, $P \in \{128, 256, 512, 1024, 2048, 4096\}$, $n \in \{2, 4, 6, 8, 10\}$
- 100 random repetitions of the procedure

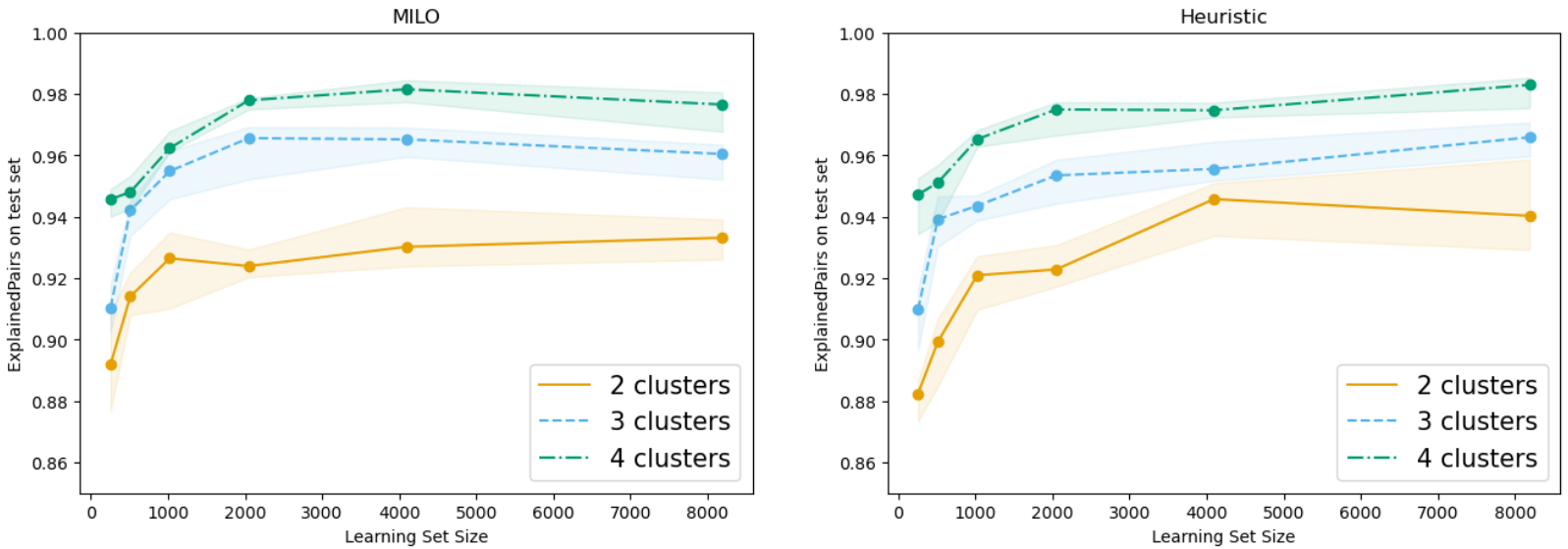
Resolution

- Set the number of clusters to the correct value
- Solve instance with MIP
- Solve instance with heuristic

