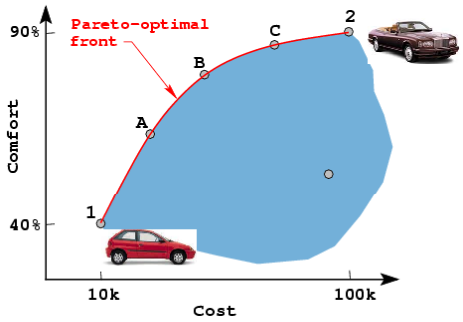
Lecture 6:
Multi-Objective Optimization

Overview:

- Introduction to multi-objective optimization
- Main classical methods
- Philosophy of evolutionary methods
- Early non-elitist methods
- Efficient elitist methods
- Constraint handling

Multi-Objective Optimization

- Really need no introduction

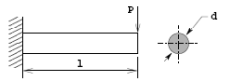


Formally, as a Mathematical Programming Problem

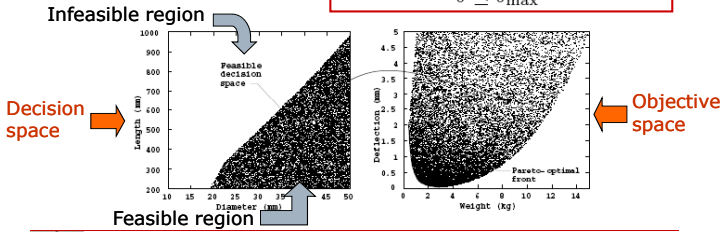
- Multiple objectives, constraints, and variables

$$\begin{aligned} \text{Min/Max} \quad & (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x})) \\ \text{Subject to} \quad & g_j(\mathbf{x}) \geq 0 \\ & h_k(\mathbf{x}) = 0 \\ & \mathbf{x}^{(L)} \leq \mathbf{x} \leq \mathbf{x}^{(U)} \end{aligned}$$

A Simple Example from Practice

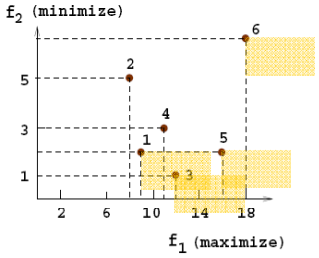


$$\begin{aligned} \text{Minimize} \quad & f_1(d,l) = \rho\frac{\pi d^2}{4}l \\ \text{Minimize} \quad & f_2(d,l) = \delta = \frac{64Pl^3}{3E\pi d^4} \\ \text{subject to} \quad & \sigma_{\max} \leq S_y \\ & \delta \leq \delta_{\max} \end{aligned}$$



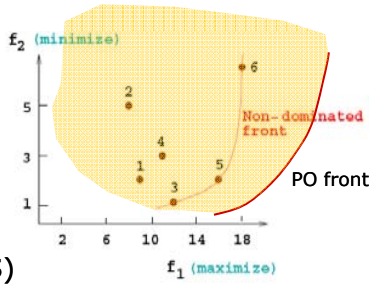
Which Solutions are Optimal?

- Relates to the concept of domination
- $x^{(1)}$ dominates $x^{(2)}$, if
  - $x^{(1)}$ is no worse than $x^{(2)}$ in all objectives
  - $x^{(1)}$ is strictly better than $x^{(2)}$ in at least one objective
- Examples:
  - 3 dominates 2
  - 3 does not dominate 5



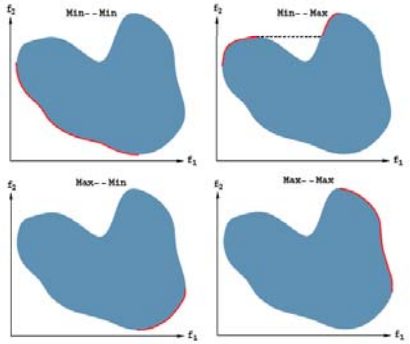
Pareto-Optimal Solutions

- P'=Non-dominated(P)
  - Solutions which are not dominated by any member of the set P
- O(N log N) algorithms exist
- Pareto-Optimal set = Non-dominated(S)
- A number of solutions are optimal



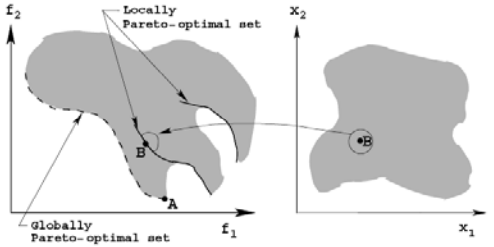
Pareto-Optimal Fronts

- Depends on the type of objectives
- Always on the boundary of feasible region



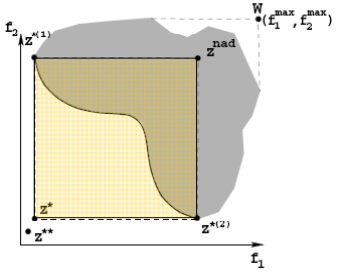
Local Versus Global Pareto-Optimal Fronts

