A Tutorial on Evolutionary Multi-Objective Optimization

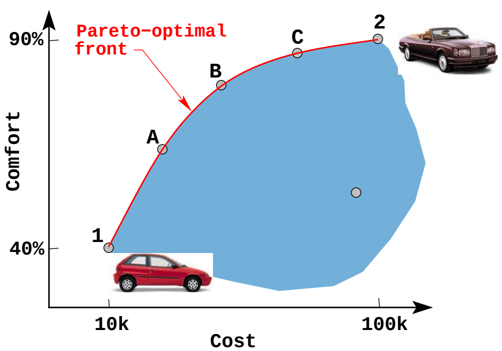
Kalyanmoy Deb
Kanpur Genetic Algorithm Laboratory (KanGAL)
Department of Mechanical Engineering
Indian Institute of Technology Kanpur
Kanpur, Pin 208016 INDIA
deb@iitk.ac.in
http://www.iitk.ac.in/kangal/deb.htm

Overview of the Tutorial

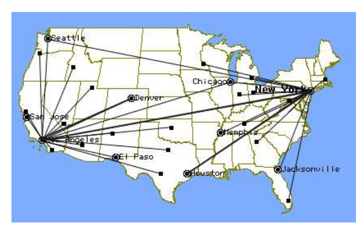
- Multi-objective optimization
- Classical methods
- Evolutionary computing methods (EMO)
  - Differences
  - Non-elitist EMO
  - Elitist EMO
  - Constrained EMO
  - Applications of EMO
  - Salient research issues
- Conclusions

Multi-Objective Optimization

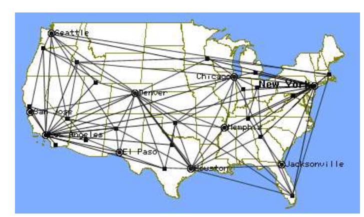
- We often face them



More Examples

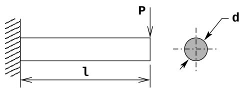


A cheaper but inconvenient flight



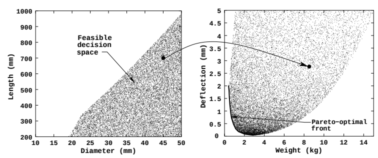
A convenient but expensive flight

Mathematical Programming Problem



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

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

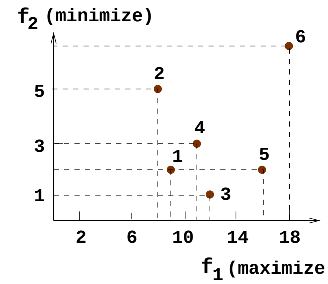


Which Solutions are Optimal?

Relates to the concept of domination

$\mathbf{x}^{(1)}$ dominates $\mathbf{x}^{(2)}$ if

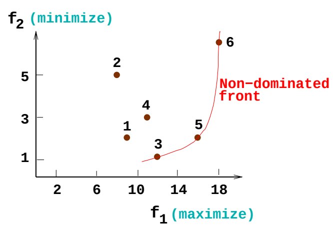
1. $\mathbf{x}^{(1)}$ is no worse than $\mathbf{x}^{(2)}$ in all objectives
2. $\mathbf{x}^{(1)}$ is strictly better than $\mathbf{x}^{(2)}$ in at least one objective



Pareto-Optimal Solutions

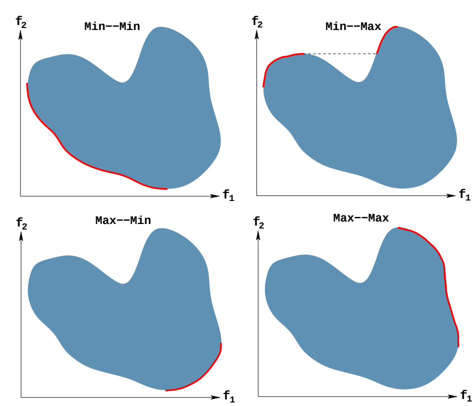
Non-dominated solutions: Among a set of solutions P, the non-dominated set of solutions P' are those that are not dominated by any member of the set P. $O(N \log N)$ algorithms exist.

Pareto-Optimal solutions: When $P = \mathcal{S}$, the resulting P' is Pareto-optimal set



A number of solutions are optimal

Pareto-Optimal Fronts

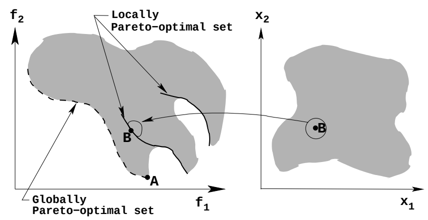


Optimality Conditions

Fritz–John Necessary Condition:
Solution $\mathbf{x}^*$ satisfy

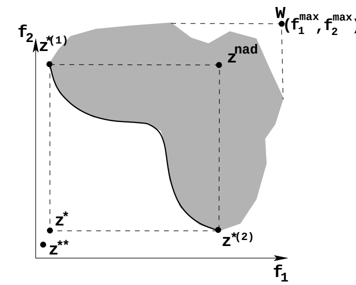
1. $\sum_{m=1}^{M} \lambda_m \nabla f_m(\mathbf{x}^*) - \sum_{j=1}^{J} u_j \nabla g_j(\mathbf{x}^*) = \mathbf{0}$, and
2. $u_j g_j(\mathbf{x}^*) = 0$ for all $j = 1, 2, \ldots, J$.

Like single-objective optimization, local and global P-O fronts exist:



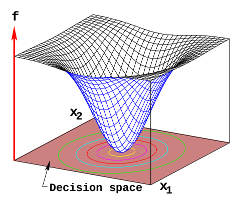
Some Terminologies

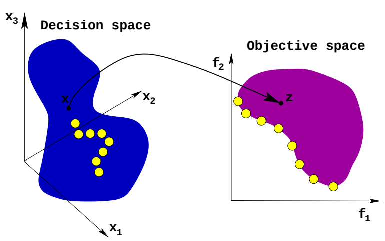
- Ideal point, $\mathbf{z}^*$: nonexistent, lower bound on Pareto-optimal set
- Utopian point, $\mathbf{z}^{**}$: nonexistent
- Nadir point, $\mathbf{z}^{\rm nad}$: upper bound on Pareto-optimal set
- Normalization: $f_i^{\rm norm} = \frac{f_i - z_i^*}{z_i^{\rm nad} - z_i^*}$



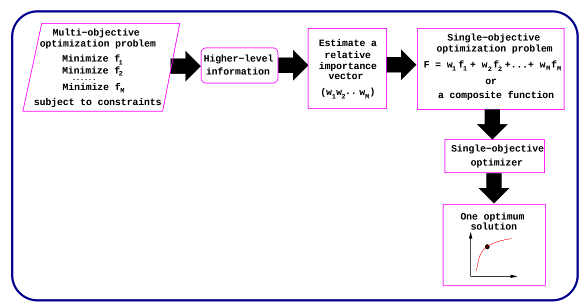
Differences with Single-Objective Optimization

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





Preference-Based Approach



- Classical approaches follow it

Classical Approaches

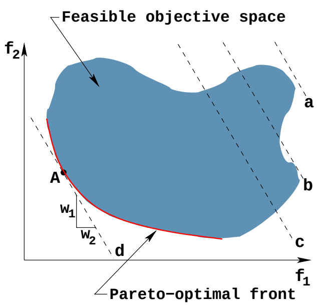
- No Preference methods (heuristic-based)
- Posteriori methods (generating solutions)
- A priori methods (one preferred solution)
- Interactive methods (involving a decision-maker)

Weighted Sum Method

- Construct a weighted sum of objectives and optimize

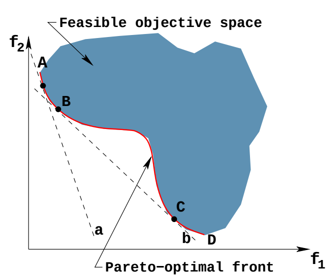
$$F(\mathbf{x}) = \sum_{m=1}^{M} w_m f_m(\mathbf{x}).$$

- User supplies weight vector $\mathbf{w}$



Difficulties with Weighted Sum Method

- Need to know $\mathbf{w}$
- Non-uniformity in Pareto-optimal solutions
- Inability to find some Pareto-optimal solutions

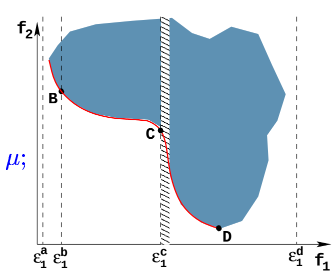


ϵ-Constraint Method

- Optimize one objective, constrain all other

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

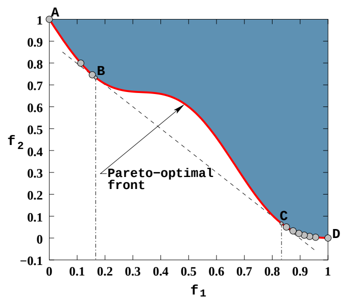
- User supplies a ϵ vector



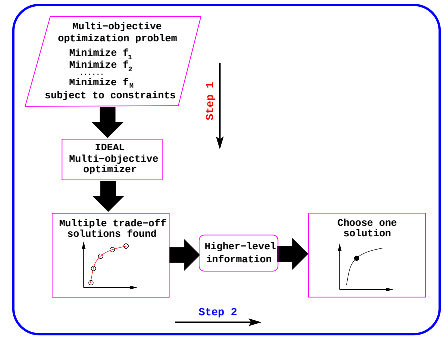
- Need to know relevant ϵ vectors
- Non-uniformity in Pareto-optimal solutions

Difficulties with Most Classical Methods

- 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



Ideal Multi-Objective Optimization

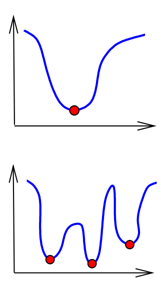


Step 1 Find a set of Pareto-optimal solutions

Step 2 Choose one from the set

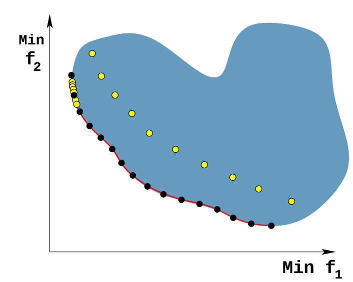
A More Holistic Approach for Optimization

- 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 is a special case of multi-objective optimization



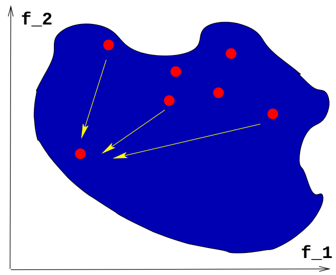
Two Goals in Ideal Multi-Objective Optimization

1. Converge on the Pareto-optimal front
2. Maintain as diverse a distribution as possible



Why Use Evolutionary Algorithms?

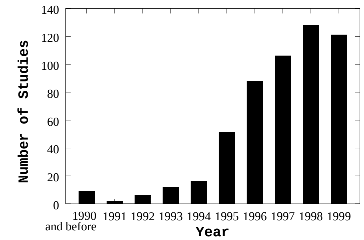
- Population approach suits well to find multiple solutions
- Niche-preservation methods can be exploited to find diverse solutions
- Implicit parallelism helps provide a parallel search



- Multiple applications of classical methods do not constitute a parallel search

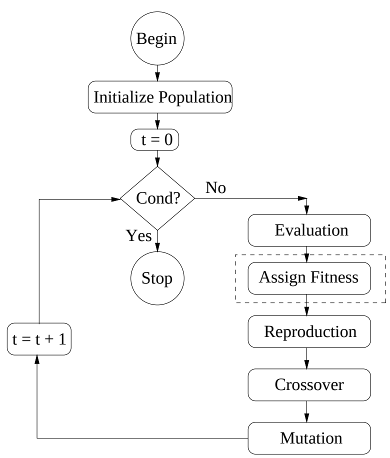
History of Evolutionary Multi-Objective Optimization (EMO)

- 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 the 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.

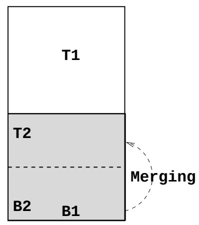
$O(MN^2)$ computational complexity

Finding the Non-dominated Set: An Efficient Approach

Kung et al.'s algorithm (1975)

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

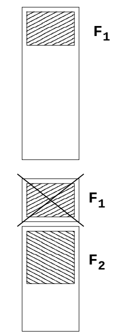
Step 2, Front(P) If $|P| = 1$, return P as the output of **Front(P)**. Otherwise, T = **Front**$(P^{(1)} - - P^{(|P|/2)})$ and B = **Front**$(P^{(|P|/2+1)} - - P^{(|P|)})$. If the i-th solution of B is not dominated by any solution of T, create a merged set $M = T \cup \{i\}$. Return M as the output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

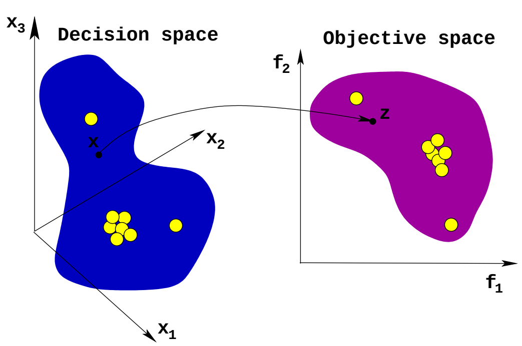
A Simple Non-Dominated Sorting Procedure

- Identify the best non-dominated set
- Discard them from population
- Identify the next-best non-dominated set
- Continue till all solutions are classified
- We discuss a $O(MN^2)$ algorithm later



Which are Less-Crowded Solutions?

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



Non-Elitist EMOs

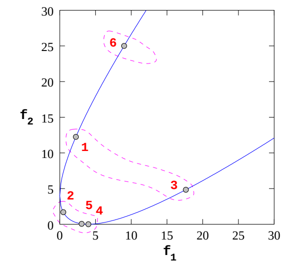
- Vector evaluated GA (VEGA) (Schaffer, 1984)
- Vector optimized EA (VOES) (Kursawe, 1990)
- Weight based GA (WBGA) (Hajela and Lin, 1993)
- Multiple objective GA (MOGA) (Fonseca and Fleming, 1993)
- Non-dominated sorting GA (NSGA) (Srinivas and Deb, 1994)
- Niched Pareto GA (NPGA) (Horn et al., 1994)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

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

Non-Dominated Sorting GA (NSGA)

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



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

Multi-Objective GA (MOGA)

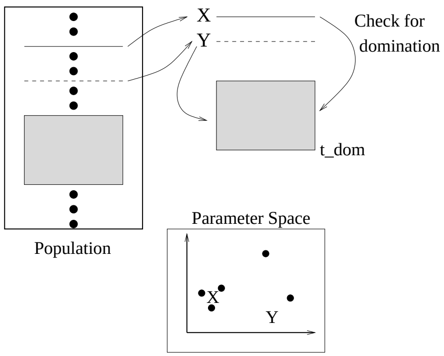
- Count the number of dominated solutions (say n)
- Fitness: $F = n + 1$
- A fitness ranking adjustment
- Niching in *fitness* space
- Rest all are similar to NSGA

	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

Niched Pareto GA (NPGA)

- Solutions in a tournament are checked for domination with respect to a small subpopulation (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 EMOs

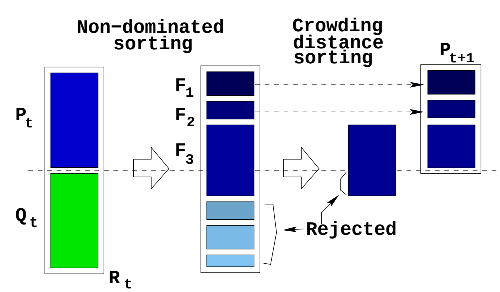
- Elite-preservation is missing
- Elite-preservation is important for proper convergence in SOEAs
- Same is true in EMOs
- Three tasks
  - Elite preservation
  - Progress towards the Pareto-optimal front
  - Maintain diversity among solutions

Elitist EMOs (cont.)

- 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 coevolutionary sharing

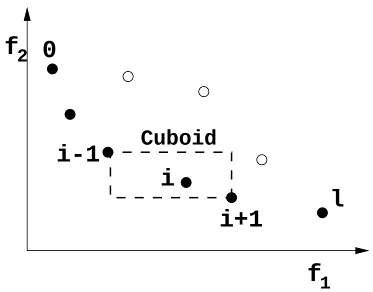
Elitist Non-dominated Sorting Genetic Algorithm (NSGA-II)

Elites are preserved



NSGA-II (cont.)

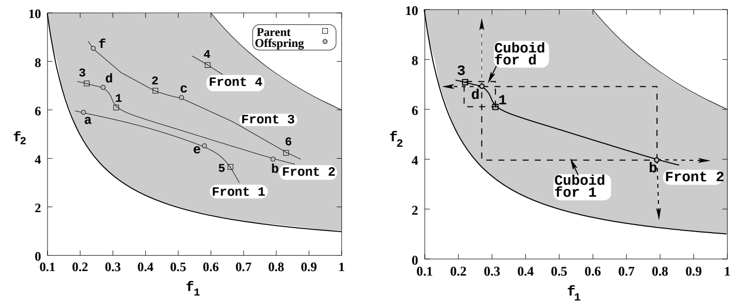
Diversity is maintained: $O(MN \log N)$



Overall Complexity: $O(MN^2)$

An Illustration of NSGA-II

Six parents and six offspring

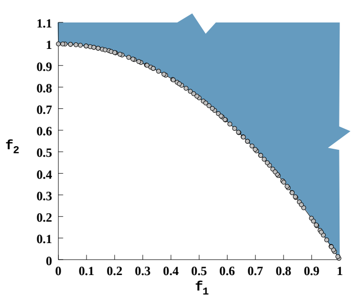


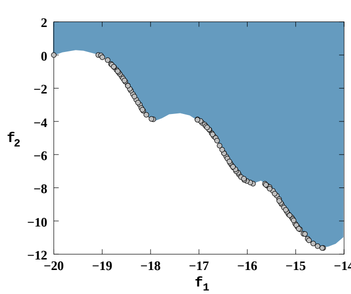
Parents after one iteration: (`a,3,1,e,5,b`)

NSGA-II on Test Problems

(Min) $f_1(\mathbf{x}) = x_1$
(Min) $f_2(\mathbf{x}) = g\left[1 - (f_1/g)^2\right]$
where $g(\mathbf{x}) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i$

(Min) $f_1(\mathbf{x}) = x_1$
(Min) $f_2(\mathbf{x}) = g\left[1 - \sqrt{\frac{f_1}{g}} - \frac{f_1}{g}\sin(10\pi f_1)\right]$
where $g(\mathbf{x}) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i$



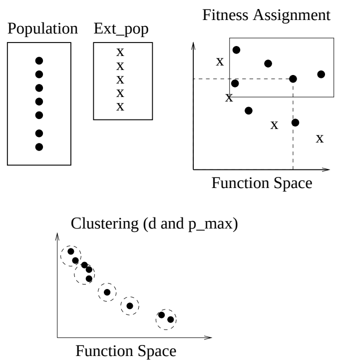


Strength Pareto EA (SPEA)

- 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 current and elite populations
- A clustering technique to maintain diversity in updated external population, when size increases a limit

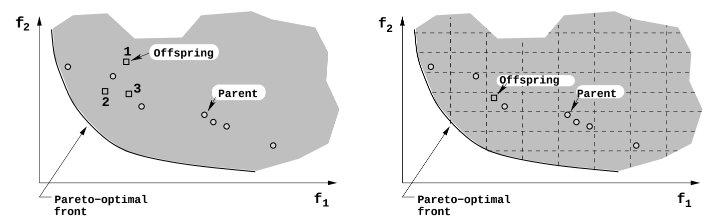
SPEA (cont.)

- Fitness assignment and clustering methods



Pareto Archived ES (PAES)

- An (1+1)-ES
- 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



Comparative Results: Convergence

Algorithm	SCH	FON	POL	KUR
NSGA-II	0.003391	0.001931	0.015553	0.028964
	0	0	0.000001	0.000018
SPEA	0.003403	0.125692	0.037812	0.045617
	0	0.000038	0.000088	0.00005
PAES	0.001313	0.151263	0.030864	0.057323
	0.000003	0.000905	0.000431	0.011989

Algorithm	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
NSGA-II	0.033482	0.072391	0.114500	0.513053	0.296564
	0.004750	0.031689	0.007940	0.118460	0.013135
SPEA	0.001799	0.001339	0.047517	7.340299	0.221138
	0.000001	0	0.000047	6.572516	0.000449
PAES	0.082085	0.126276	0.023872	0.854816	0.085469
	0.008679	0.036877	0.00001	0.527238	0.006664

Comparative Results: Diversity

Algorithm	SCH	FON	POL	KUR
NSGA-II	0.477899	0.378065	0.452150	0.411477
	0.003471	0.000639	0.002868	0.000992
SPEA	1.021110	0.792352	0.972783	0.852990
	0.004372	0.005546	0.008475	0.002619
PAES	1.063288	1.162528	1.020007	1.079838
	0.002868	0.008945	0	0.013772

Algorithm	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
NSGA-II	0.390307	0.430776	0.738540	0.702612	0.668025
	0.001876	0.004721	0.019706	0.064648	0.009923
SPEA	0.784525	0.755148	0.672938	0.798463	0.849389
	0.004440	0.004521	0.003587	0.014616	0.002713
PAES	1.229794	1.165942	0.789920	0.870458	1.153052
	0.004839	0.007682	0.001653	0.101399	0.003916

Constrained Handling

- Penalty function approach

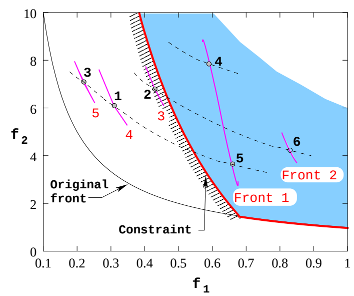
$$F_m = f_m + R_m \Omega(\vec{g}).$$

- 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

Constrain-Domination Principle

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

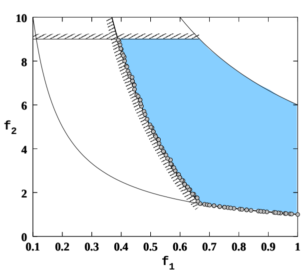
1. Solution i is feasible and solution j is not.
2. Solutions i and j are both infeasible, but solution i has a smaller overall constraint violation.
3. Solutions i and j are feasible and solution i dominates solution j.

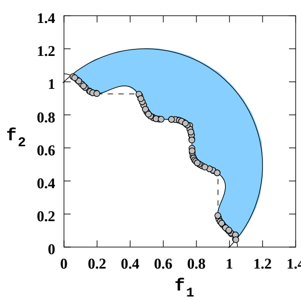


Constrained NSGA-II Simulation Results

(Min) $f_1(\mathbf{x}) = x_1$
(Min) $f_2(\mathbf{x}) = \frac{1+x_2}{x_1}$
$x_2 + 9x_1 \geq 6$
$-x_2 + 9x_1 \geq 1$

(Min) $f_1(\mathbf{x}) = x_1$
(Min) $f_2(\mathbf{x}) = x_2$
$x_1^2 + x_2^2 - 1 - \frac{1}{10}\cos\left(16\tan^{-1}\frac{x_1}{x_2}\right) \geq 0$
$(x_1 - 0.5)^2 + (x_2 - 0.5)^2 \leq 0.5$





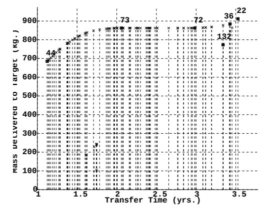
EMO Applications

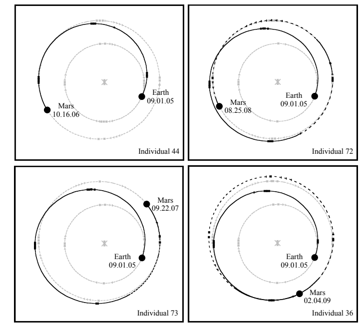
1. Identify different trade-off solutions for choosing one
2. Understanding insights about the problem
   - Reveal common properties among P-O solutions
   - Identify what causes trade-offs
   - Such information are valuable to users
   - May not exist other means of finding above
3. To aid in other optimization tasks

For a Better Decision-Making

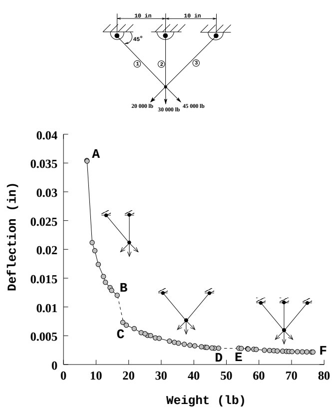
- Spacecraft trajectory optimization (Coverstone-Carroll et al. (2000) with JPL Pasadena)
- Three objectives for inter-planetary trajectory design
  - Minimize time of flight
  - Maximize payload delivered at destination
  - Maximize heliocentric revolutions around the Sun
- NSGA invoked with SEPTOP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

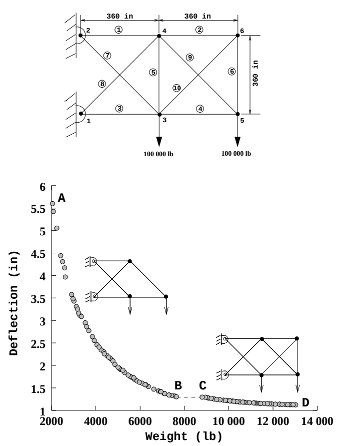
Earth–Mars Rendezvous



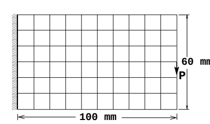


Revealing Salient Insights: Truss Structure Design

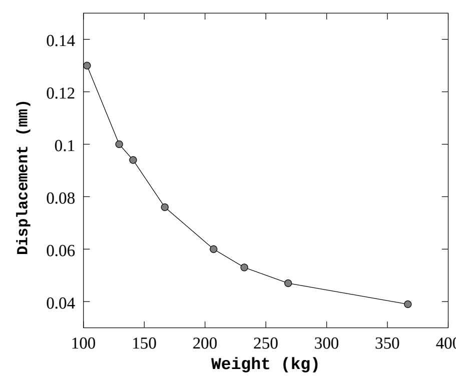




Revealing Salient Insights: A Cantilever Plate Design



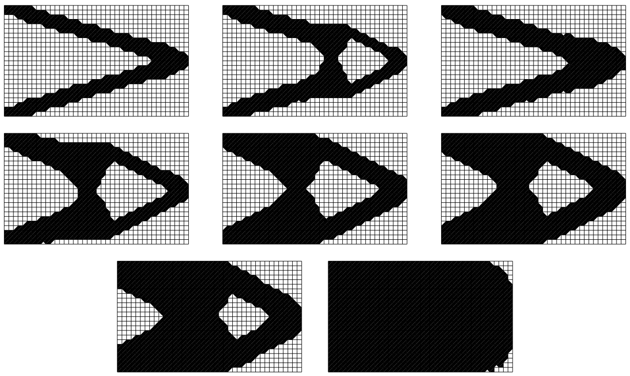
Base plate



Eight trade-off solutions are chosen

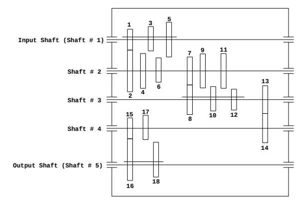
Trade-Off Solutions

- Symmetry in solutions about mid-plane, discovery of stiffener

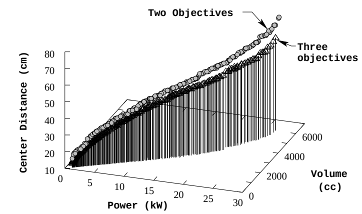


Revealing Salient Insights: Gear-box Design

- A multi-spindle gear-box design
- 29 variables (integer, discrete, real-valued)
- 101 non-linear constraints

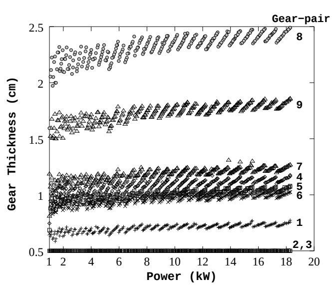


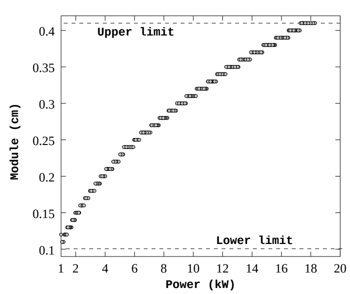
- Important insights obtained (larger module for more power)



Outcome of an Analysis of Solutions

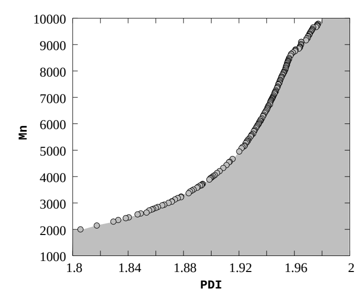
- Module varies proportional to square-root of power ($m \propto \sqrt{p}$)
- Not known earlier





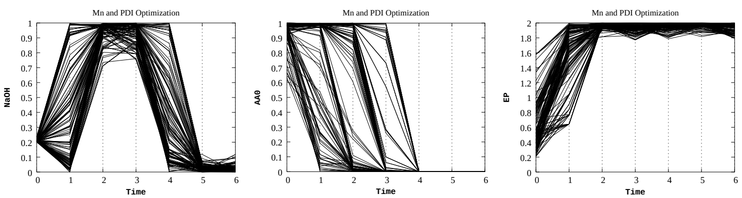
Revealing Salient Insights: Epoxy Polymerization

- Three ingredients (NaOH, EP and AA0) added hourly
- 54 ODEs solved for a 7-hour simulation
- Maximize high chain length (Mn) and minimize polydispersity index (PDI)
- NaOH and AA0 varies in [0,1] and EP in [0,2]
- Total 3×7 or 21 variables



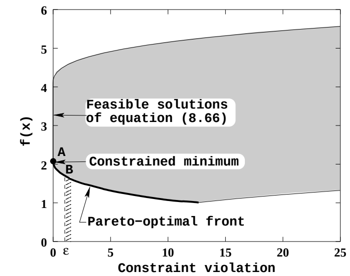
Epoxy Polymerization (cont.)

- A problem having a non-convex Pareto-optimal front
- Some patterns emerge among obtained solutions
- Need to check their chemical significance



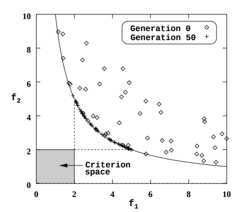
EMO for Other Optimization Tasks

- Constrained handling
  - Constraint violations as additional objectives (Surry, Radcliffe and Boyd, 1995, Coello (2000))
- Find partial front near zero-CV
- May provide a flexible search



Goal Programming and Others

- Goal programming to find multiple solutions
  - Avoids fixing a weight vector (Deb, 2001)



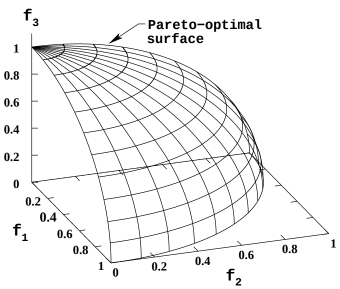
- Genetic programming to reduce bloating: Program size as a second objective (Bleuler et al., 2001)
- Reducing the chance of getting trapped in local optima (Knowles et al., 2001)
- Use secondary objectives for maintaining diversity (Abbass and Deb, 2003, Jensen, 2003)

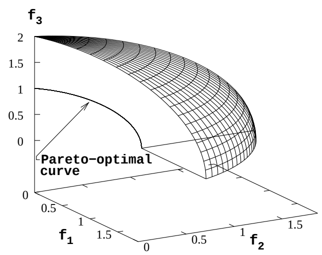
Salient Research Tasks

- Scalability of EMOs to handle more than two objectives
- Mathematically convergent algorithms with guaranteed spread of solutions
- Test problem design
- Performance metrics and comparative studies
- Other EMOs – Multi-modal EMOs, Dynamic EMOs
- Controlled elitism
- Developing practical EMOs – Hybridization, parallelization
- More application case studies

Scalability Issues

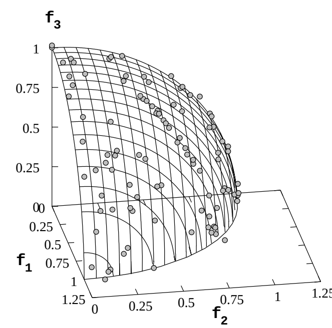
- Pareto-optimal region is a higher-dimensional surface
- Pareto-optimal front may be of smaller dimension

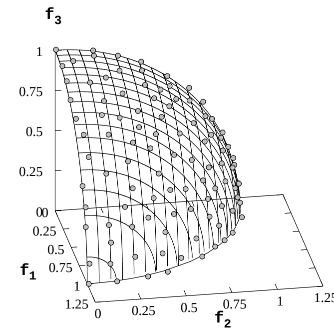




Scalability Issues (cont.)

- Complexity of niching procedures – Who is one's neighbor?
- Algorithms differ in maintaining diversity (NSGA-II vs. SPEA)



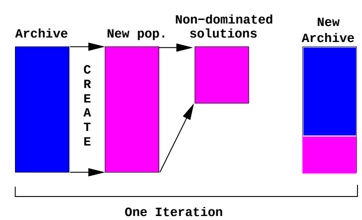


Some Results on Scalability of EMOs

- PESA, SPEA2, and NSGA-II compared up to 8 objectives (Khare, Yao, Deb, 2003)
- PESA best for convergence, but poor in diversity and running time (exponential)
- SPEA2 good for diversity, but poor in convergence and running time
- NSGA-II best for running time and good for diversity, but poor in convergence in higher objectives
- Very different outcome for large number of objectives

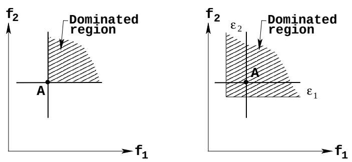
Convergence Issues

- Lukewarm interest till to date
- NSGA-II, SPEA etc. have problem of convergence
  - Pareto-optimal solutions can be lost to maintain a well-diverse set
- Rudolph and Agapie's algorithm for guaranteed convergence



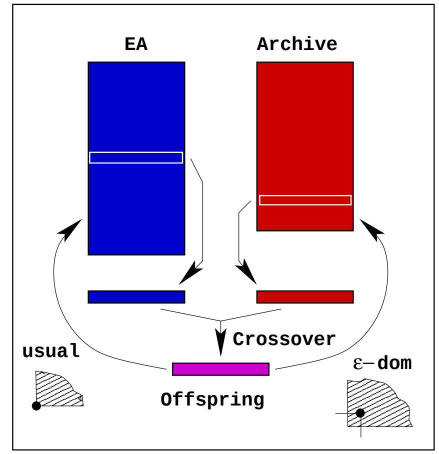
Convergence Issues (cont.)

- Shortcomings of Rudolph and Agapie's algorithm
  - No guarantee on spread of solutions
  - No time complexity measure
- Laumanns et al. (2001) suggest a remedy
  - ϵ-dominance and diversity through hyper-box dominance
  - A new solution is compared with an archive in each iteration
  - ϵ-dominance concept is practical



The ϵ-MOEA

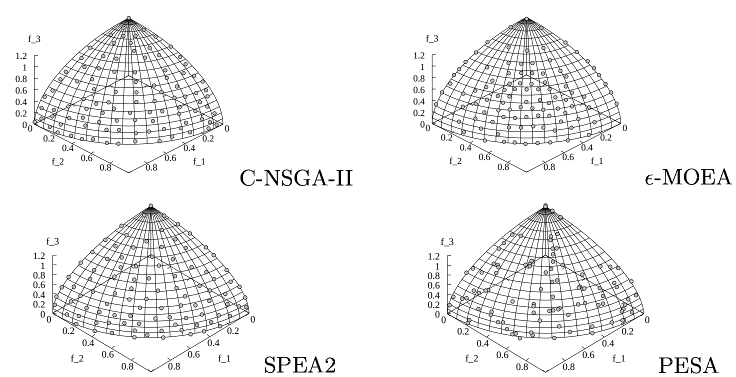
- EA and archive populations evolve
- One EA and one archive member are mated
- Archive update using ϵ-dominance
- EA update using usual dominance



Comparative Study on DTLZ Functions

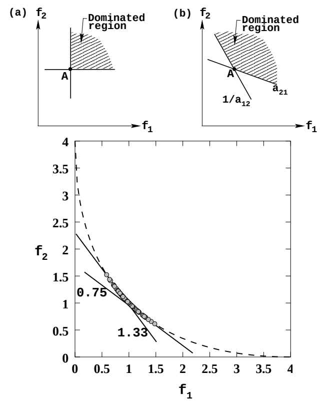
EMO	Convergence measure		Sparsity		Time (sec)	
	Average	Std. Dev.	Average	Std. Dev.	Average	Std. Dev.
DTLZ2						
NSGA-II	0.0137186	0.0020145	0.931111	0.0124474	17.16	0.196
C-NSGA-II	0.0107455	0.0008424	**0.999778**	0.0004968	7837.42	81.254
PESA	**0.0106292**	0.0025483	0.945778	0.0309657	88.01	12.901
SPEA2	0.0126622	0.0009540	0.998889	0.0007855	2164.42	19.858
ϵ-MOEA	0.0108443	0.0002823	0.999104	0.0009316	**2.01**	0.032
DTLZ3						
NSGA-II	0.0149156	0.01028	0.839228	0.02961	136.45	31.080
C-NSGA-II	0.0202315	0.00898	0.995521	0.00613	24046.03	4690.032
PESA	0.0130633	0.00449	0.722296	0.02785	89.49	12.527
SPEA2	0.0122429	0.00194	**0.999771**	0.00031	9080.81	963.723
ϵ-MOEA	**0.0122190**	0.00223	0.993207	0.00974	**9.42**	2.180
DTLZ5						
NSGA-II	0.00208342	11.976e-05	0.953778	0.00992	11.49	0.036
C-NSGA-II	0.00256138	30.905e-05	0.996667	0.00314	1689.16	81.365
PESA	0.00094626	11.427e-05	0.772110	0.02269	53.27	11.836
SPEA2	0.00197846	16.437e-05	**1.000000**	0.00000	633.60	14.082
ϵ-MOEA	**0.000953623**	4.892e-05	0.980867	0.01279	**1.45**	0.051

Test Problem DTLZ2



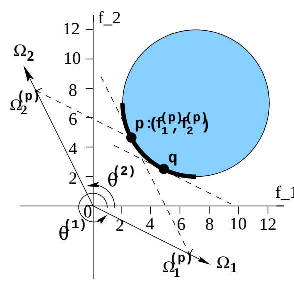
Finding a Partial Pareto-Optimal Set

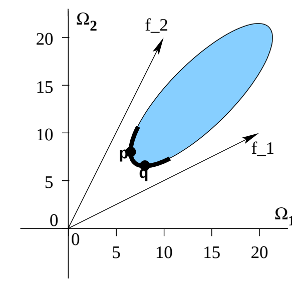
- Using a DM's preference (not for a solution but for a region)
- Guided domination principle (Branke et al., 2000)
- Biased niching approach (Deb, 2002)
- Weighted domination approach (Parmee et al., 2000)



Distributed Computing of Pareto-Optimal Set

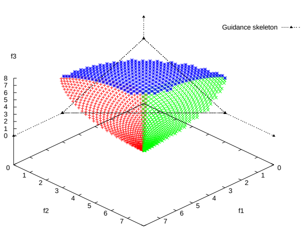
- Guided domination concept to search different parts of P-O region
- Usual island model with migration



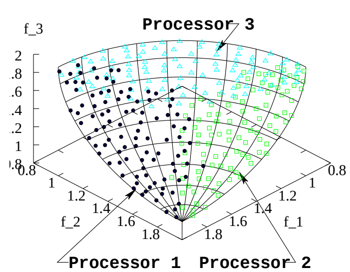


Distributed Computing: A Three-Objective Problem

- Spatial computing, not temporal



Theory



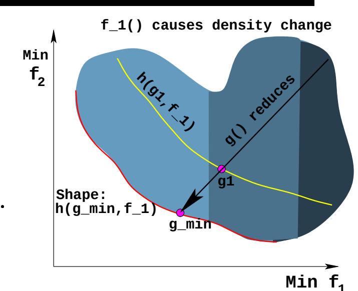
NSGA-II Simulations

Two-Objective Test Problems

- Pareto-optimal front is controllable and known
- ZDT problems:

$$\text{Min.} \quad f_1(\mathbf{x}) = f_1(\mathbf{x}_I),$$
$$\text{Min.} \quad f_2(\mathbf{x}) = g(\mathbf{x}_{II})h(f_1, g).$$

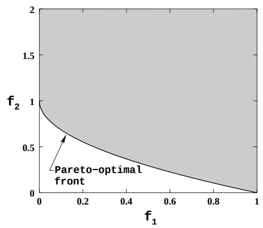
- Choose $f_1()$, $g()$ and $h()$ to introduce various difficulties



Zitzler–Deb–Thiele's Test Problems

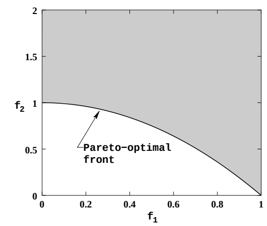
ZDT1

$$f_1(\mathbf{x}) = x_1,$$
$$g(\mathbf{x}) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i,$$
$$h(f_1, g) = 1 - \sqrt{f_1/g}.$$



ZDT2

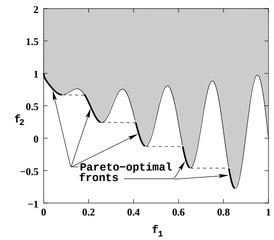
$$f_1(\mathbf{x}) = x_1,$$
$$g(\mathbf{x}) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i,$$
$$h(f_1, g) = 1 - (f_1/g)^2.$$



Zitzler–Deb–Thiele's Test Problems

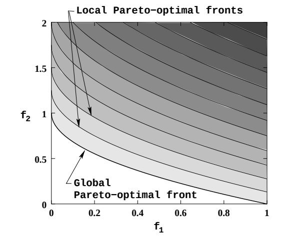
ZDT3

$f_1 = x_1,$
$g = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i,$
$h = 1 - \sqrt{f_1/g} - (f_1/g)\sin(10\pi f_1).$



ZDT4

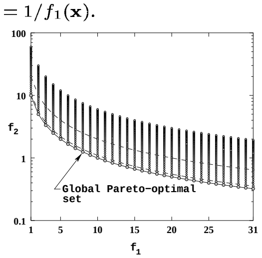
$f_1 = x_1,$
$g = 10n - 9 + \sum_{i=2}^{n}(x_i^2 - 10\cos(4\pi x_i)),$
$h = 1 - \sqrt{f_1/g}.$



Zitzler–Deb–Thiele's Test Problems

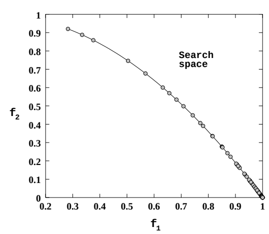
ZDT5

$f_1 = 1 + u(x_1)$
$g = \sum_{i=2}^{11} v(u(x_i))$
$$v = \begin{cases} 2 + u(x_i) & \text{if } u(x_i) < 5, \\ 1 & \text{if } u(x_i) = 5, \end{cases}$$
$h = 1/f_1(\mathbf{x}).$



ZDT6

$f_1 = 1 - \exp(-4x_1)\sin^6(6\pi x_1),$
$g = 1 + 9\left[(\sum_{i=2}^{10} x_i)/9\right]^{0.25},$
$h = 1 - (f_1/g)^2.$



Parameter Interactions

- More difficult problems using parameter interactions
- True variables (y_i) are linearly related to other auxiliary variables (x_i):
$$\vec{x} = M\vec{y}$$
- Fitness computed using $\vec{x}$
- All parameters must change to remain Pareto-optimal

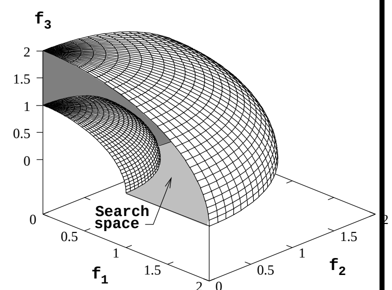
Scalable Test Problems (Deb et al. 2001)

Step 1 Define Pareto-optimal front mathematically

Step 2 Build the objective search space using it

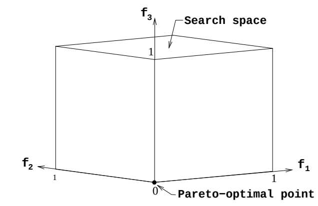
Step 3 Map variable space to objective space

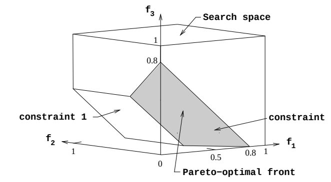
- Scalable DTLZ problems suggested



Constraint Surface Approach

- Define a rectangular hyper-box
- Chop off regions using constraints
- Adv: Easy to construct
- Disadv: Difficult to define Pareto-optimal front





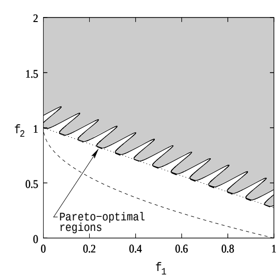
Constrained Test Problem Generator

- Some test problems in Veldhuizen (1999)
- More controllable test problems are called for

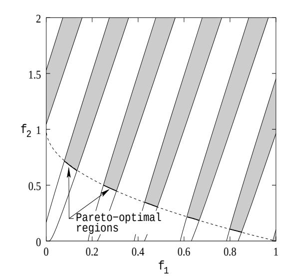
$$
\begin{aligned}
\text{Minimize} \quad & f_1(\mathbf{x}) = x_1 \\
\text{Minimize} \quad & f_2(\mathbf{x}) = g(\mathbf{x}) \left(1 - \frac{f_1(\mathbf{x})}{g(\mathbf{x})}\right) \\
\text{Subject to} \quad & c(\mathbf{x}) \equiv \cos(\theta)(f_2(\mathbf{x}) - e) - \sin(\theta) f_1(\mathbf{x}) \geq \\
& \quad a \left|\sin\left(b\pi \left(\sin(\theta)(f_2(\mathbf{x}) - e) + \cos(\theta) f_1(\mathbf{x})\right)^c\right)\right|^d
\end{aligned}
$$

Various Parameter Settings

$$\theta = -0.2\pi, \quad b = 10, \quad c = 1, \quad e = 1.$$



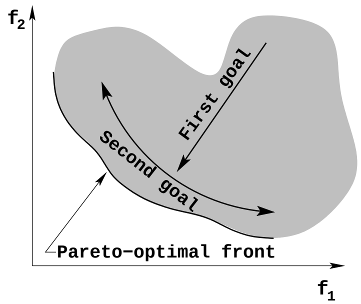
CTP2: $d = 6$ and $a = 0.2$



CTP 7: $\theta = -0.05\pi, a = 40, b = 5, c = 1, d = 6, e = 0$

Performance Metrics

- A recent study by Zitzler et al. suggests at least M metrics
- Two essential metrics (functionally)
  - Convergence measure
  - Diversity measure



Metrics for Convergence

- Error ratio:

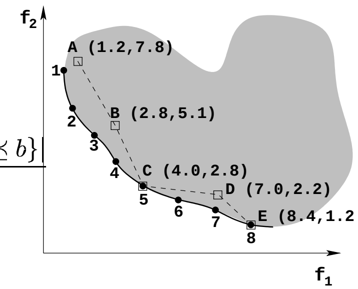
$$ER = \frac{\sum_{i=1}^{|Q|} e_i}{|Q|}$$

- Set Coverage:

$$\mathcal{C}(A,B) = \frac{|\{b \in B | \exists a \in A : a \preceq b\}|}{|B|}$$

- Generational distance:

$$GD = \frac{(\sum_{i=1}^{|Q|} d_i^p)^{1/p}}{|Q|}$$



Metrics for Diversity

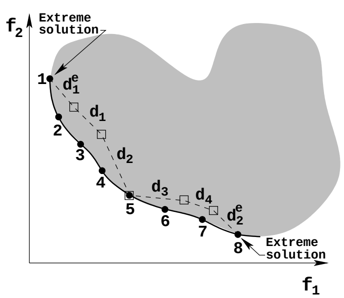
- Spacing:

$$S = \sqrt{\frac{1}{|Q|}\sum_{i=1}^{|Q|}(d_i - \overline{d})^2}$$

- Spread:

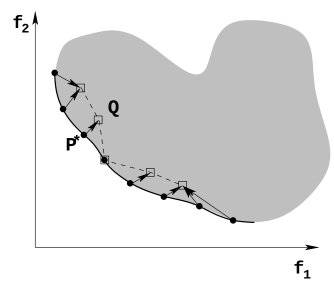
$$\Delta = \frac{\sum_{m=1}^{M} d_m^e + \sum_{i=1}^{|Q|} |d_i - \overline{d}|}{\sum_{m=1}^{M} d_m^e + |Q|\overline{d}}$$

- Chi-square like deviation measure

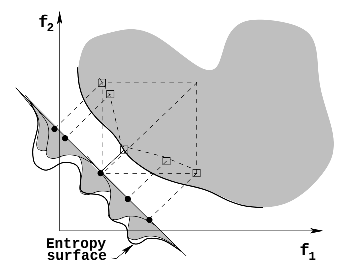


Metrics for Diversity (cont.)

- Distance from P^*

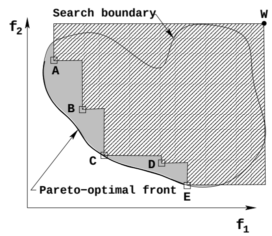


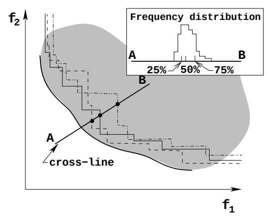
- Entropy measure



Metrics for Convergence and Diversity

- Hypervolume
- Attainment surface method



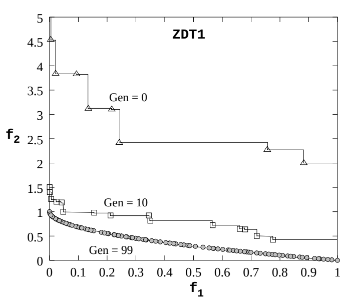


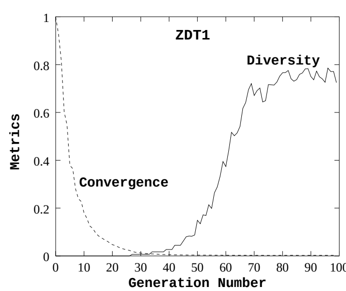
Running Metrics

- Like SGA, define metric that shows generation-wise variation
- Identify non-dominated set $F^{(t)}$ of each population $P^{(t)}$
- Comparison Set (H):
  - If exact P-O front is known, $H = P^*$
  - Else $H = \text{Non-dominated}\left(\cup_{t=0,1,\ldots} F^{(t)}\right)$
- Convergence metric $C^{(t)}$: Average distance of each member of $F^{(t)}$ from H
- Diversity metric $D^{(t)}$: Similar to entropy measure

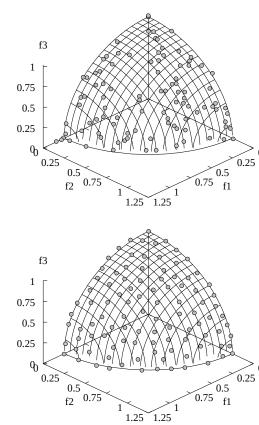
Running Metrics on ZDT1

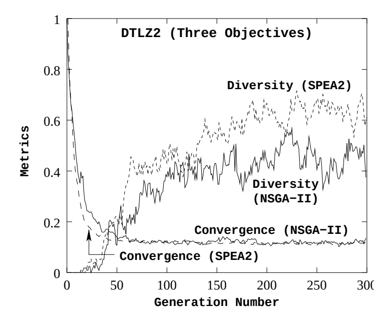
- Using Pareto-optimal solutions





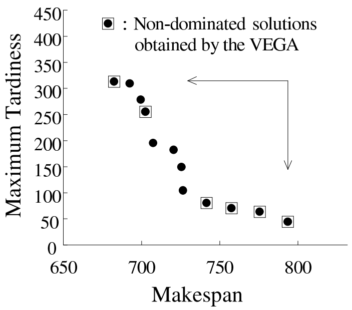
Running Metrics (cont.)

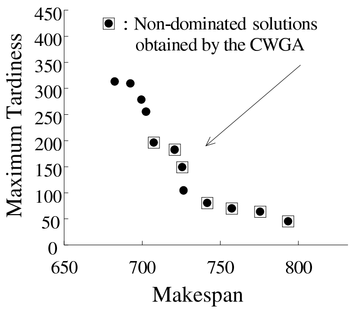




Scheduling EMOs

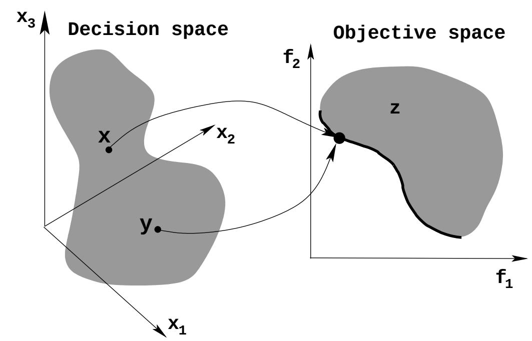
- Objective space niching allows a straightforward application
- Most techniques use a local search
- Job-shop scheduling (Ishibuchi and Murata, 1998)





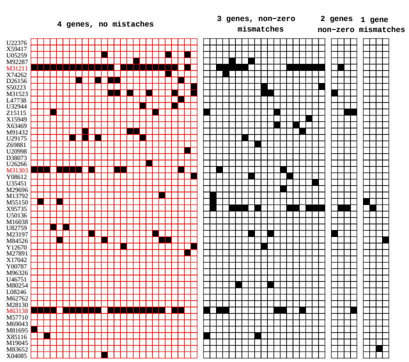
Multi-Modal EMOs

- Different solutions having identical objective values
- Multi-modal Pareto-optimal solutions: Design, Bioinformatics



Multiple Gene Subsets for Leukemia Samples

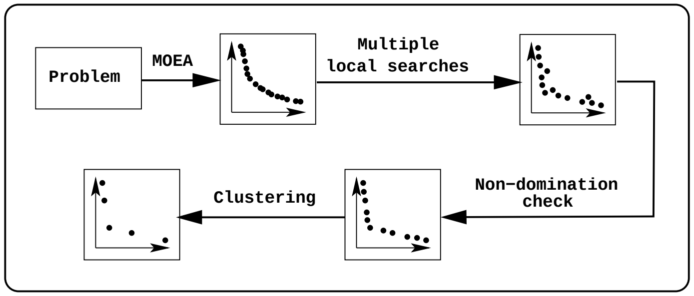
- Deb and Reddy (BioSystems, 2003)
- Multiple (26) four-gene combinations for 100% classification
- Discovery of some common genes



Hybrid EMOs

- Combine EAs with a local search method
  - Better convergence
  - Faster overall optimization
- Two hybrid approaches
  - Local search to update each solution in an EA population (Ishubuchi and Murata, 1998; Jaskiewicz, 1998)
  - First EA and then apply a local search (Deb and Goel, 2000)

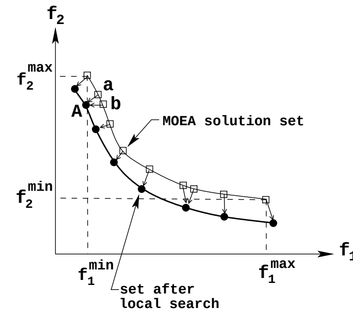
Posteriori Approach in an EMO



- Which objective to use in local search?

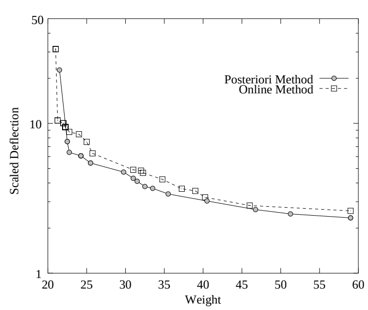
An Idea for Local Search

- Extreme solutions are assigned extreme weights
- Linear relation between weight and fitness
- Many solution can converge to same solution after local search



Posteriori Versus Online Approaches

- Cantilever plate design
- Compared for identical evaluations
- Posteriori finds a better front



Which Pareto-Optimal Solution to Choose?

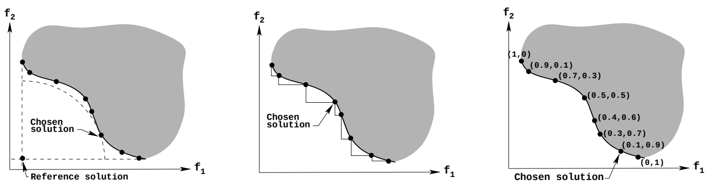
- Needs to involve a decision-maker (DM)
- Interactive EMO is called for – Not much study yet
- A few difficulties:
  - The act of a DM makes it a single-obj. problem
  - But, obj. is not known precisely and changes with iteration
  - EMO finds many solutions, but only one is desired
  - Is DM interested in evaluating more than one solution?

- EMO as a starter, then a classical approach

A Possible Interactive EMO

EMO: Find potentially good solutions – robust, knee-like, etc.

Classical: Concentrate in an area based DM's preference



Conclusions

- Ideal multi-objective optimization is generic and pragmatic
- Evolutionary algorithms are ideal candidates
- Many efficient algorithms exist, more efficient ones are needed
- With some salient research studies, EMOs will revolutionize the act of optimization
- EAs have a definite edge in multi-objective optimization and should become more useful in practice in coming years

EMO Resources

Books

- C. A. C. Coello, D. A. VanVeldhuizen, and G. Lamont. *Evolutionary Algorithms for Solving Multi-Objective Problems.* Boston, MA: Kluwer Academic Publishers, 2002.
- K. Deb. *Multi-objective optimization using evolutionary algorithms.* Chichester, UK: Wiley, 2001. (Second edition, with exercise problems)
- Paper Repository: `http://www.lania.mx/~ccoello/EMOO/`

Conference Proceedings

- