The 25th International Conference on Artificial Intelligence and Statistics

AIStats-2022, 28-30.03.2022

k-Pareto Optimality-Based Sorting with Maximization of Choice

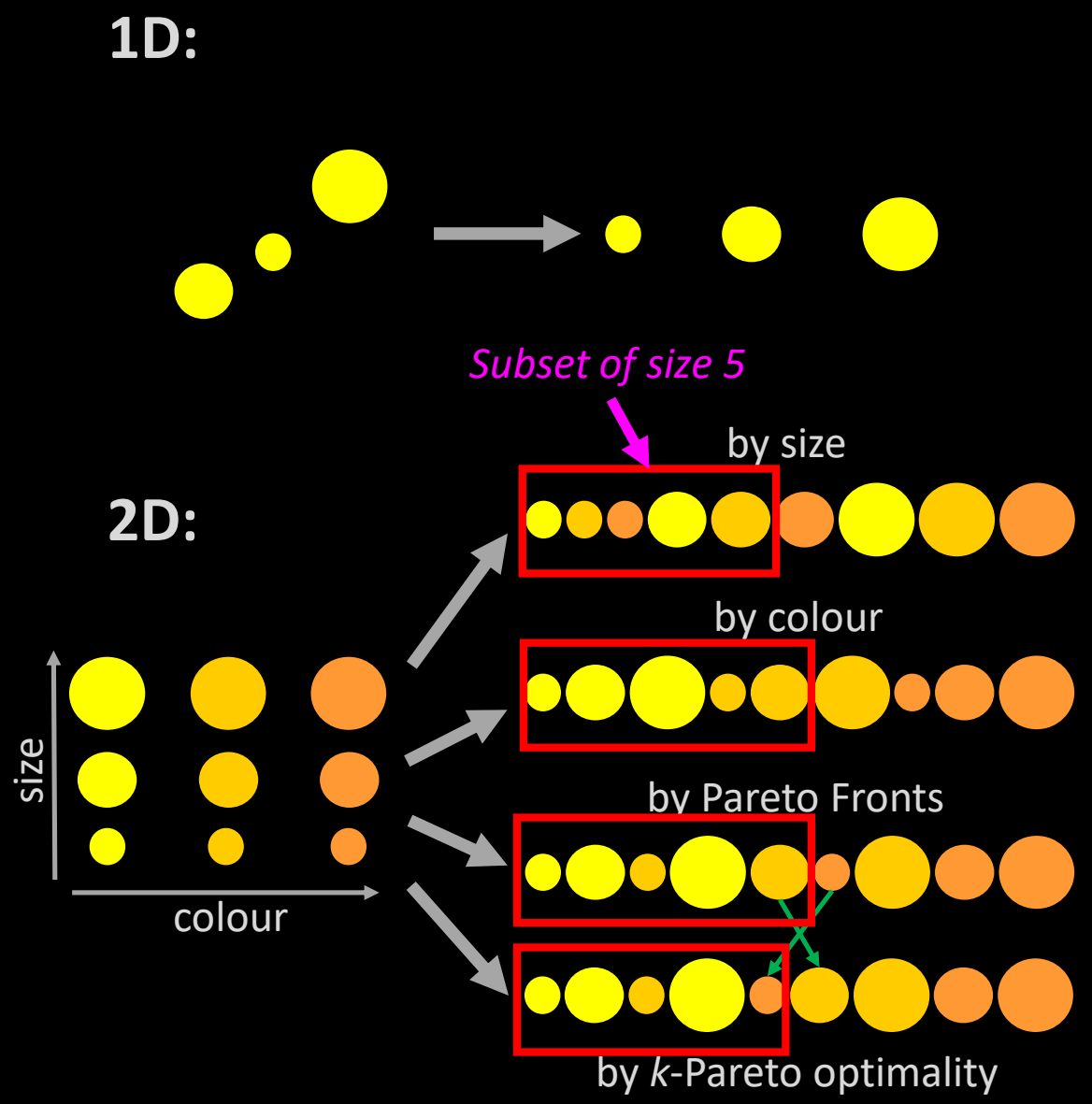
Jean Ruppert
jean.ruppert@mathcomp.lu
Mathematics and Computing S.à.r.l.
Luxembourg

Dr. Marharyta Aleksandrova
marharyta.aleksandrova@uni.lu
University of Luxembourg