Local Pareto-optimal Front: Domination check is restricted within a neighborhood (in decision space) of P



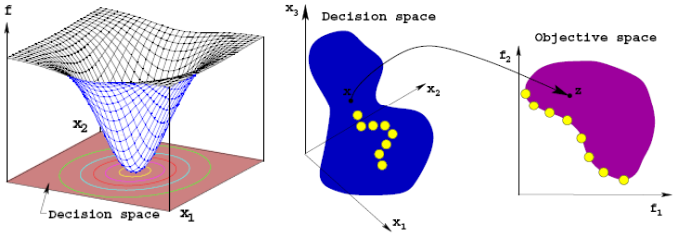
Some Terminologies

- Ideal point (z*)
  - nonexistent, lower bound on Pareto-optimal set
- Utopian point (z**)
  - nonexistent
- Nadir point (z^{nad})
  - Upper bound on Pareto-optimal set
- Normalization: $f_i^{norm} = \frac{f_i - z_i^*}{z_i^{nad} - z_i^*}$

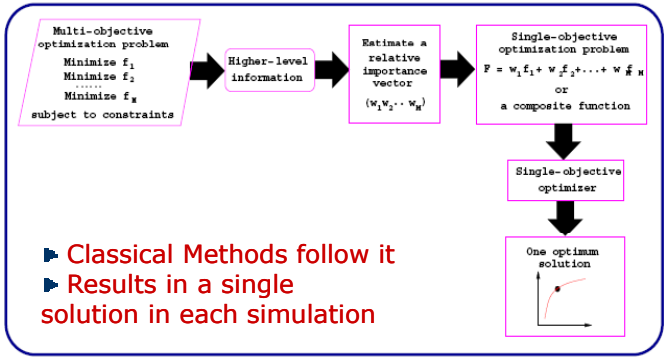


Differences with Single-Objective Optimization

- One optimum versus multiple optima
- Requires search and decision-making
- Two spaces of interest, instead of one



Preference-Based Methods



- Classical Methods follow it
- Results in a single solution in each simulation

Classical Approaches

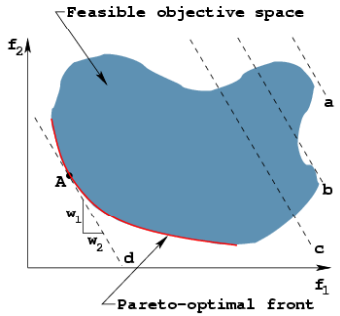
- No Preference methods (heuristic-based)
- Posteriori methods (generating solutions) discussed later
- A-priori methods (one preferred solution)
- Interactive methods (involving a decision-maker)

Classical Approach: Weighted Sum Method

- Construct a weighted sum of objectives and optimize

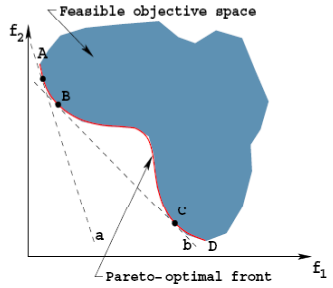
$$F(x) = \sum_{i=1}^{M} w_i f_i(x)$$

- User supplies weight vector *w*



Difficulties with Weighted-Sum Method

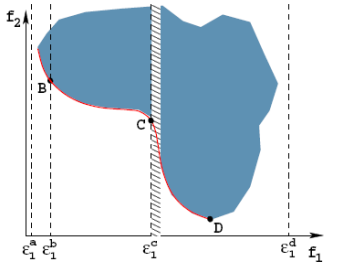
- Need to know **w**
- Non-uniformity in Pareto-optimal solutions
- Inability to find some Pareto-optimal solutions (those in non-convex region)
- However, a solution of this approach is Pareto-optimal



ε-Constraint Method

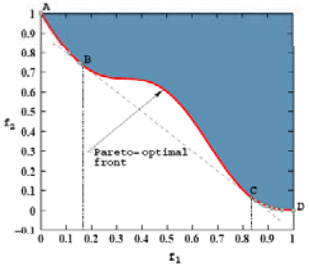
- Constrain all but one objective
- Need to know relevant **ε** vectors
- Non-uniformity in Pareto-optimal solutions
- However, any Pareto-optimal solutions can be found with this approach

$$\text{Minimize} \quad f_\mu(\mathbf{x}),$$
$$\text{subject to} \quad f_m(\mathbf{x}) \le \epsilon_m, \; m \neq \mu;$$



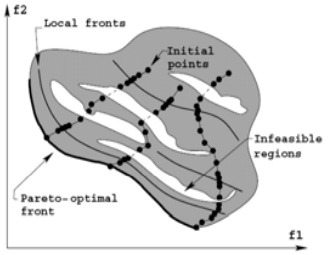
Difficulties with Most Classical Approaches

- Need to run a single-objective optimizer many times
- Expect a lot of problem knowledge
- Even then, good distribution is not guaranteed
- Multi-objective optimization as an application of single-objective optimization



Classical Generating Methods

- One-at-a-time and repeat
- Population approaches
  - Timmel's method
  - Schaffler's method
- Absence of parallel search is a drawback
- EMO finds multiple solutions with an implicit parallel search



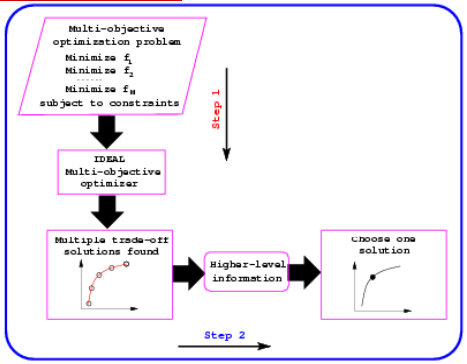
Ideal Multi-Objective Optimization

Step 1 :

Find a set of Pareto-optimal solutions

Step 2 :

Choose one from the set



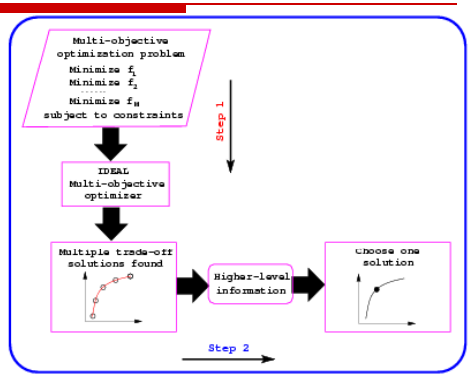
Ideal Multi-Objective Optimization

Step 1 :

Find a set of Pareto-optimal solutions

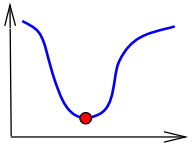
Step 2 :

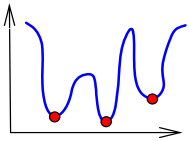
Choose one from the set



A More Holistic Approach for Optimization

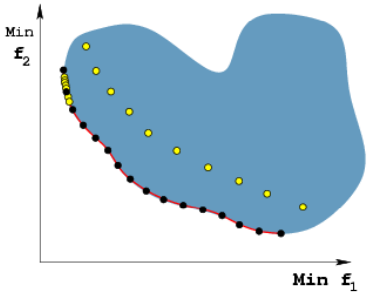
- Decision-making becomes easier and less subjective
- Single-objective optimization is a degenerate case of multi-objective optimization
  - Step 1 finds a single solution
  - No need for Step 2
- Multi-modal optimization possible





Two Goals in Ideal Multi-Objective Optimization

- Converge to the Pareto-optimal front
- Maintain as diverse a distribution as possible

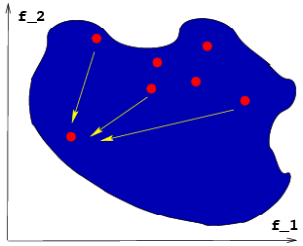


Evolutionary Multi-Objective Optimization (EMO)

- Principle:
  - Find multiple Pareto-optimal solutions simultaneously
- Two main reasons:
  - Help in choosing a particular solution
  - Unveil salient optimality properties of solutions
  - Assist in other problem solving

EMO Methodologies

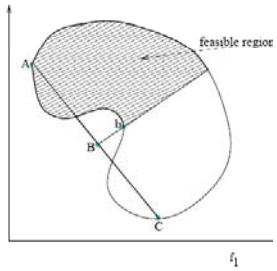
- Population approach suits well to find multiple solutions simultaneously
- Implicit parallelism through EA operators including recombination helps provide a parallel and faster search



EMO Versus Normal Constraint (NC) Method (Messac and Mattson, 2006)

- A modification to ε-constraint method
- Find anchor points N_k as minimum of k-th obj (A,C)
- Create a set of indicator points (Das and Dennis, 1998)

$$\min_{(\mathbf{x})} \quad \overline{f}_M(\mathbf{x}),$$
$$\text{subject to} \quad \overline{N}_k^T(\overline{F} - Z_\beta) \leq 0, \qquad \text{for } 1 \leq k \leq M\text{-}1$$
$$\mathbf{x} \in S$$

