

Metaheuristics for Multi-objective Optimization

July 2007

E.-G. Talbi and the ParadisEO team

 **ParadisEO**

<http://paradisEO.gforge.inria.fr>
paradisEO-help@lists.gforge.inria.fr

Outline

- Multi-objective optimization: definitions, problems, etc
- A taxonomy of the resolution methods
- Landscapes and performance analysis
- Conclusion

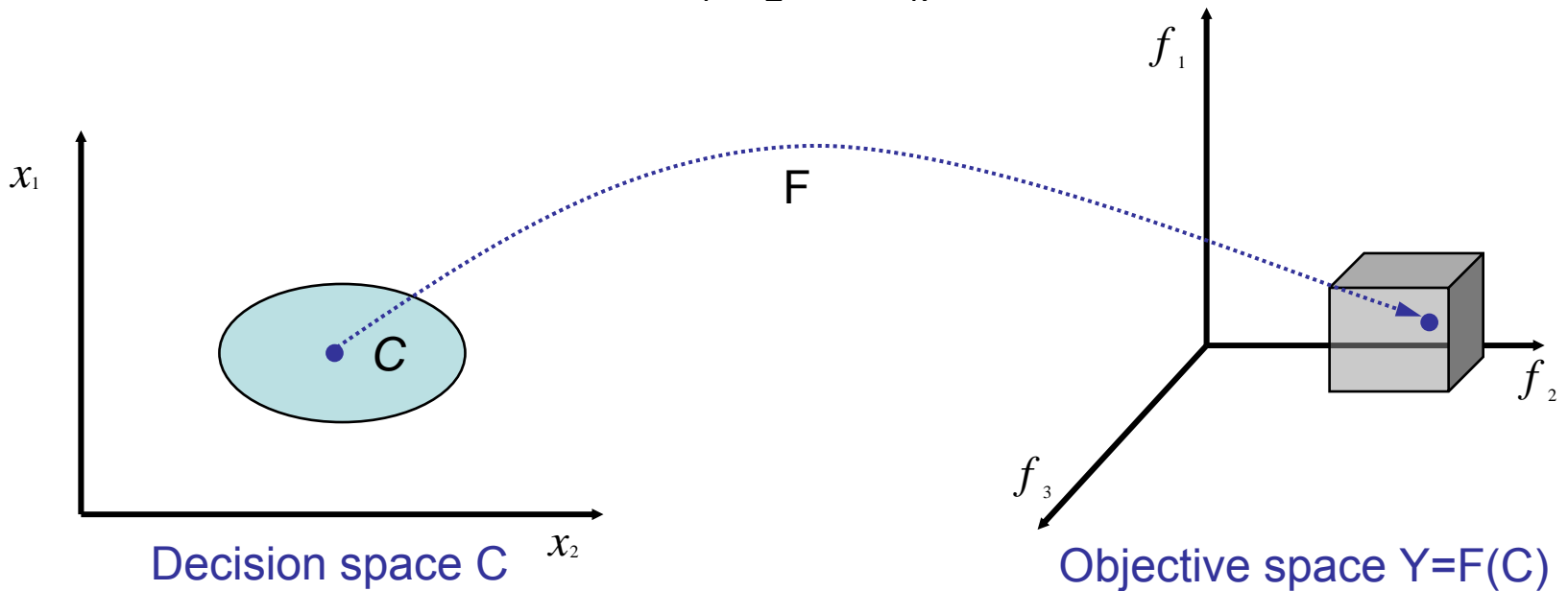
Multi-objective Optimization

- Many areas of industry are concerned
 - Telecommunication
 - Transport
 - Aeronautics
 - ...
- Roots in the 19th century in **economy** from Edgeworth and Pareto
 - Management
 - Engineering
 - ...
- Linear/Non linear Multi-objective Optimization [Steuer 86, White 90]
- Multi-objective Combinatorial Optimization

Multi-objective Optimization

$$(MOP) \begin{cases} \min F(x) = (f_1(x), f_2(x), \dots, f_n(x)) & n \geq 2 \\ \text{s.t. } x \in C \end{cases}$$

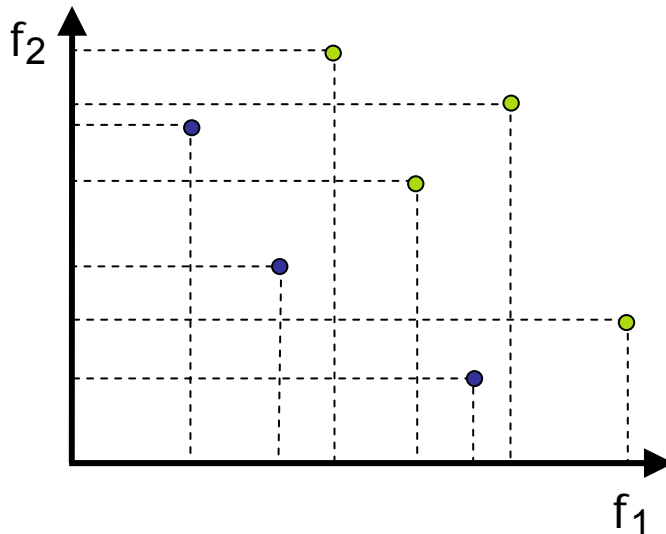
Decision variables $x = (x_1, x_2, \dots, x_k)$



Definitions

Dominance: A solution x_1 dominates a solution x_2 if and only if:
 $\forall i \in [1, \dots, n], f_i(x_1) \leq f_i(x_2)$ and $\exists i \in [1, \dots, n], f_i(x_1) < f_i(x_2)$

Non-dominated solution: A solution x is non-dominated if a solution which dominates x does not exist

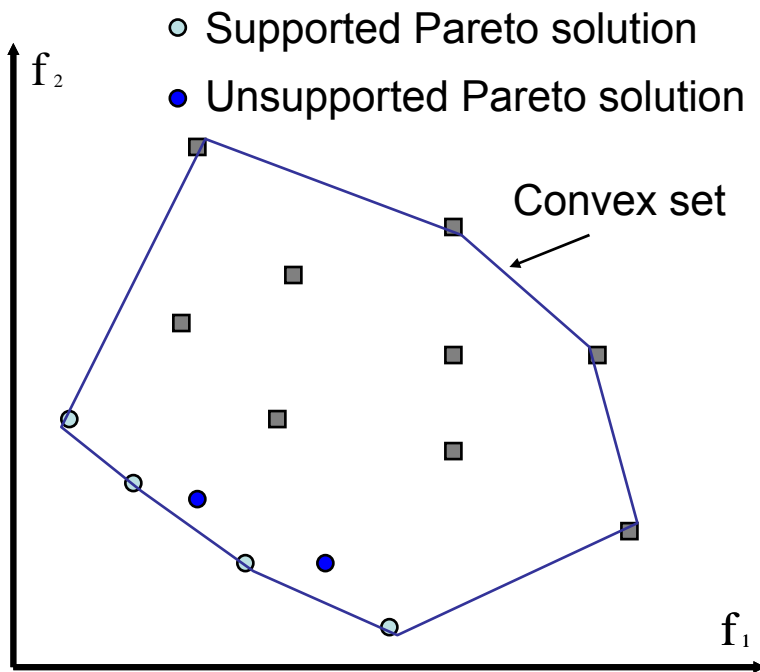


- Non-dominated solution (eligible, efficient, non inferior, Pareto optimal)
- Dominated feasible solution

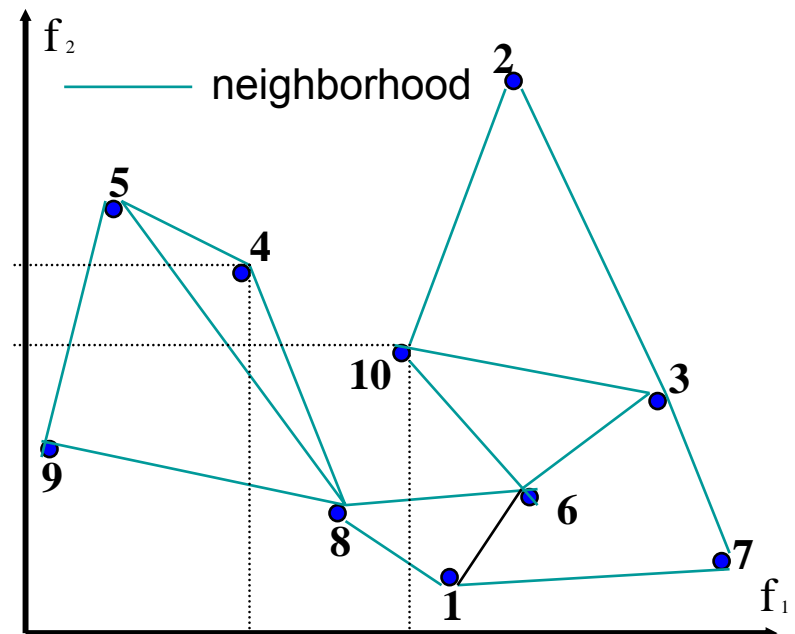
Goal: To find a well-converged and well-diversified Pareto set

Definitions

- Supported / Unsupported solutions
- Optimal vector y^*
Reference vector z^*



- **Locally Pareto optimal**
- Neighborhood N: Local search
 - 1,8,9 : Pareto optimal
 - 4,10 : Locally Pareto optimal



Main Issues

- **What is optimality?** Partial order, dynamic environment, the last choice belongs to the decision maker
- **Number of Pareto solutions increases** function of the dimension(s) of the tackled problems and the number of objectives
- For **non-convex MOPs**, Pareto solutions are located on the frontiers and in the convex hulls

Solver-Decision Maker Cooperation

- A priori

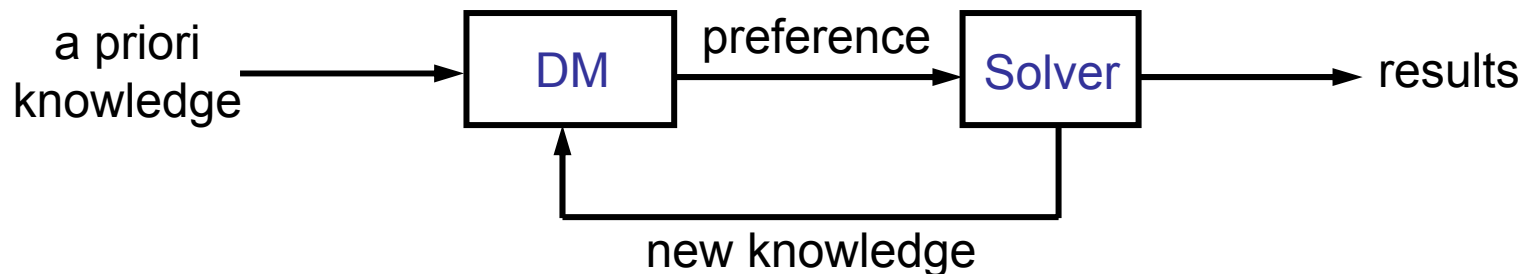
- Preference → Search
- A priori knowledge of the tackled problem