Evaluation

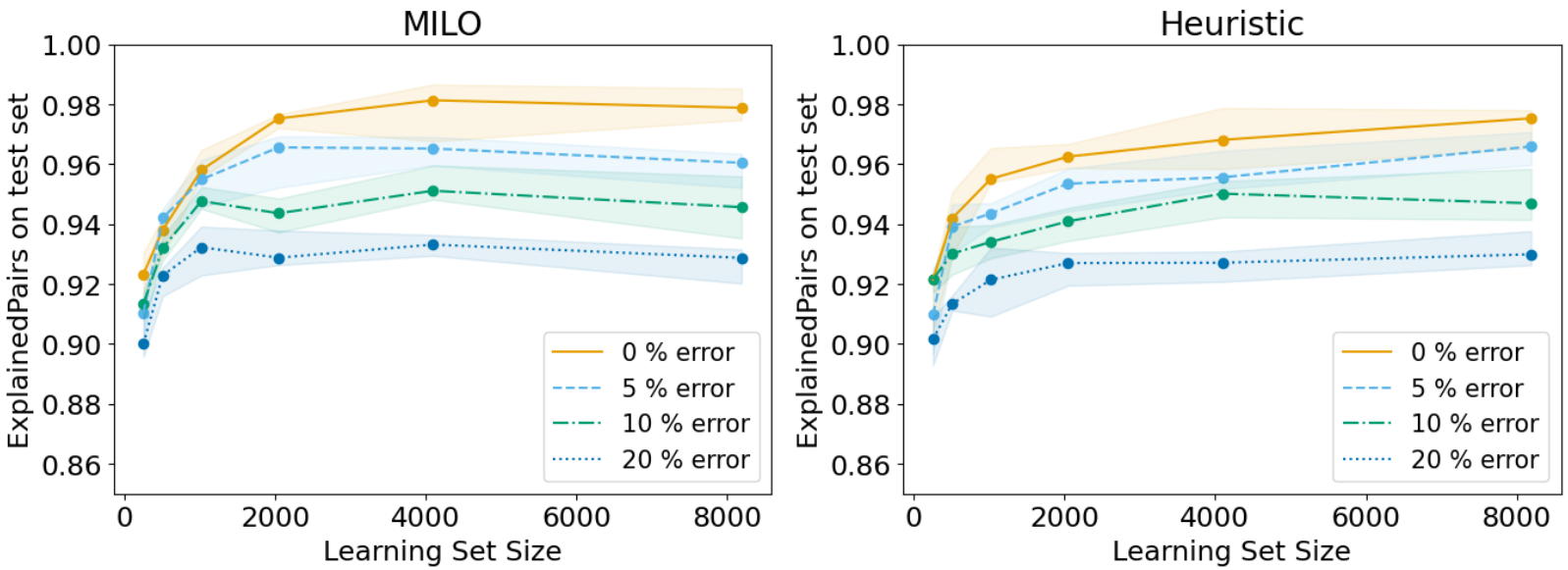
- Assess computing time
- Evaluate inferred models on a test set

Restoring Preferences



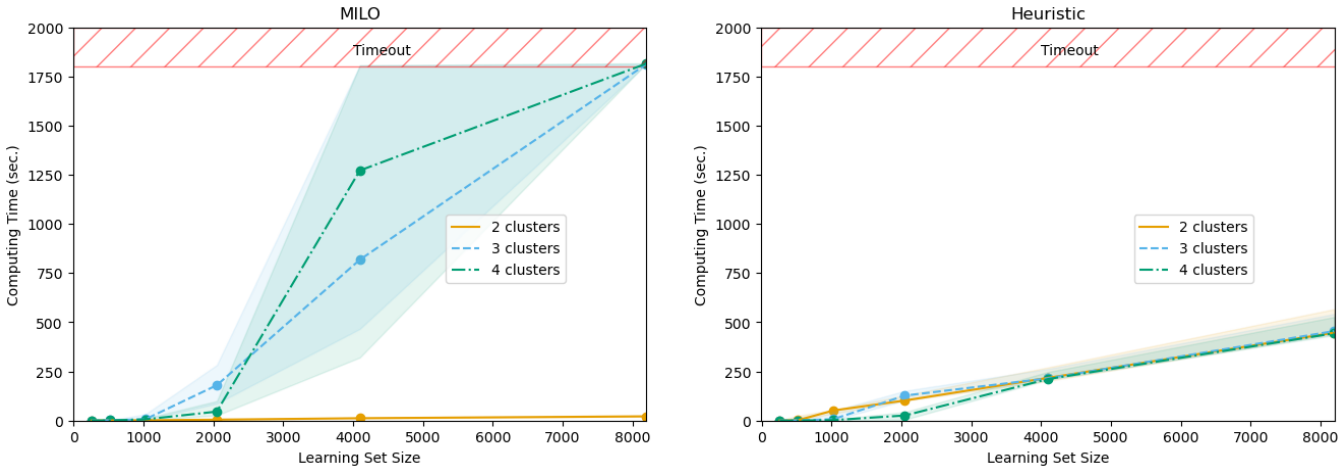
Proportion of ExplainedPairs on test set with 8 criteria, 5% noise and different cluster numbers. Shaded areas represent the values of the first and 3rd quartiles over experiment repetitions

Noise robustness

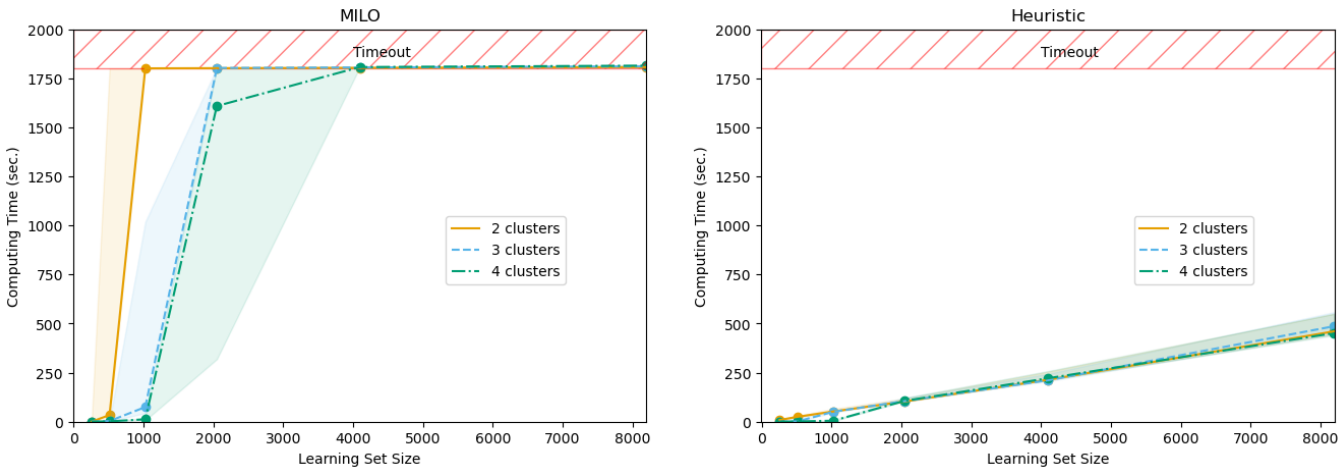


Proportion of ExplainedPairs on test set for 3 clusters, 8 criteria and different noise levels. Shaded areas represent the values of the first and 3rd quartiles over experiment repetitions