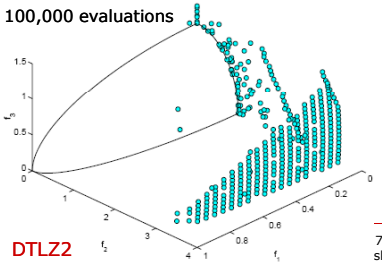


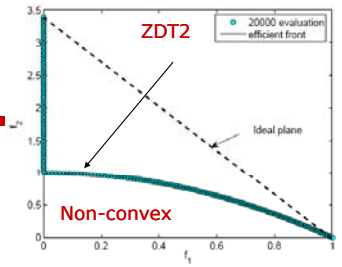
- Employ SQP to solve each subproblem

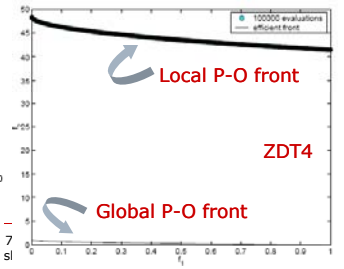
NCM Results

Difficulties:

- Weak P-O points
- Local P-O fronts
- More objectives

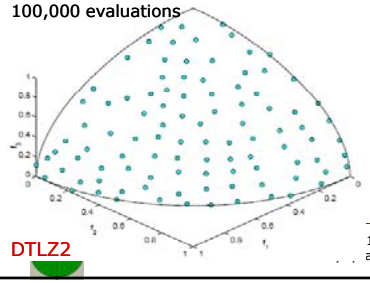


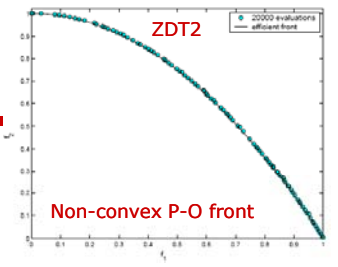


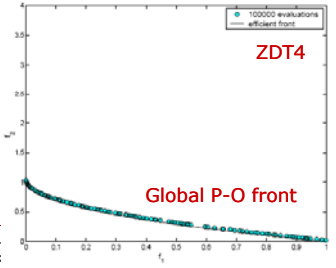


NSGA-II Results

- Population size=100
- Standard parameter setting
- Seems to work well

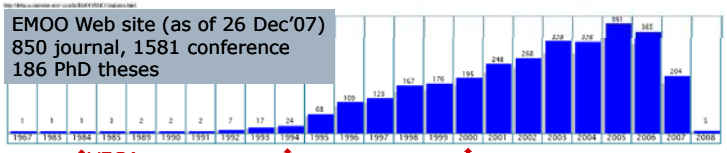






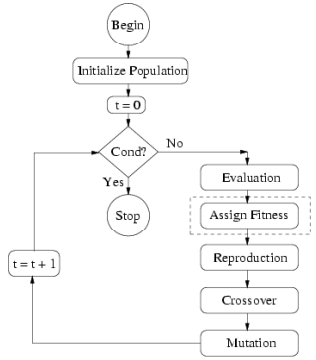
Evolution of EMO

- Early penalty-based approaches
- VEGA (1984)
- Goldberg's (1989) suggestion
- MOGA, NSGA, NPGA (1993-95) used Goldberg's suggestion
- Elitist EMO (SPEA, NSGA-II, PAES, MOMGA etc.) (1998 -- Present)



What to Change in a Simple GA?

- Modify the fitness computation
- Emphasize non-dominated solutions for convergence
- Emphasize less-crowded solutions for diversity



Identifying Non-Dominated Set

Step 1 Set $i = 1$ and create an empty set P'.

Step 2 For a solution $j \in P$ (but $j \neq i$), check if solution j dominates solution i. If yes, go to Step 4.

Step 3 If more solutions are left in P, increment j by one and go to Step 2; otherwise, set $P' = P' \cup \{i\}$.

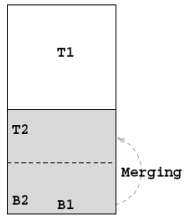
Step 4 Increment i by one. If $i \leq N$, go to Step 2; otherwise stop and declare P' as the non-dominated set.

- O(MN^2) computational complexity

Kung et al.'s (1975) Approach

Step 1 Sort the population in descending order of importance of f_1

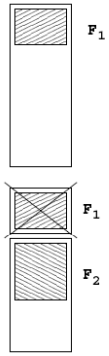
Step 2, Front(P) If $|P| = 1$, return P as the output of **Front**(P). Otherwise, $T = \textbf{Front}(P^{(1)} -\!- P^{(\lfloor |P|/2 \rfloor)})$ and $B = \textbf{Front}(P^{(\lfloor |P|/2+1 \rfloor)} -\!- P^{(|P|)})$. If the i-th solution of B is not dominated by any solution of T, create a merged set $M = T \cup \{i\}$. Return M as the output of **Front**(P).