Prof. Dr. Thomas Engel
thomas.engel@uni.lu
University of Luxembourg

Introduction

- Topological sorting is important in many applications
  - Challenging in high dimensional space
- Choosing a subset of best elements with maximum diversity (choice)? ➔ Sorting with choice maximization
  - Can be useful in practice: recommender systems, selection step of genetic algorithms, scheduling, …
  - NOVEL: Was not considered in the literature before



k-Pareto optimality

$$po(x) = \mu(\{y | yR^*x\})$$

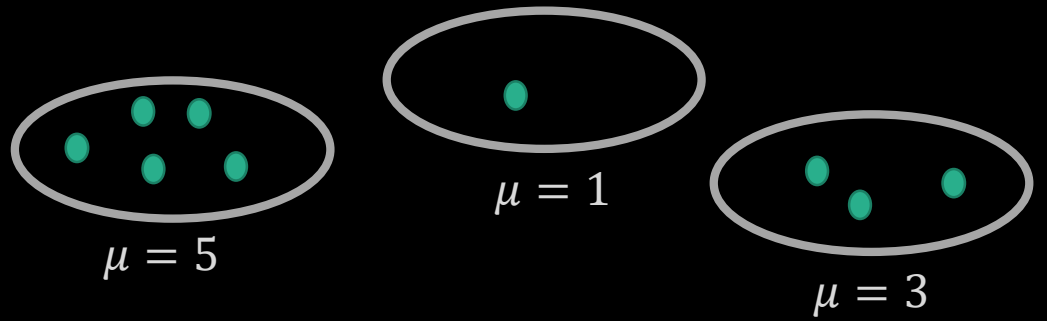
- R – preference relation; defines best
  - yRx – y is preferable to x
  - yR^*x – y is strictly preferable to x
- Set of elements $x \in X$ can be discrete or continuous
- μ – measure indicating set size
  - Counting
  - Probability
- Choice – measure of incomparable pairs;

Choice = 1 

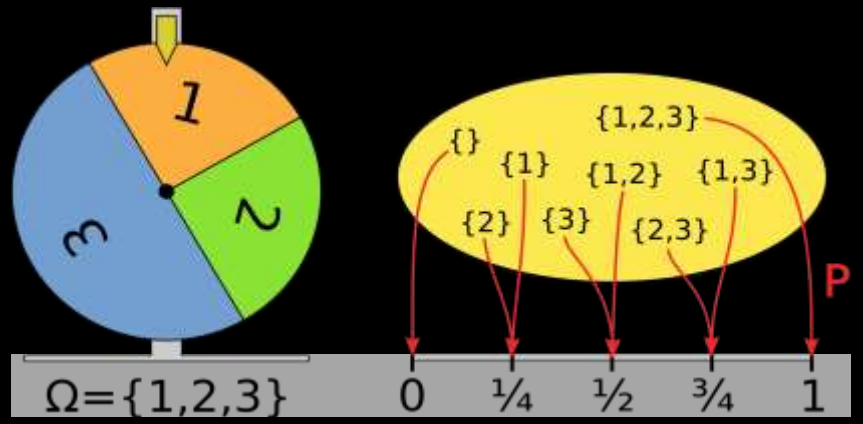
Choice = 0 

- Choice is functionally related to diversity

Counting measure:



Probability measure:



Source: https://en.wikipedia.org/wiki/Probability_measure#/media/File:Probability-measure.svg

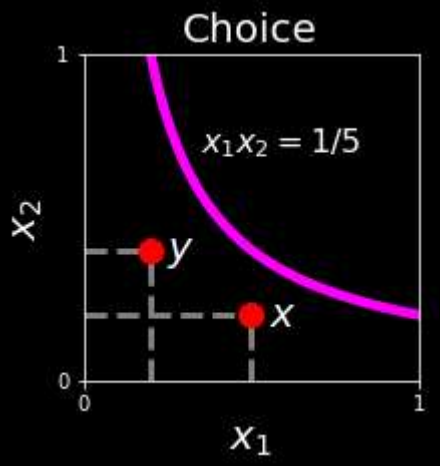
Maximum choice problem and solution

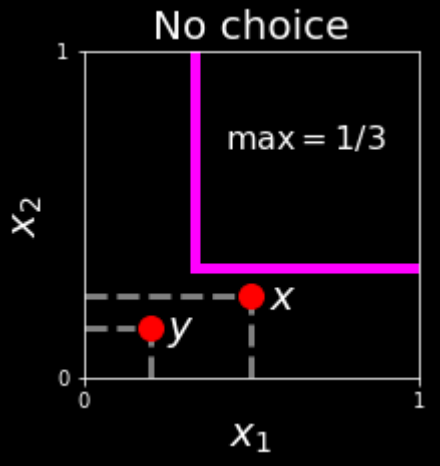
- Maximum choice problem:
  Choose a subset of best elements of a given measure with maximum choice (diversity).

- Theorem: Sorting by *k*-Pareto optimality is a solution.