Computing Time : MILO vs Heuristic



Comparison of approaches for 2 clusters, 0% noise in data



Comparison of approaches for 2 clusters, 5% noise in data

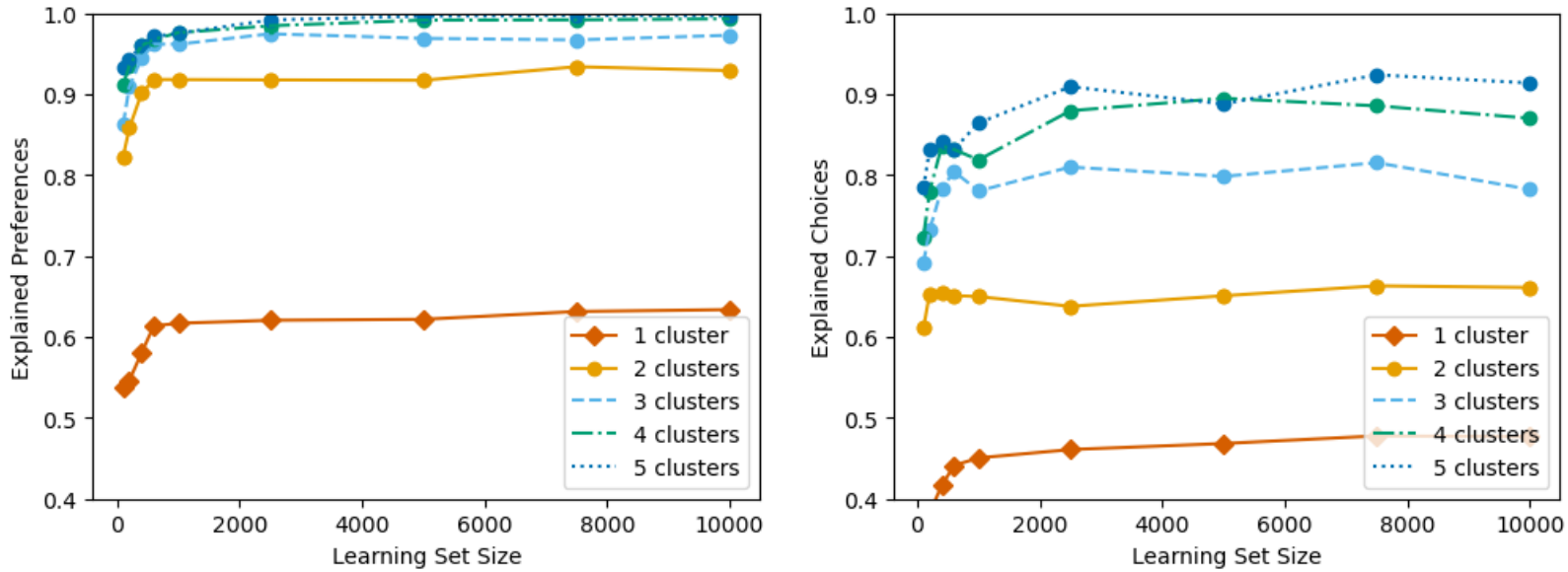
Car preferences dataset

Criterion	Objective
price of vehicule	min
range between refuelings	max
acceleration	min
maximum speed	max
tailpipe emissions	min
luggage space	max
cost per mile	min
Fraction of stations that can refuel/recharge vehicle	max

Table – Car preferences criteria description.