$O\left(N(\log N)^{M-2}\right)$ for $M \geq 4$ and $O(N \log N)$ for $M = 2$ and 3

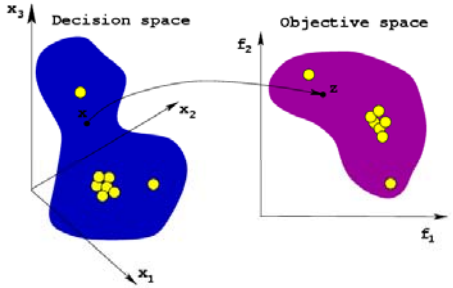
Non-Dominated Sorting: A Naive Approach

- Identify the best non-dominated set
- Discard them from population
- Identify the next-best non-dominated set
- Continue till all solutions are classified



Which are Less-Crowded Solutions?

- Crowding can be in decision variable space or in objective space



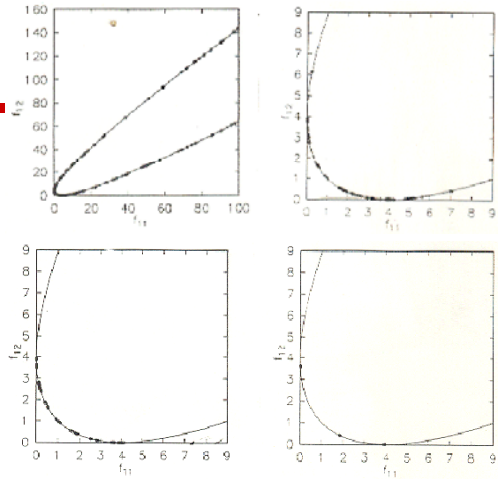
Non-Elitist EMO Procedures

- Vector evaluated GA (VEGA) (Schaffer, 1984)
- Vector optimized EA (VOES) (Kursawe, 1990)
- Weight based GA (WBGA) (Hajela and Lin, 1993)
- Multiple objective GA (MOGA) (Fonseca and Fleming, 1993)
- Non-dominated sorting GA (NSGA) (Srinivas and Deb, 1994)
- Niched Pareto GA (NPGA) (Horn et al., 1994)
- Predator-prey ES (Laumanns et al., 1998)
- Other methods: Distributed sharing GA, neighborhood constrained GA, Nash GA etc.

Schaffer's (1984) Vector-Evaluated GA (VEGA)

- Divide population into M equal blocks
- Each block is reproduced with one objective function
- Complete population participates in crossover and mutation
- Bias towards to individual best objective solutions
- A non-dominated selection: Non-dominated solutions are assigned more copies
- Mate selection: Two distant (in parameter space) solutions are mated
- Both necessary aspects missing in one algorithm

VEGA Results

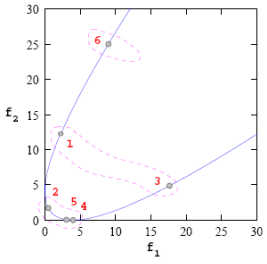


Srinivas and Deb's (1995) Non-dominated Sorting GA (NSGA)

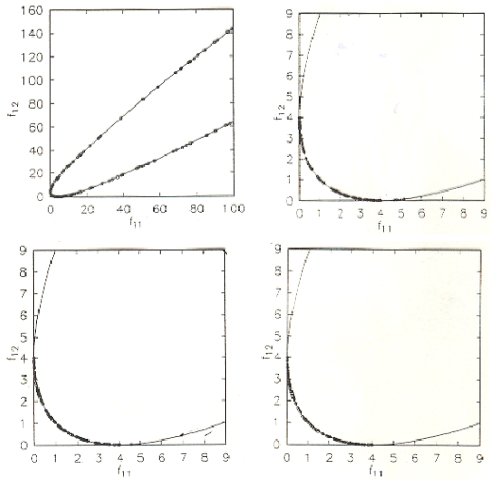
- Niching in parameter space
- Non-dominated solutions are emphasized
- Diversity among them is maintained

Example:
$f_1(x)=x^2$
$f_2(x)=(x-2)^2$

x	f_1	f_2	Front	Fitness before	Fitness after
−1.50	2.25	12.25	2	3.00	3.00
0.70	0.49	1.69	1	6.00	6.00
4.20	17.64	4.84	2	3.00	3.00
2.00	4.00	0.00	1	6.00	3.43
1.75	3.06	0.06	1	6.00	3.43
−3.00	9.00	25.00	3	2.00	2.00



NSGA Results



Fonseca and Fleming's (1993) Multi-Objective GA (MOGA)

- Count the number of dominated solutions (say n)
- Fitness: $F=n+1$
- A fitness ranking adjustment
- Niching in fitness space
- Rest all are similar to NSGA

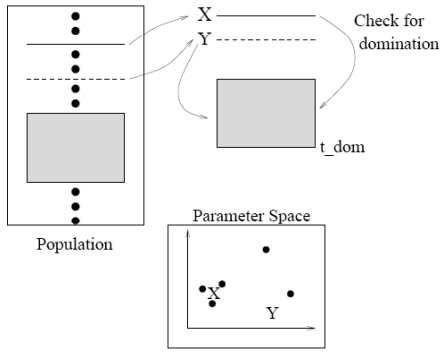
Example:

	F	Asgn.	Fit.
1	2	3	2.5
2	1	6	5.0
3	2	2	2.5
4	1	5	5.0
5	1	4	5.0
6	3	1	1.0

Horn et al.'s (1995) Niched Pareto GA (NPGA)

- Tournament within a small subpopulation (of size t_{dom})
- If one dominated and other non-dominated, select second
- If both non-dominated or both dominated, choose the one with smaller niche count in the subpopulation
- Algorithm depends on t_{dom}
- Nevertheless, it has both necessary components

NPGA (cont.)



Shortcomings of Non-Elitist EMO Procedures

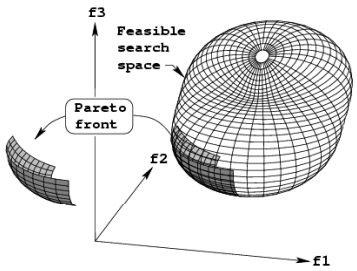
- Elite-preservation is missing
- Elite-preservation is important for proper convergence in single-objective EAs
- Same is true in EMO procedures
- Three tasks
  - Elite preservation
  - Progress towards the Pareto-optimal front
  - Maintain diversity among solutions

Elitist EMOs

- Distance-based Pareto GA (DPGA) (Osyczka and Kundu, 1995)
- Thermodynamical GA (TDGA) (Kita et al., 1996)
- Strength Pareto EA (SPEA) (Zitzler and Thiele, 1998)
- Non-dominated sorting GA-II (NSGA-II) (Deb et al., 1999)
- Pareto-archived ES (PAES) (Knowles and Corne, 1999)
- Multi-objective Messy GA (MOMGA) (Veldhuizen and Lamont, 1999)
- Other methods: Pareto-converging GA, multi-objective micro-GA, elitist MOGA with co-evolutionary sharing

Elitist Non-dominated Sorting Genetic Algorithm (NSGA-II)

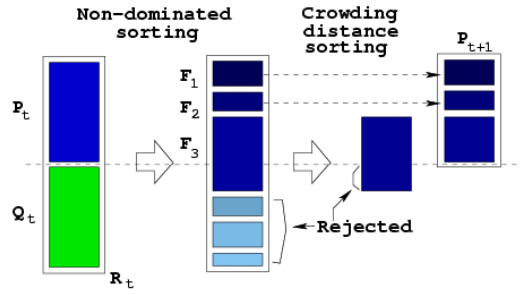
- Deb et al. (IEEE TEC, 2002)
- NSGA-II can extract Pareto-optimal frontier
- Also find a well-distributed set of solutions
- iSIGHT and modeFrontier adopted NSGA-II



Fast-Breaking Paper in Engineering by ISI Web of Science (Feb'04)

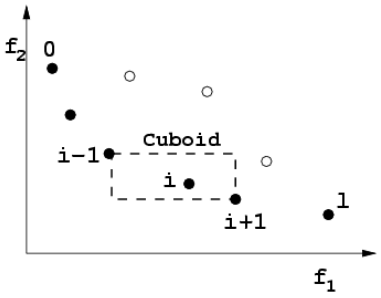
NSGA-II Procedure

Elites are preserved
Non-dominated solutions are emphasized



NSGA-II (cont.)

Diversity is maintained

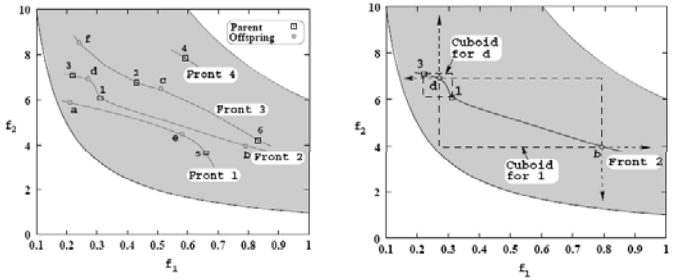


Overall Complexity
$O(N \log^{M-1} N)$

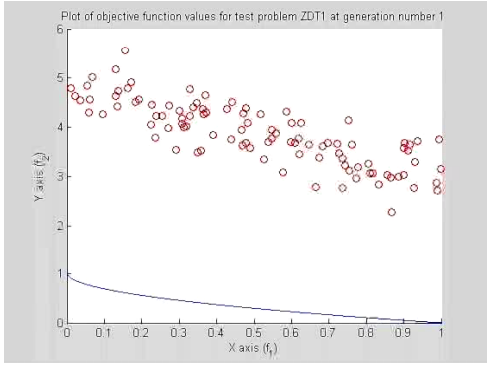
Improve diversity by
- k-mean clustering
- Euclidean distance measure
- Other techniques

One Iteration of NSGA-II

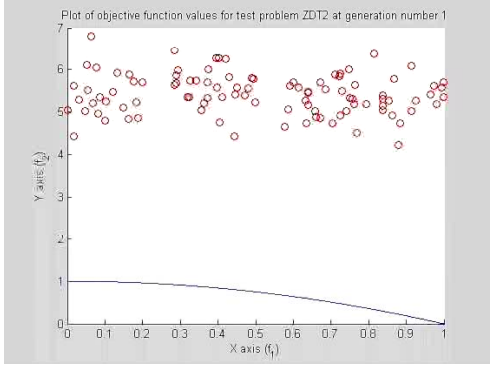
- Six parents and six offspring
- Parents after one iteration: (a,3,1,e,5,b)



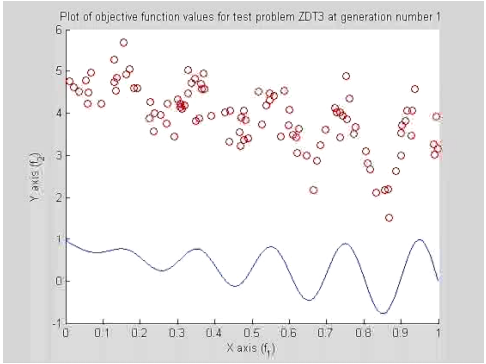
Simulation on ZDT1



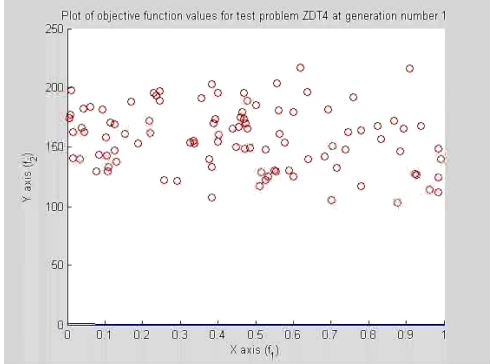
Simulation on ZDT2



Simulation on ZDT3



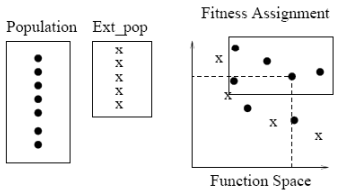
Simulation on ZDT4

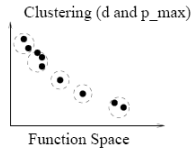


Zitzler and Thiele's (1999) Strength Pareto EA (SPEA)

- Stores non-dominated solutions externally
- Pareto-dominance to assign fitness
- External members: Assign number of dominated solutions in population (smaller -> better)
- Population members: Assign sum of fitness of external dominating members (smaller->better)
- Tournament selection and recombination applied to combined population
- A clustering technique to maintain diversity in updated external population, when size increases a limit

Fitness Assignment and Clustering in SPEA

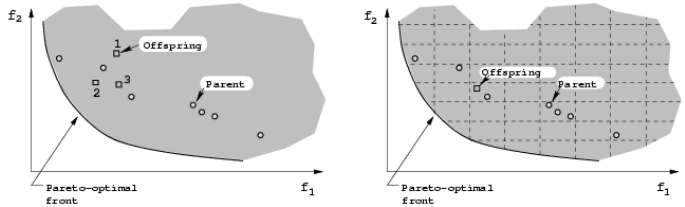




- SPEA:
  - O(MN^3) operation
- SPEA2
  - Improved clustering for not losing boundary points
  - $k=\sqrt{N+\bar{N}}$
  - Fixed archive size
  - Mating within archive members
  - Dominating points use different fitness

Pareto Archived ES (PAES)

- A point-by-point approach
- Parent p_t and child c_t are compared with an external archive A_t
- Child can enter the archive and can become a parent



Constrained Handling

- Penalty function approach

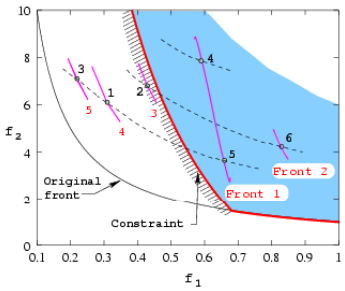
$$F_m = f_m + R_m \Omega\left(\vec{g}\right)$$

- Explicit procedures to handle infeasible solutions
  - Jimenez's approach
  - Ray-Tang-Seow's approach
- Modified definition of domination
  - Fonseca and Fleming's approach
  - Deb et al.'s approach

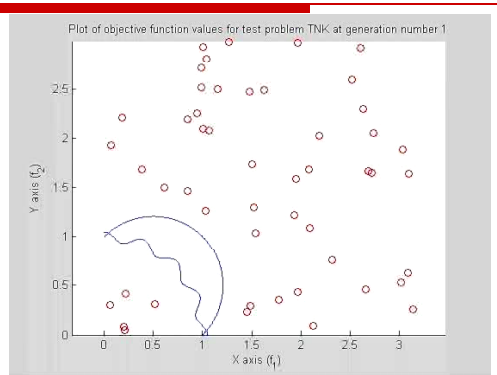
Constraint-Domination Principle

A solution *i* constraint-dominates a solution *j*, if any is true:

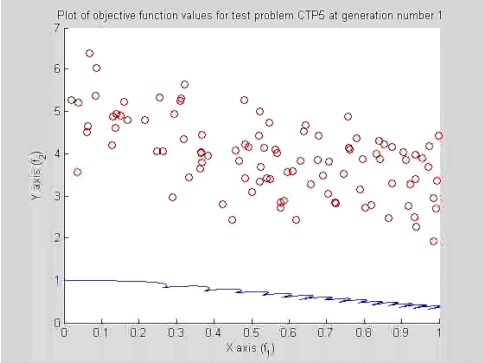
1. *i* is feasible and *j* is not
2. *i* and *j* are both infeasible, but *i* has a smaller overall constraint violation
3. *i* and *j* are feasible and *i* dominates *j*



Simulation on TNK

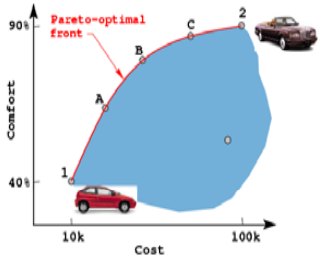


Simulation on CTP5



Application to Aid a Better Decision-Making

- Existence of multiple trade-off solutions
  - Provide trade-off information
  - A better idea of the nature of Pareto-optimal front
  - An idea of range of solutions
- Weighted scheme with 70-30
  - Always wonder what if
  - 69-31 or 71-29?

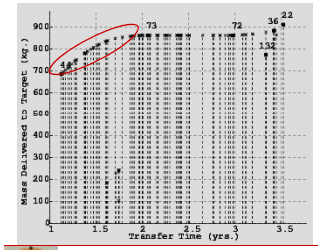


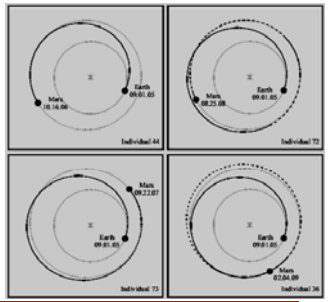
An Application Case Study

- Spacecraft trajectory optimization (Coverstone-Carroll et al. (2000) with JPL Pasadena)
- Three objectives for inter-planetary trajectory design
  - Minimize time of flight
  - Maximize payload delivered at destination
  - Maximize heliocentric revolutions around the Sun
- NSGA invoked with SEPTOP software for evaluation

Case Study (cont.)

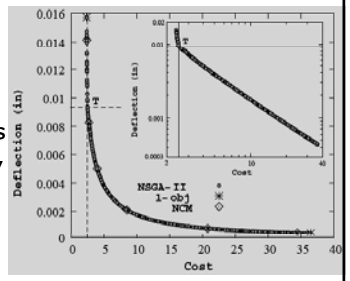
- Identify different trade-off solutions for choosing one (Better and more confident decision-making)
- Inter-planetary trajectory design





Performing Reliable and Confident EMO Simulations

- EMO is a numerical method
- Verify by single-objective optimizations
  - Verify with ε-constraint method
  - Verify the extreme solutions
  - Verify by other means (NBI, NC, etc.)
- Verify by smaller-objective optimizations
- Cluster the frontier
  - Check to see if they are KKT points
  - Norm of gradient expression is close to zero



Coordination of Two Aspects

- Search
  - Exploration and exploitation
  - Recombination+mutation and selection+elitism operators
- Multi-objective preferences
  - Non-domination preference
  - Diversity preservation
- Intertwined, but both must be chosen appropriately for a problem
- Faster approach: Introduce local search

Conclusions

- EMO procedures can find multiple Pareto-optimal solutions in one simulation run
  - Parallel search
  - Computationally faster approach than classical generating methods
- An optimal front provides an idea of objective interactions
  - Decision-making better and easier
  - Not possible before