- Direct consequence
  Characterization of hyperbola: among all descending functions delimiting an area of a fixed size, the arc of hyperbola provides the highest probability of the event that 2 points randomly chosen from the delimited area are incomparable (offer choice).

Characterization of hyperbola





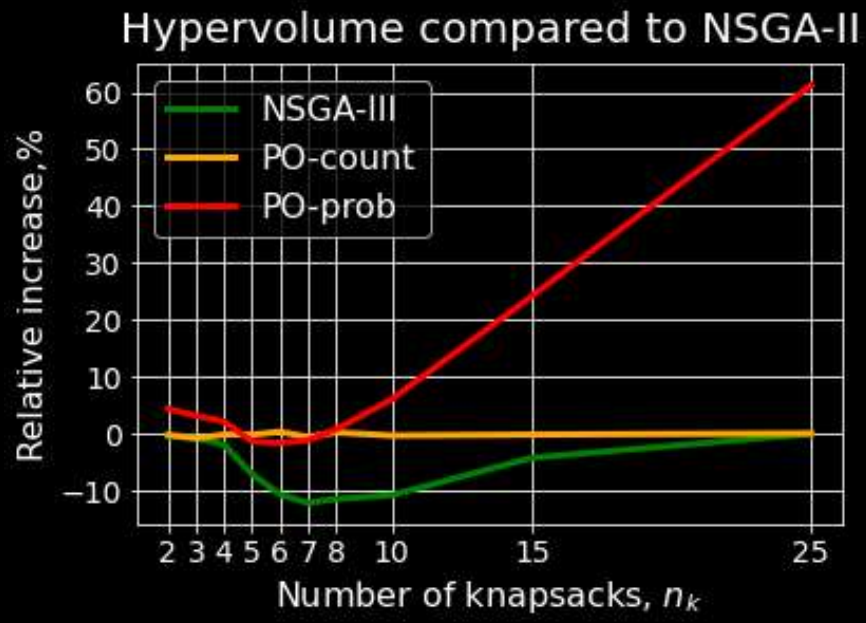
Applications: genetic optimization

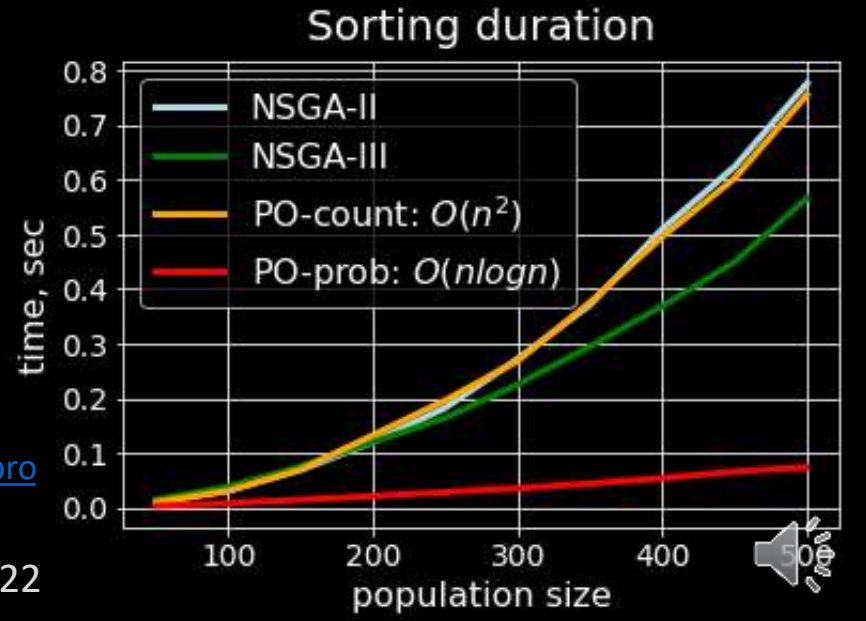
Sorting with choice (diversity) maximization is beneficial for genetic optimization

Experimental setup:

- NSGA-II with modified sorting:
  - μ - counting: *PO-count*
  - μ - probability: *PO-prob*
- 0/1 knapsack problem with ind. obj.
- 30 runs
- Baseline:
  - NSGA-II
  - NSGA-III
- Implementation: https://github.com/marharyta-aleksandrova/kPO

Image source: https://en.wikipedia.org/wiki/Knapsack_problem#/media/File:Knapsack.svg





Ruppert J., Aleksandrova M., Engel T., AIStats-2022

Future work

1. Further theoretical analysis:
   - Behaviour for different types of elements, relations and measures

2. Other applications:
   - Constrained genetic optimization
   - Genetic optimization of complex structures:
     - Decision trees
     - Artificial neural networks
     - …
   - Recommender systems
   - Nonparametric statistical tests