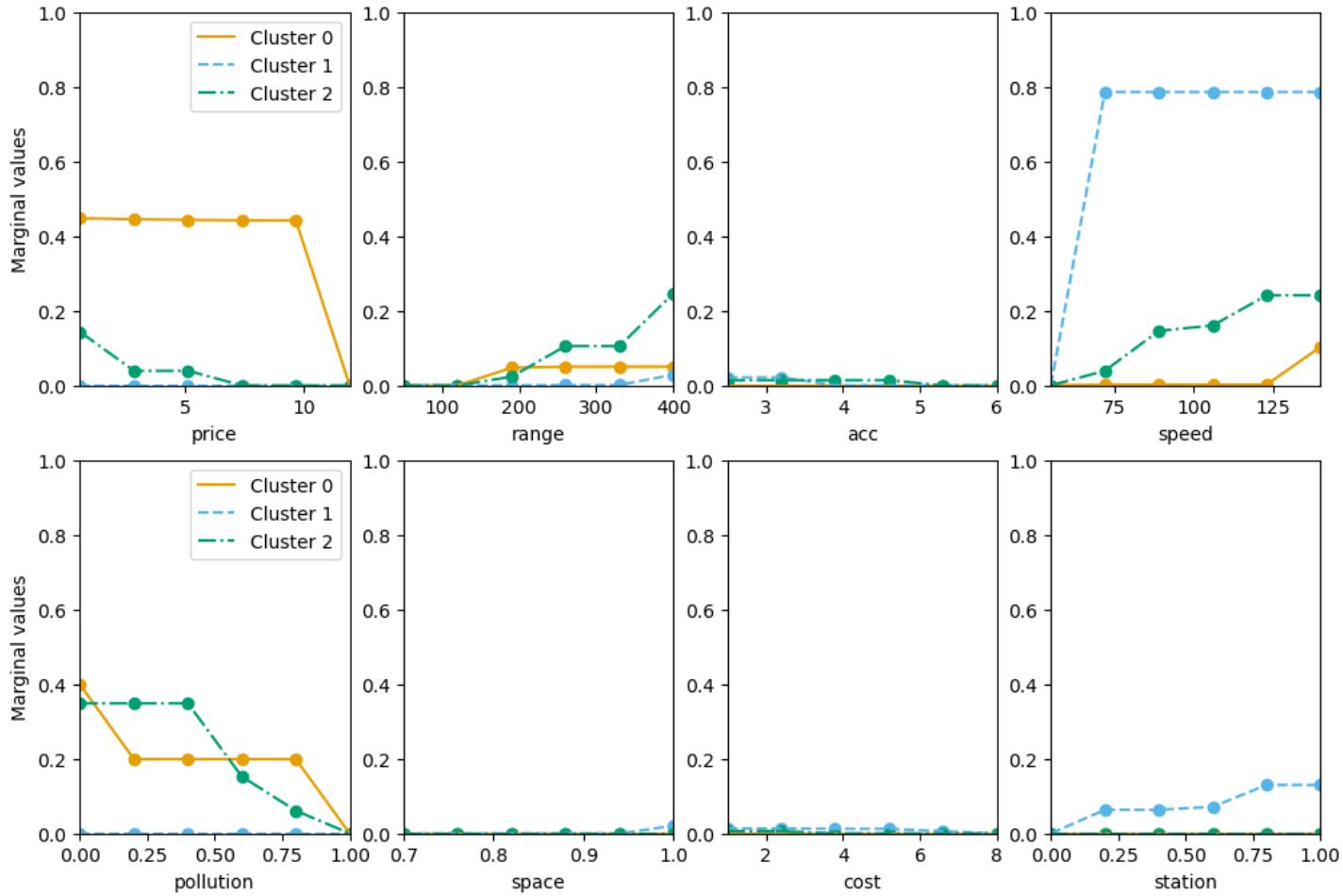
-> Six alternatives presented, one chosen

Model performances with different number of clusters



Average *ExplainedPairs* and *ExplainedChoices* for the heuristic and different settings.

Results example



Example of resulting UTA marginal values on the stated car preferences dataset.

Conclusions & Future Work

Preference Learning and Segmentation
- Motivations
- Definition

Algorithm to learn from heterogeneous preferences
- Notations
- The UTA Model
- Proposed MILO formulation for PLS
- Proposed tractable heuristic for PLS

Numerical results
- Synthetic experiments
- Real-World experiments

Conclusions & Future Work

Conclusions & Future Work

Conclusions

- Formulation of clustering + preference learning problem
- A MIP formulation yields satisfactory results on small instances
- A simple heuristic approximates the results on larger instances

Future Research

- Improvement of the heuristic
- Extension to non-monotone criteria & sorting problems
- Get some insight on number of clusters choice

Thank you for listening !





Questions

Heuristic for scalability

Initialization

- Random partition of the pairs in K clusters

Expectation-like step

- Assignment of each pair to a cluster
- cluster $k = \arg\max_{k \in [\![1,K]\!]} u_k(x^{(j)}) - u_k(y^{(j)})$

Maximization-like step

- Estimation of a UTA model on each cluster
- LP is scalable (while MIP is not)

Assessing the quality of the partition

$$RandIndex = \frac{TP + TN}{TP + FP + FN + TN}$$

with :

TP	Nb of pairs in the same cluster in Truth and Estimation
TN	Nb of pairs in different clusters in Truth and Estimation
FN	Nb of pairs in different clusters in Truth and same in Estimation
FP	Nb of pairs in same clusters in Truth and different in Estimation