- A posteriori

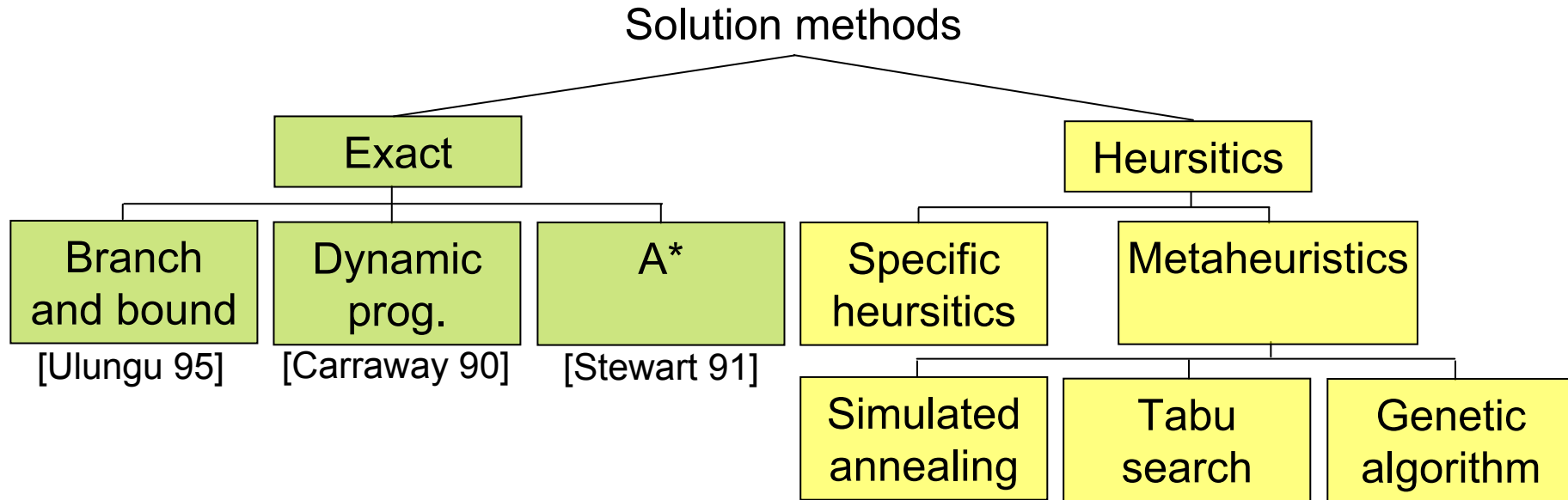
- Search → Preference
- Cardinality of the reduced PO set

- Interactive

- Search ↔ Preference
- Decision helping is not addressed



Taxonomy of Optimization Methods



- **Exact algorithms:** bi-criterion problems of small size
- **Metaheuristics:**
 - **Solution-based:** simulated annealing, tabu search, ...
 - **Population-based:** genetic algos, scatter search, ants systems, ...

PO*: Approximation of PO

Academic MOPs

- Scheduling [Sayin et al. 99]
- Flow [Warburton 87], Spanning tree [Zhou et Gen 99]
- TSP [Serafini 92], Affectation [Teghem 95]
- Vehicle Routing [Park et Koelling 89], ...
- Bag packing [Teghem et al. 97]:

$$\left\{ \begin{array}{l} \text{Max } (f_i(x)) = \sum_{j=1}^m c_j^i x_j \\ (i = 1, \dots, n), x \in C \\ C = x / \sum_{j=1}^m w_j x_j < W \\ x_j \in \{0, 1\}, \forall j = 1, \dots, m \end{array} \right.$$

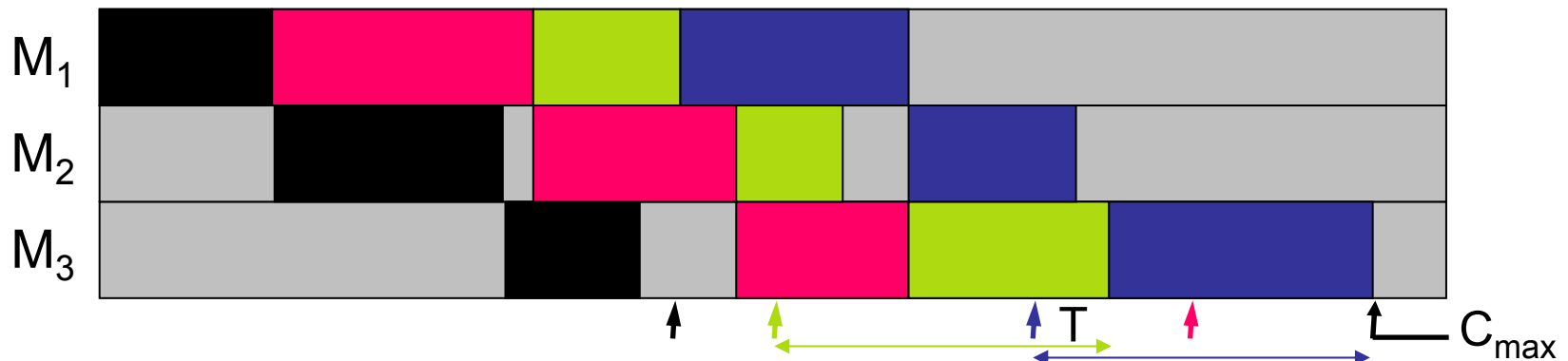
$$x_j = \begin{cases} 1 & \text{if } j \text{ is in the bag} \\ 0 & \text{else} \end{cases}$$

w_j : weight of j

c_j^i : usefulness of j / objective i

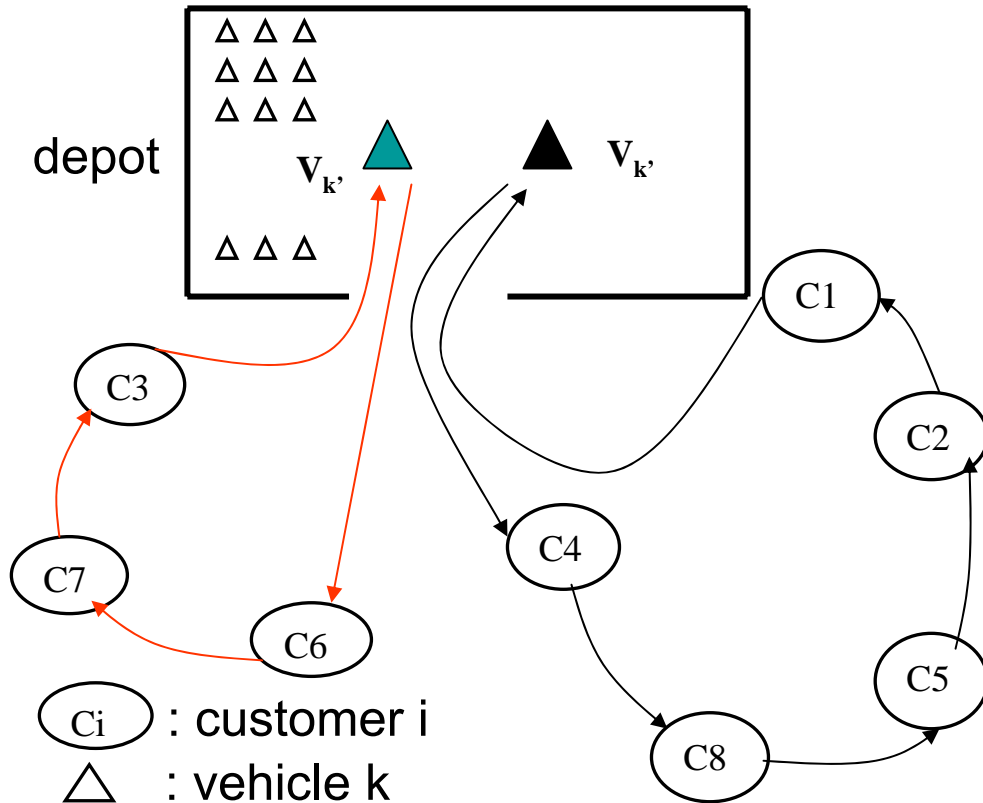
Bi-objective Flow-shop Scheduling Problem

- Scheduling problem $F/\text{perm}, d_i/(C_{\max}, T)$ [Graham 79]
- N jobs to schedule on M machines
- 2 objectives to minimize
 - Makespan (C_{\max})
 - Total tardiness (T)



M. Basseur, F. Seynhaeve, E-G. Talbi: “**Design of multiobjective evolutionary algorithms: Application to the flow-shop scheduling problem**”, *CEC'2002*, pp. 1151-1156, Hawaii, 2002

Bi-objective VRP



A multi-objective problem
by nature

Bi-objective model:

- First criteria

- Minimizing the tour length

$$\sum_{i,j=1}^N d_{ij} \sum_{k=1}^M x_{ijk}$$

- Second criteria

- Minimizing the balance

$$\max_k \sum_{i,j=1}^N d_{ij} x_{ijk} - \min_{k'} \sum_{i,j=1}^N d_{ij} x_{ijk'}$$

N. Jozefowiez, F. Semet, E.-G. Talbi: "The bi-objective covering tour problem", *Computers and Operations Research (COR)*, vol. 34, pp. 1929-1942, 2007

Industrial MOPs

- Telecommunications
 - Frequencies assignment [Dahl et al. 95], Wire antenna geometry design [Sandlin et al. 97]...
- Bioinformatics
- Aeronautics
 - Wings [Obayashi 98], Motors [Fujita 98]...
- Environment
 - Air quality [Loughlin 98], Water distribution...
- Transports
 - Motorway layout , Container Management [Tamaki 96]...
- Finances, Computer-integrated manufacturing, Robotics, Mechanics...

Cellular Network Design

[France Telecom]

- **Network design** (antennae on sites, configuration of antennae)
- Objectives
 - Minimizing the **number of sites**
 - Minimizing **interferences**
 - Minimizing **traffic loss**
 - Maximizing **resource use**
- Constraints
 - **Covering**
 - **Handover**

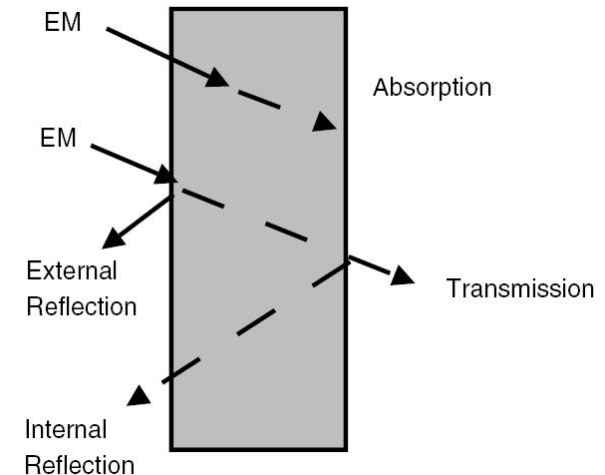


E.-G. Talbi, H. Meunier: “**Hierarchical parallel approach for GSM mobile network design**”,
Journal of Parallel and Distributed Computing, vol. 66(2), pp. 274-290, 2005

Design of Conducting Polymer Composites for Electromagnetic Shielding

4 objectives:

- Maximizing the **electromagnetic shielding**
- Min/Max the **high reflecting coefficient**
- Minimizing the **mass percentage**
- Minimizing the **thickness**



O. Schuetze, L. Jourdan, T. Legrand, E.-G. Talbi, J.-L. Wojkiewicz: **"A Multi-Objective Approach to the Design of Conducting Polymer Composites for Electromagnetic Shielding"**, EMO'2007, LNCS vol. 4408, pp. 590-603, Matsushima, Japan, 2007

Knowledge Discovery in Biological Data

- Exponential growth of amount of **available data**
- Analyzing **microarray** is a major issue in **genomics**
- Often used techniques
 - **Clustering and classification**
- **New high throughput technologies**
 - **Association rules**

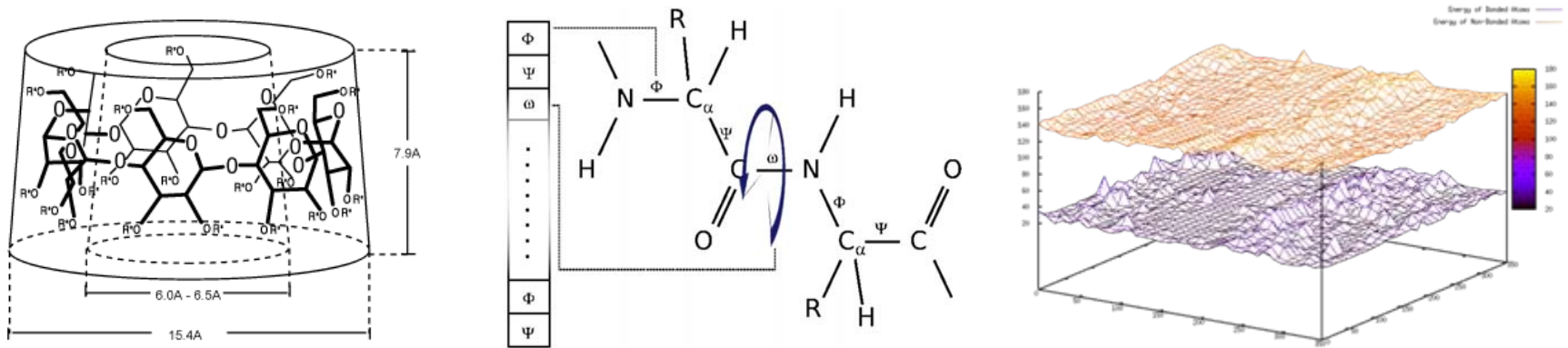


→ **Multi-objective rule-mining problem**

L. Jourdan, M. Khabzaoui, C. Dhaenens, E.-G. Talbi: “**A hybrid metaheuristics for knowledge discovery in microarray experiments**”, *Handbook of Bioinspired Algorithms and Applications*, Chapter 28, pp. 489-505, CRC Press, USA, 2005

Bicriterion Molecular Structure Prediction

- A molecule with 40 residues and 10 conformations per residue
 - 1040 conformations are obtained (in average)
 - 1018 years are required (with 1014 conformations per second)
- Island-deployed Lamarckian genetic algorithm



A.-A. Tantar, N. Melab, E.-G. Talbi, B. Toursel: “Solving the Protein Folding Problem with a Bicriterion Genetic Algorithm on the Grid”, *BioGrid'06*, Singapore, 2006

A Taxonomy of the Resolution Methods

Methods classification

- **Scalar approaches**
 - Aggregation method
 - ϵ -constraint method
 - Goal programming method
- **Criterion-based approaches**
 - Each objective is treated separately
- **Pareto-based approaches**
 - The concept of Pareto dominance is used
- **Indicator-based approaches**
 - The optimization goal is given in term of a quality indicator

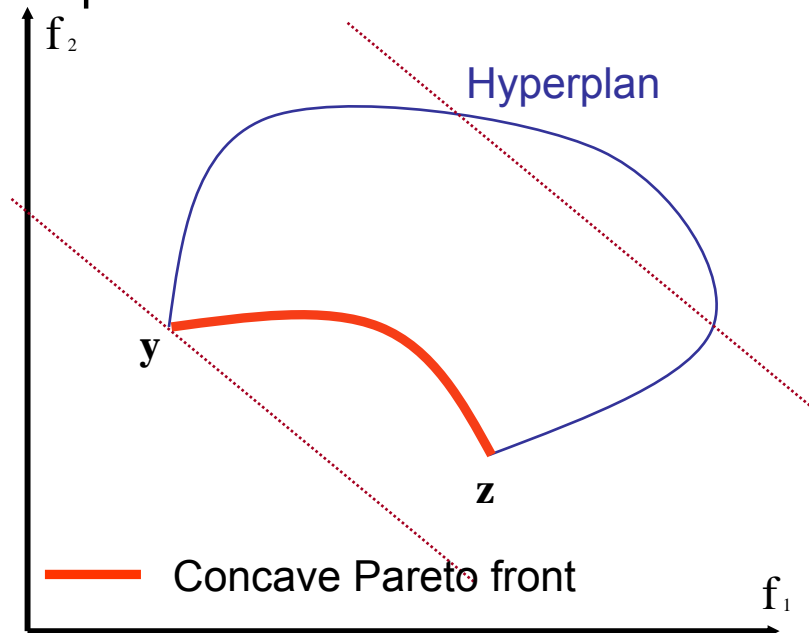
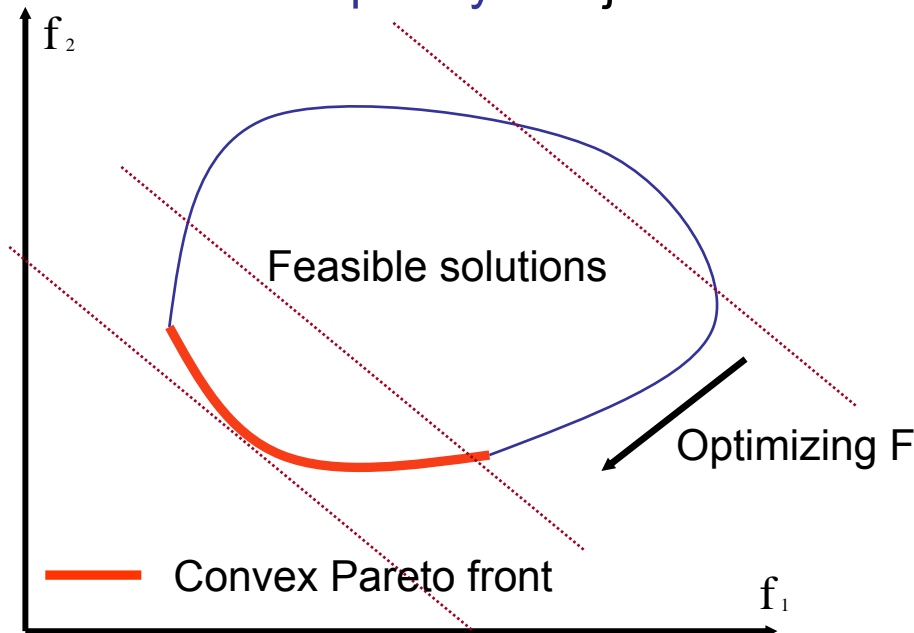
Aggregation Approach

[Gass & Saaty 55]

$$(MOP_w) \begin{cases} \min F(x) = \sum w_i f_i(x) \\ \text{s.t. } x \in C \\ \text{where } w_i \geq 0 \text{ for all } i=1, \dots, n \text{ and } \sum w_i = 1 \end{cases}$$

linear aggregation

Complexity: subjacent combinatorial problem



Aggregation Metaheuristics

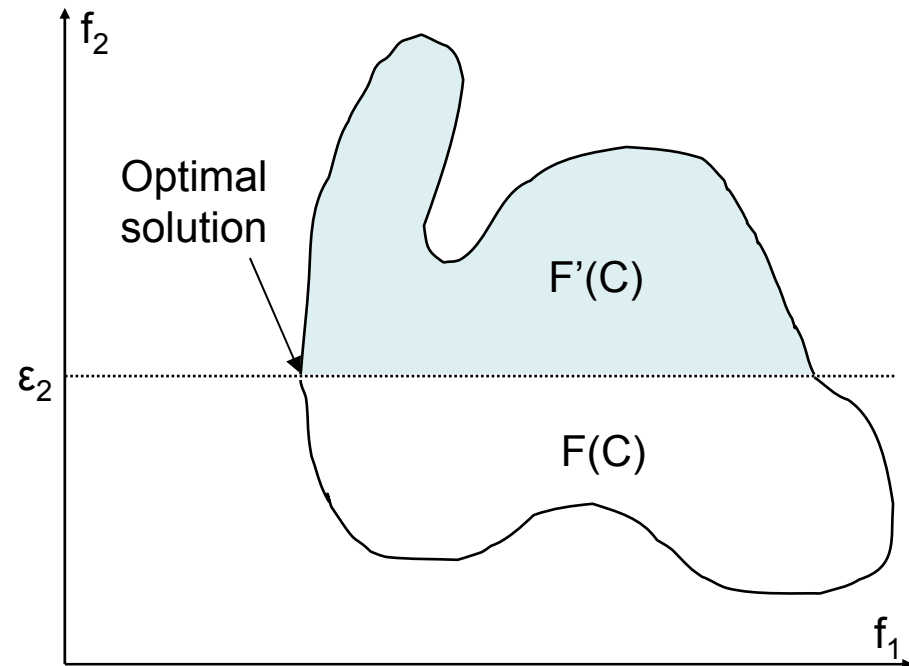
- Genetic algorithms [Hajela et Lin 92]
 - Individual representation: solution + λ
 - Goal: generating various Pareto solutions
- Simulated annealing [Serafini 92]
 - Acceptance probability
- Tabu search [Dahl et al. 95]
 - Weight = priority of the objective
- Hybrid metaheuristics [Talbi 98]
 - Greedy algorithm + Simulated annealing [Tuyttens 98]
 - Genetic algorithm (Local search) [Ishibuchi et Murata 98]
 - Selection with different weights
 - Local search on the produced individual (same weights)

ϵ -constraint Approach

[Haimes et al. 71]

$$(MOP_k(\epsilon)) \begin{cases} \min f_k(x) \\ \text{s.t. } x \in C, \\ f_j(x) \leq \epsilon_j \text{ for all } j=1, \dots, n, j \neq k \end{cases} \quad \epsilon = (\epsilon_1, \dots, \epsilon_n)$$

- Variations of ϵ
- $F(C)$ = objective space of the initial problem
- $F'(C)$ = objective space of the modified problem



ε -constraint Metaheuristics

- Genetic algorithms
 - Individual = ε [Loughlin 98]

$$\varepsilon_i = \varepsilon_{\min} + \frac{(i-1) \cdot (\varepsilon_{\max} - \varepsilon_{\min})}{(k-1)}$$

k = size of the population

- Tabu search [Hertz et al. 94]
 - Set of sub-problems
 - Priority order
 - Threshold (f') = Optimum (f^*)
 - + Allowed deterioration

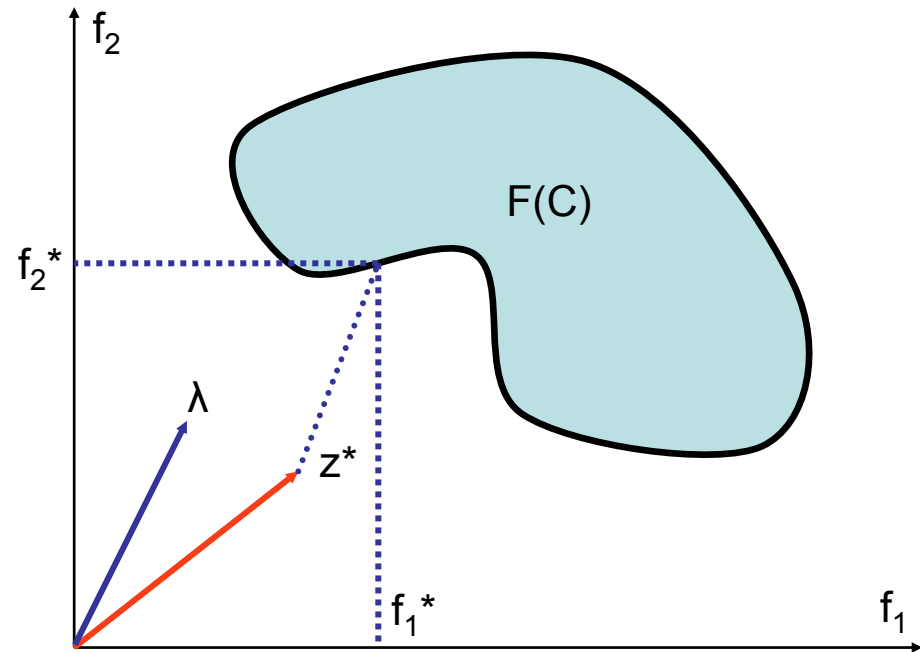
$$(MOP_q) \begin{cases} f_q^* = \min f_q(x) \\ \text{s.t. } x \in C \\ f_r(x) \leq f'_r, r=1, \dots, q-1 \end{cases}$$

- Hybrid metaheuristics [Quagliarella et al. 97]
 - Genetic algorithm + Local search

Goal Programming Approach

$$(MOP_{\lambda, z^*}) \begin{cases} \min (\sum \lambda_i |f_i(x) - z_i^*|^p)^{1/p} & z^*: \text{ideal or reference vector} \\ \text{s.t. } x \in C \end{cases}$$

- Tchebycheff metric
- L_p metric
- $p=2 \rightarrow$ Euclidian metric
- $p=\infty \rightarrow$ Min-Max



Goal Programming Metaheuristics

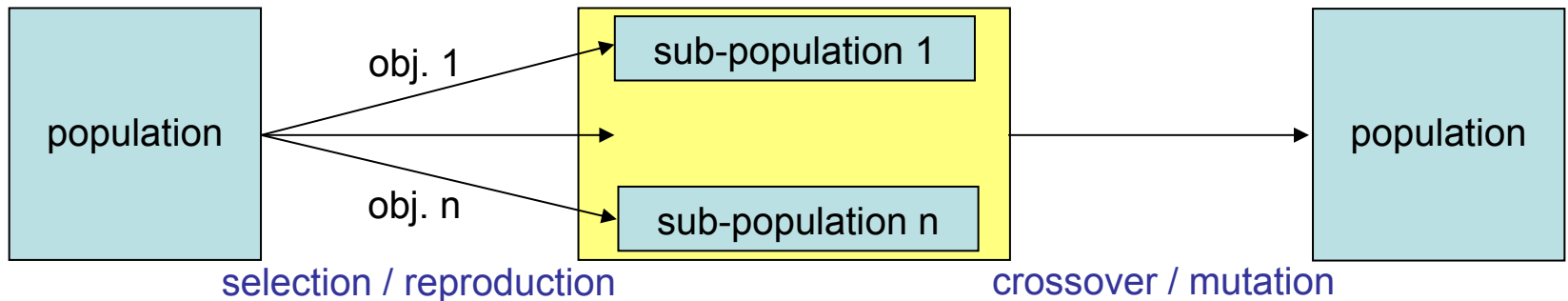
- Genetic algorithms
 - Min-max function + parallel GA with various weights [Coello Coello 98]
 - Fuzzy rules during the evaluation of F [Reardon 98]
- Simulated annealing [Serafini 92]
 - Acceptance probability
 - Iteration = $\lambda_i = \lambda_i + [-0.05, +0.05]$
- Tabu search [Gandibleux 96]
 - Best neighbor = best non-tabu compromise
 - Compromise = L_p norm of Tchebycheff to the ideal vector
 - Saving the M best individuals

Scalar Approaches: Critical Analysis

- Research space \neq initial problem
 - The problem loses its eventual properties
- An execution produces **one solution only**
- **A priori knowledge** (to determine parameters)
- **Sensitive** to the **landscape** of the Pareto front
 - Convexity, discontinuity...
- **Sensitive** to various **parameters**
 - Weights, constraints, goals...
- **Noisy** objectives or **uncertain** data
- **Cost** associated to the algorithms (multiple executions)

Non-Pareto Approaches

- Parallel selection (VEGA) [Schaffer 85]



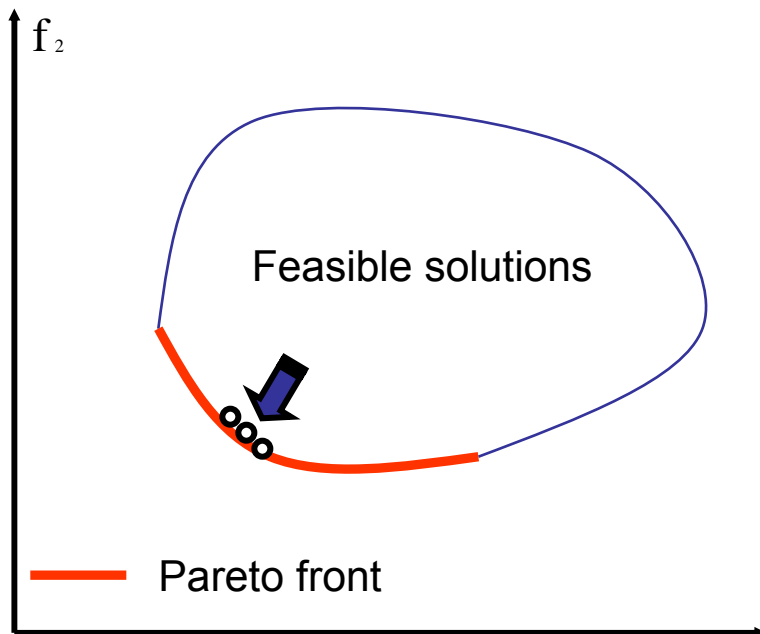
- Lexicographic selection (priority order)
 - Genetic algorithms [Fourman 85]
 - Evolutionary strategies [Kursawe 91]
 - Multi-sexual reproduction [Lis & Eiben 96]
 - One class per objective
 - Reproduction (crossover) over several individuals
- ➔ Tends to ignore compromised solutions

Pareto Approaches

Goals

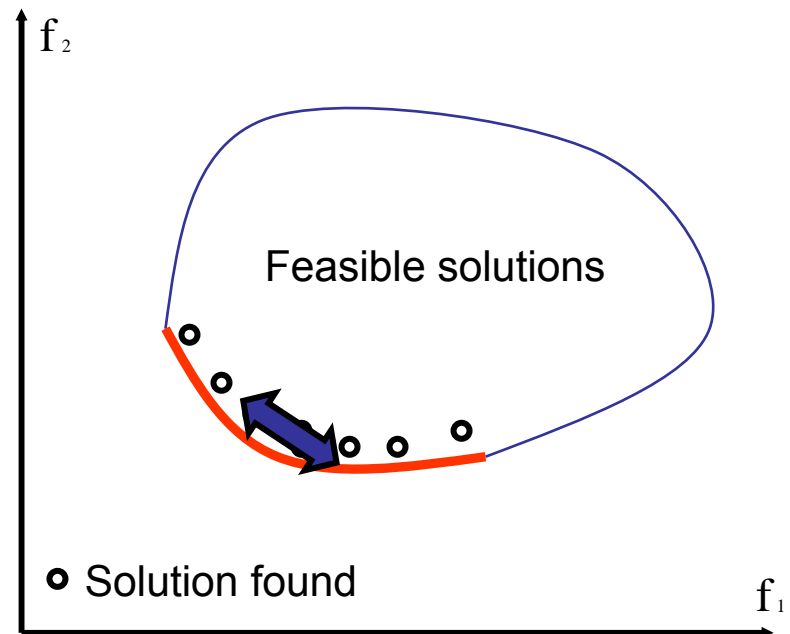
Convergence (fitness)

- Ranking, elitism, ...

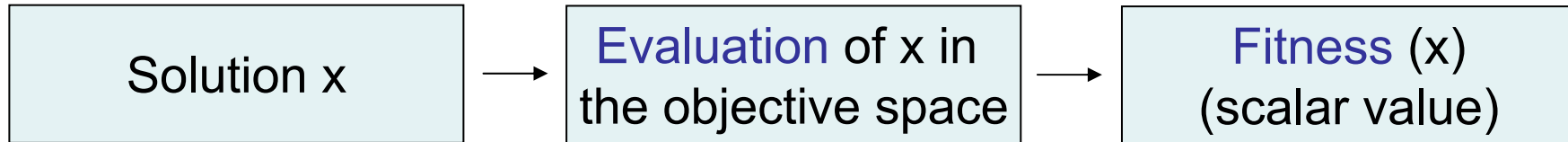


Diversification (diversity)

- Niching, sharing, ...



Fitness assignment

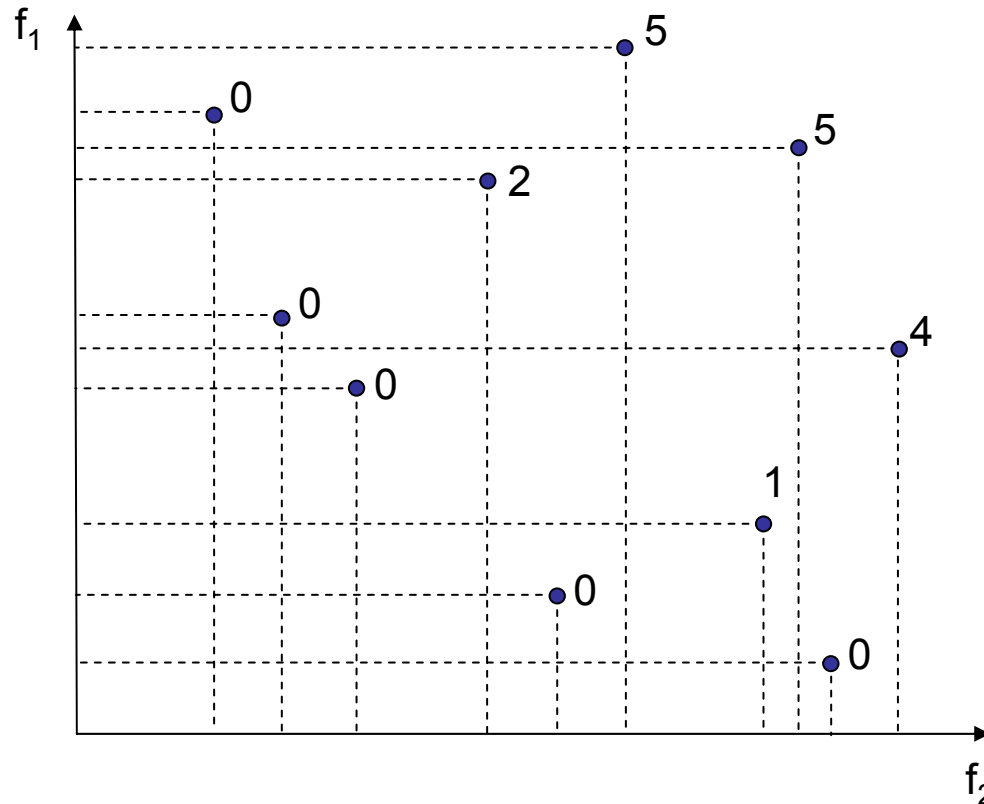


- Fitness (x) reflects the **quality** of x in term of **convergence**
- Pareto-based fitness assignment strategies
 - Dominance **rank** (e.g. used in MOGA)
 - Dominance **depth** (e.g. used in NSGA and NSGA-II)
 - Dominance **count** (e.g. combined with dominance rank in SPEA and SPEA2)
 - **Indicator**-based (e.g. used in IBEA)

Pareto Ranking

[Fonseca & Fleming 93]

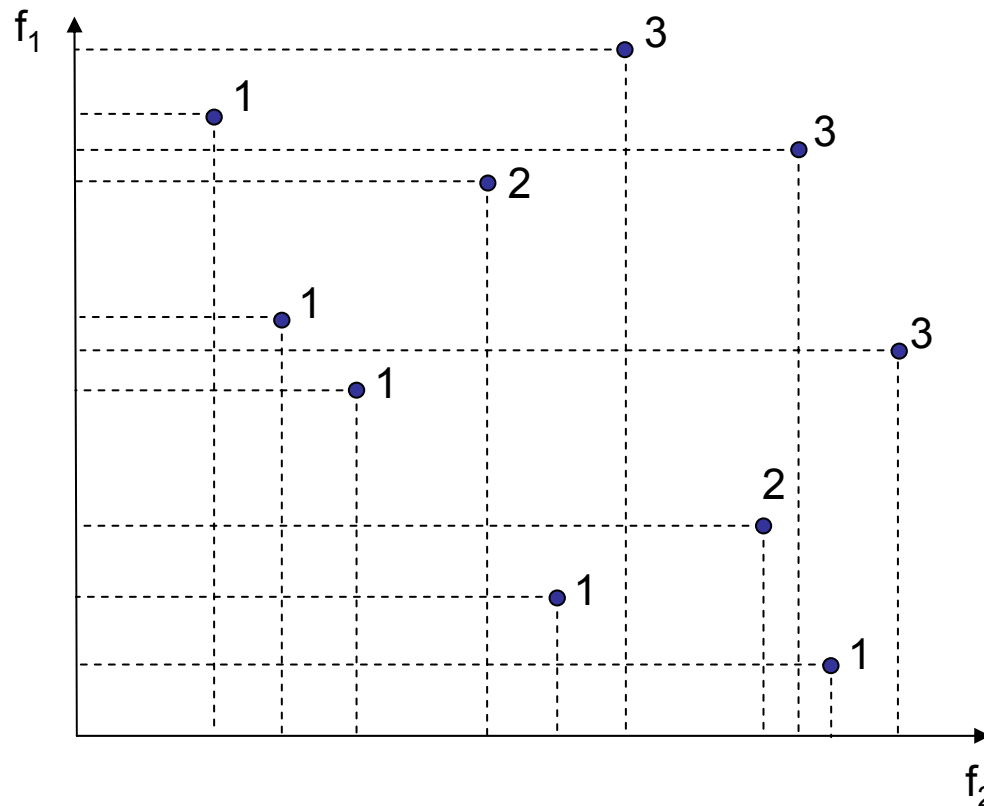
Fitness (x) = number of solutions by which the solution x is dominated



Fast Non-dominated Sorting

[Srinivas & Deb 94]

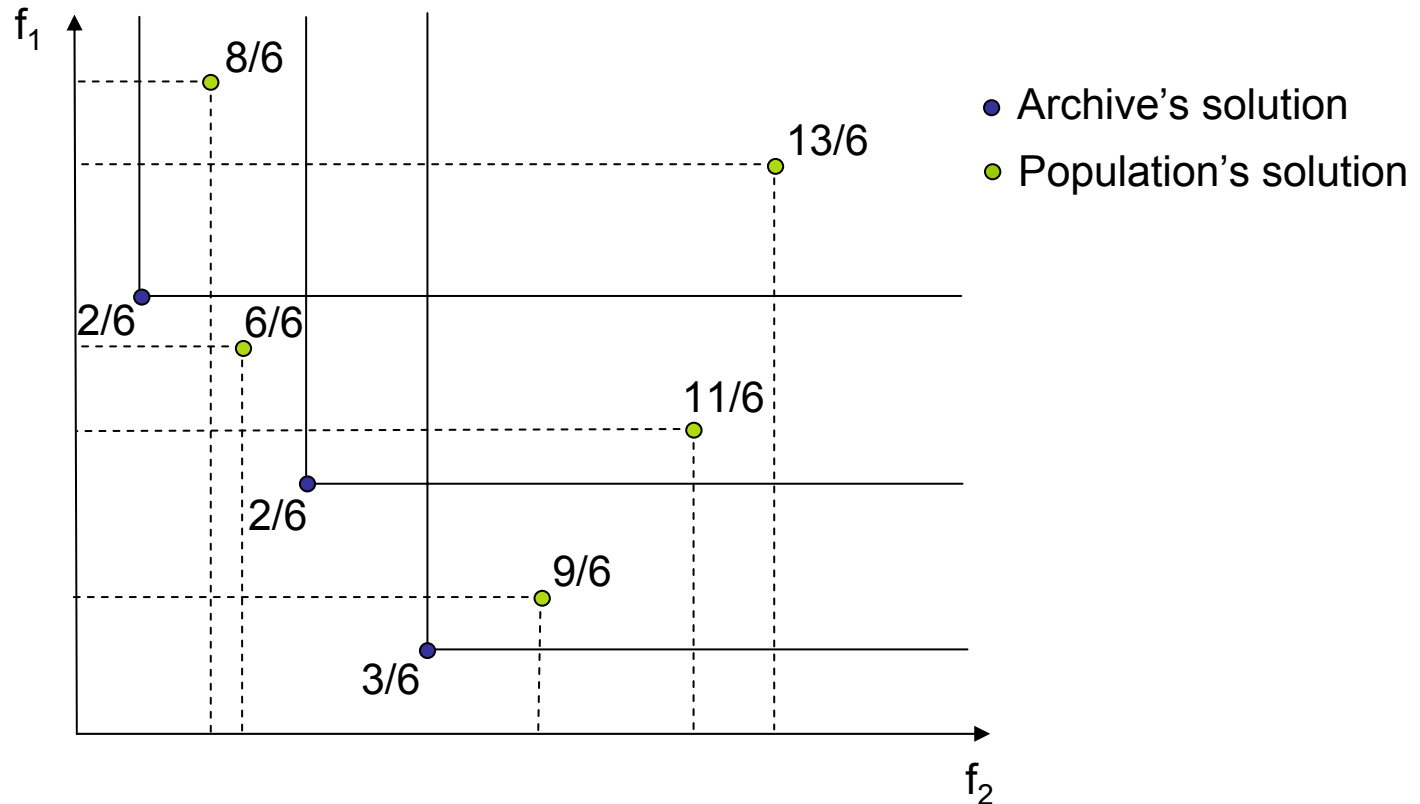
Solutions classified into several fronts



Strength Pareto Fitness Assignment

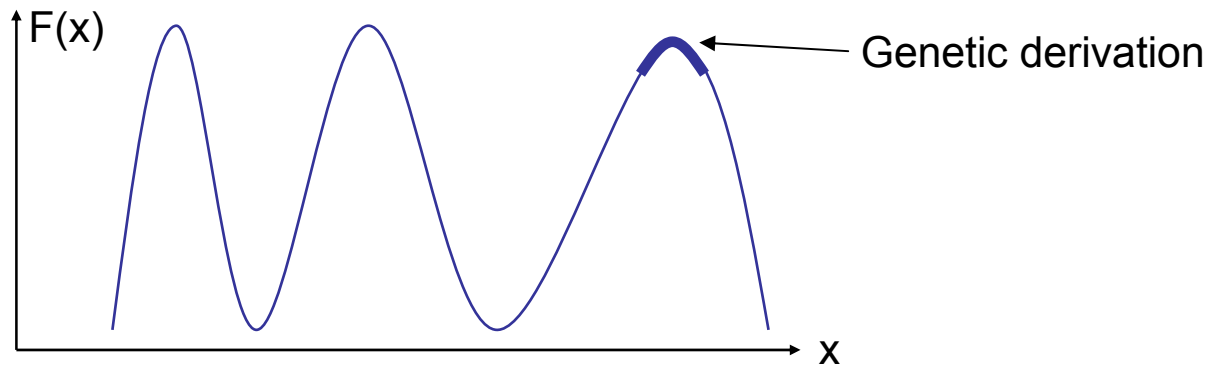
[Zitzler & Thiele 99]

Fitness (x) depends on the strength values of the archive's members that dominate x



Diversity

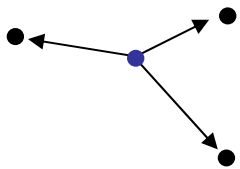
Multi-modal optimization: locating every optima of the problem



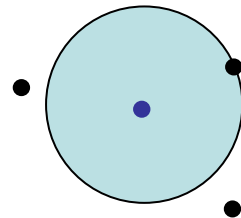
- Independent iterative executions
- Sequential niching
 - Iterative execution with a penalization of the optima already found
- Parallel niching (sharing, crowding)
 - Only one execution

Diversity

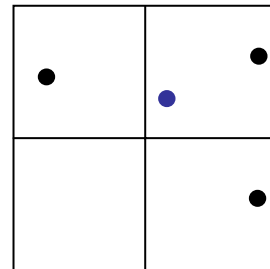
- Kernel methods (**sharing**)
 - Neighborhood of a solution in term of a function taking a distance as argument
- Nearest neighbour techniques
 - Distance of a solution to its k^{th} nearest neighbour
- Histograms
 - Space divided onto neighbourhoods by an hypergrid



kernel



nearest neighbor



histogram

→ decision / objective space

Sharing

→ Reduce the cost of an individual / number of **similar** individuals

$$f(x) = f(x) / \sum \text{sh}(\text{dist}(x,y))$$

$$\text{sh}(\text{dist}(x,y)) = \begin{cases} 1 - (\text{dist}(x,y) / \sigma_{\text{sh}})^\alpha & \text{if } \text{dist}(x,y) < \sigma_{\text{sh}} \\ 0 & \text{else} \end{cases}$$

- Objective space and/or decision space?
- Distance used (dist)?
- Niches size (σ_{sh})?

Pareto-based Metaheuristics

- Mainly evolutionary algorithms
 - MOGA [Fonseca & Fleming 1993]
 - NSGA-II [Deb et al. 2000]
 - SPEA2 [Zitzler et al. 2001]
 - IBEA [Zitzler & Künzli 2004]
- Genetic programming
 - MOGP [Rodriguez-Vasquez et al. 97]
- Ant colony

Pareto-based Metaheuristics

- Scatter search
 - SSPMO [Molina et al. 05]
 - MOSS [Beausoleil 06]
- Local search
 - PLS [Basseur et al. 03]
- Tabu search
 - MOTS [Hansen et al. 97]
 - TAPaS [Jozefowicz et al. 03]
- Simulated annealing

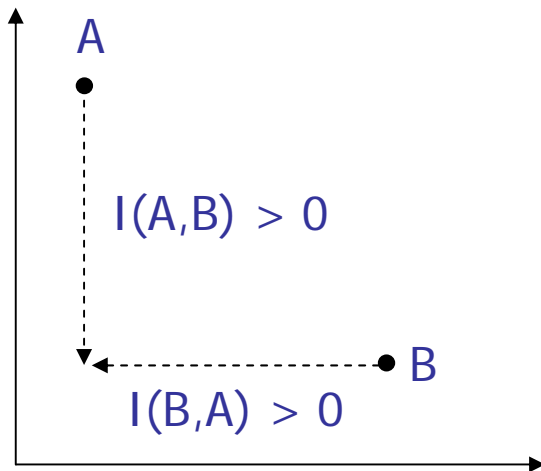
Indicator-Based Approaches

[Zitzler & Künzli 04]

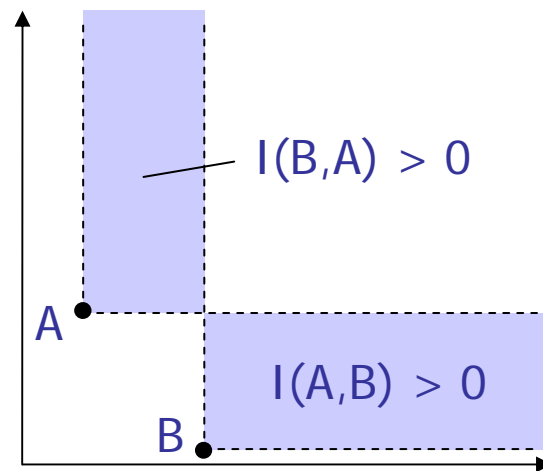
Solutions compared on the basis of a quality indicator I

Fitness (x) = usefulness of x according to the optimization goal (I)

Examples of binary quality indicators:

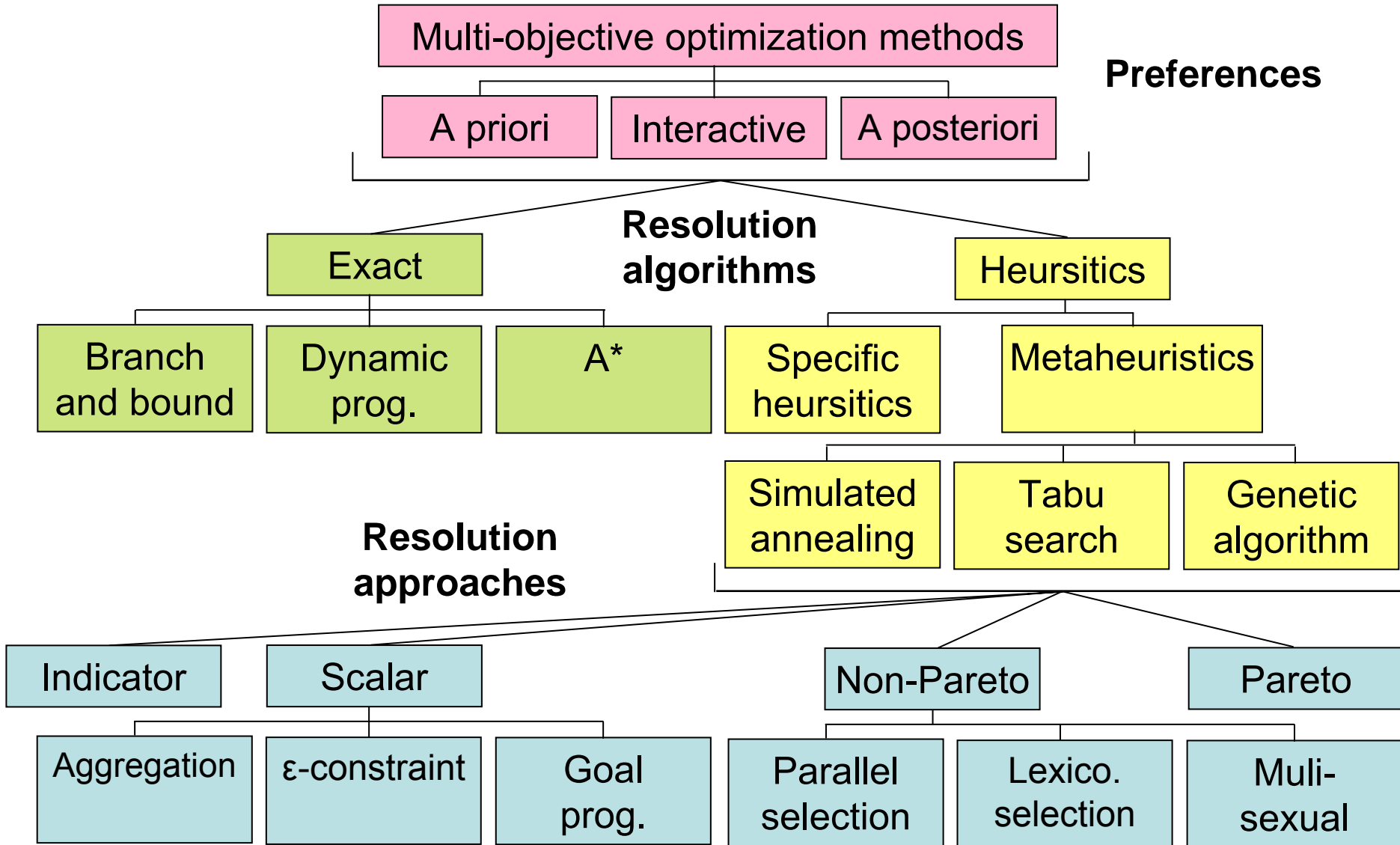


Additive epsilon indicator ($I_{\epsilon+}$)



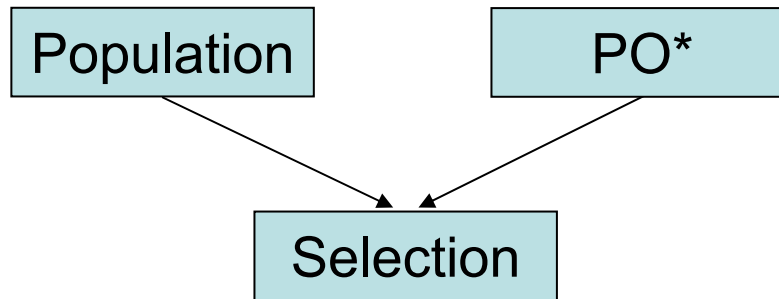
Hypervolume indicator (I_{HD})

Classification



Advanced Techniques

- Elitism (archive PO^*) [Zitzler & Thiele 98]



- Clustering of the PO^* set [Roseman et Gero 85]
- Hybridization
 - Genetic Algorithm + Local Search (Tabu, Archive PO^*)

E-G. Talbi, M. Rahoual, M-H. Mabed, C. Dhaenens: “**A Hybrid Evolutionary Approach for Multicriteria Optimization Problems: Application to the Flow Shop**”, *EMO'01, LNCS 1993*, Springer-Verlag, pp.416-428, Zurich, Switzerland, 2001

Advanced Techniques

- **Parallel models** (genetic algorithms)
 - Quicker research
 - Results improved (cooperation)

N. Jozefowicz, F. Semet, E-G. Talbi: “**Parallel and hybrid models for multi-objective optimization: Application to the vehicle routing problem**”, *PPSN VII, LNCS 2439, Springer-Verlag*, pp. 271-280, Granada, Spain, 2002

- **Adaptive operators**
 - Mutation, crossover, ...

M. Basseur, F. Seynhaeve, E-G. Talbi: “**Adaptive mechanisms for multi-objective evolutionary algorithms**”, *CESA'2003, Lille, France, 2003*

- **Path relinking**

M. Basseur, F. Seynhaeve, E-G. Talbi: “**Path Relinking in Pareto Multi-objective Genetic Algorithms**”, *EMO'05, LNCS 3410, Springer-Verlag*, pp.120-134, Guanajuato, Mexico, 2005

Landscapes and Performance Analysis



PO known

[Teghem et al.]

- Absolute efficiency

- Proportion of Pareto solutions within PO^*

$$AE = \frac{|PO^* \cap PO|}{|PO|}$$

- Distance (PO^* , PO)

- Worst distance

$$WD = \max(d(PO^*, y), y \in PO)$$

- Mean distance

$$MD = \frac{\sum_{y \in PO} d(PO^*, y)}{|PO|}$$

- Uniformity

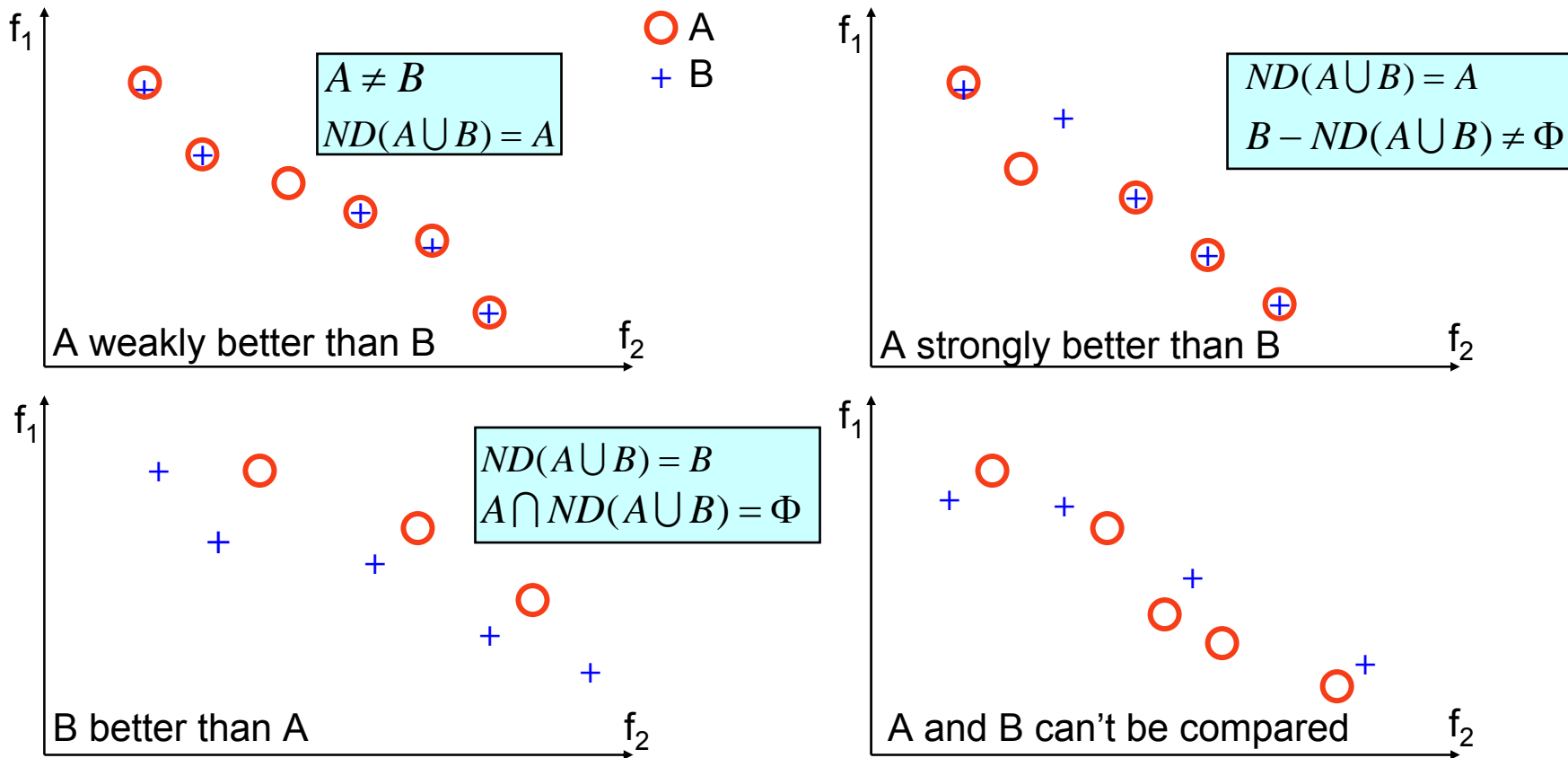
$$DIV = \frac{WD}{MD}$$

$$d(PO^*, y) = \min(d(x, y)), x \in PO^*$$

$$d(x, y) = \sum_{i=1}^n \lambda_i |f_i(x) - f_i(y)|$$

PO unknown

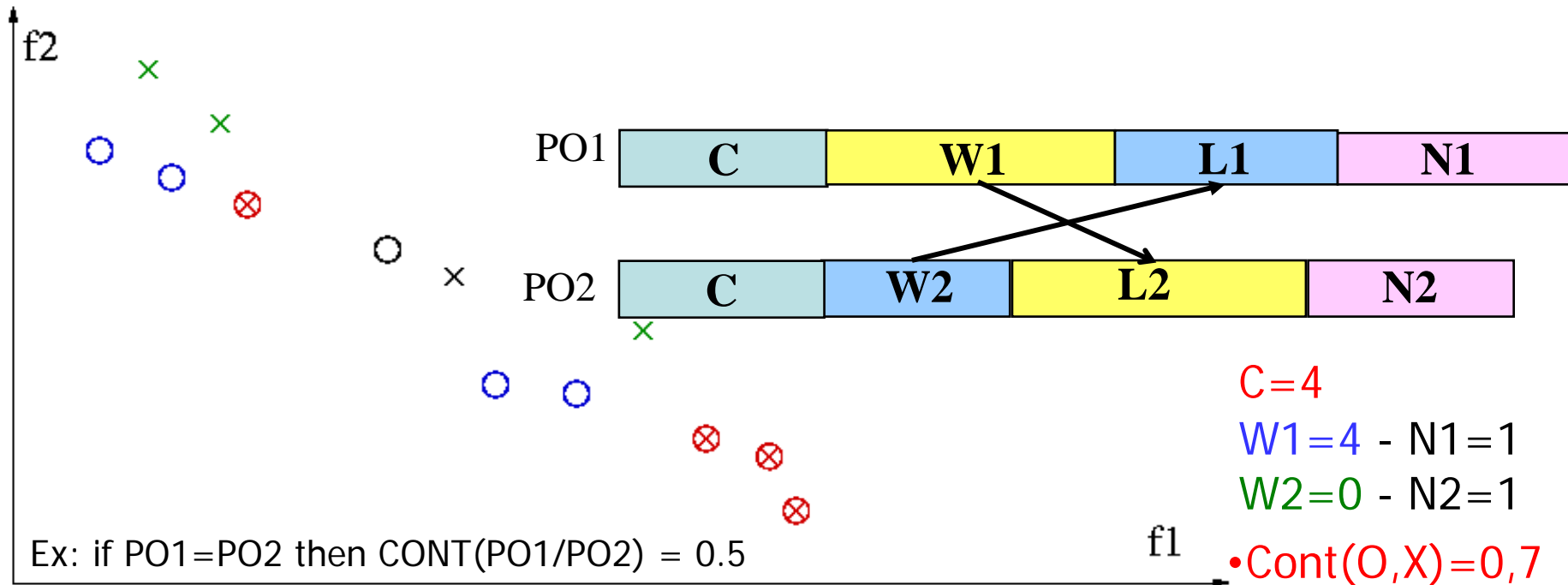
- **Relative efficiency:** number of solutions from A dominated by B



PO unknown

Contribution: Evaluating the quality of the solutions from a set towards another one

$$Cont(PO_1/PO_2) = \frac{\|C\|/2 + \|W_1\| + \|N_1\|}{\|C\| + \|W_1\| + \|N_1\| + \|W_2\| + \|N_2\|}$$



C=4
W1=4 - N1=1
W2=0 - N2=1

• **Cont(O,X)=0,7**
 • **Cont(X,O)=0,3**

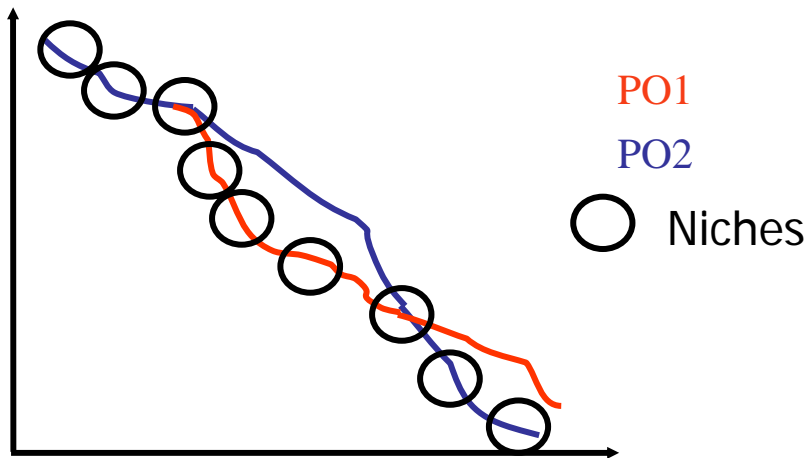
PO unknown

- **Entropy**: builds a niche around every solution of

$ND(PO_1 \cup PO_2) = PO^*$

- $E(PO_1, PO_2)$: diversity of the solutions of PO_1 in comparison of those in the niches of PO^*

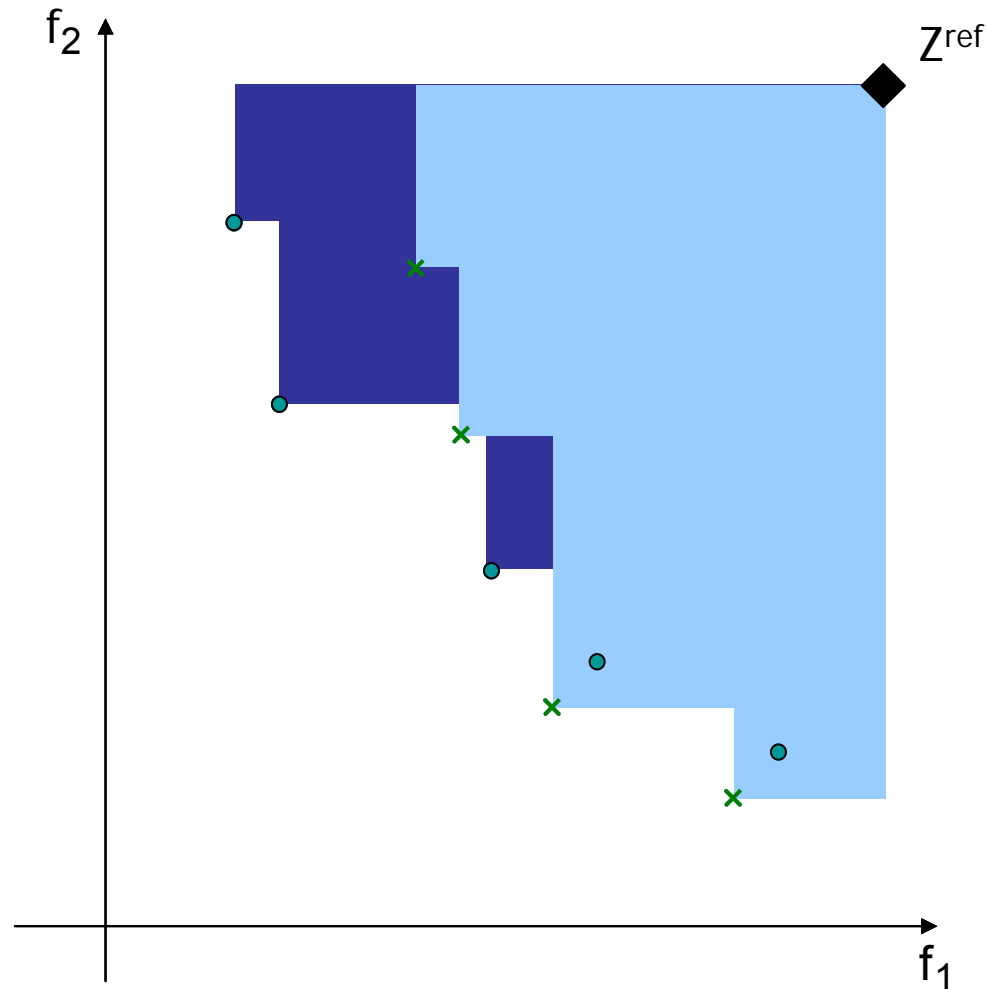
$$E(PO_1, PO_2) = \frac{-1}{\ln(\gamma)} \sum_{i=1}^{|PO^*|} \left(\frac{1}{Ni} \frac{n_i}{\|PO_1\|} \ln \frac{n_i}{\|PO_1\|} \right)$$



PO unknown

- S-metric / Hypervolume
[Zitzler 99]

Size of the objective space enclosed by PO^* and a reference point Z^{ref}



Landscapes

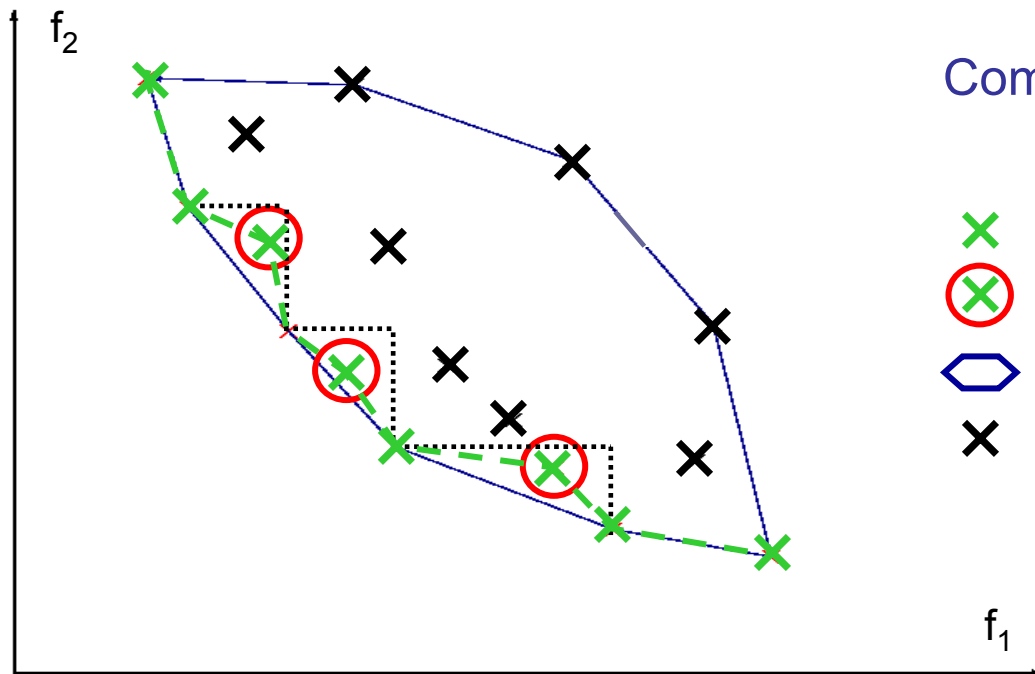
How to describe a Pareto front?

- Convexity of PO (supported or not)
- Multi-modality
- Deception (deceptive attractors)
- Isolated optimum (Flat space)
- Discontinuity
- Uniform distribution

Landscapes

Aggregation: supported solutions only

Convexity: Proportion of Pareto solutions belonging to the convex hull



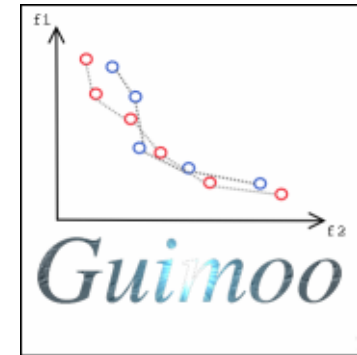
Complexity: $O(n \cdot \log(n))$

- ✕ Non-dominated solutions
- ✕ Unsupported solutions
- ◊ Convex hull
- ✕ Dominated solutions

GUIMOO

a Graphical User Interface for
Multi-objective Optimization

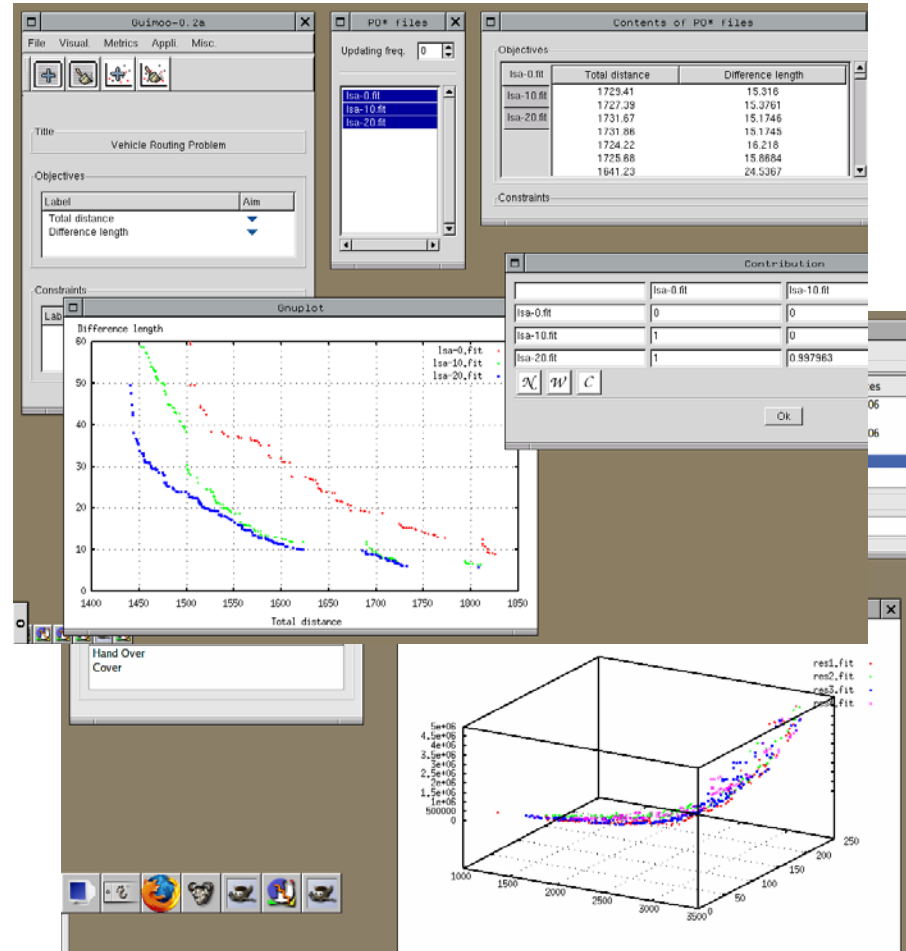
<http://guimoo.gforge.inria.fr>



Paradiseo

GUIMOO

- Graphical tool dedicated to the analysis of multi-objective optimization results
- Free software (open source)
- Platform-independent
- Different input/output formats
- Easily customized for specific applications
 - Telecom, genomics, engineering design, ...
- On-line user manual



GUIMOO

- On-line and off-line visualization (2D or 3D) of Pareto fronts
 - (dis)continuity, (dis)convexity, multi-modality, ...
 - Distribution of the solutions among the objective space
- Metrics for quantitative and qualitative performance evaluation
 - Contribution
 - Entropy
 - Generational distance
 - Spacing
 - Coverage of two sets, coverage difference
 - S-metric, D-metric
 - R-metrics

Conclusion



Conclusion

- Important research area orientation for scientists and engineers (real problems, open questions)
- Population-based metaheuristics naturally efficient
- Hybrid approaches
- Performance analysis
- Interactive approach
- **ParadisEO-MOEO**: a framework for multi-objective optimization
 - <http://paradiseo.gforge.inria.fr>
- **GUIMMO**: a graphical user interface for multi-objective optimization
 - <http://guimoo.gforge.inria.fr>

Conclusion

- Academic problems (benchmarks : www.lifl.fr/OPAC/)
 - TSP / VRP
 - Scheduling problems
- Real-world applications
 - Bioinformatics
 - Protein docking
 - Protein folding
 - Knowledge discovery in biological data
 - Telecom
 - Cellular networks design
 - Polymer composites in the microwave band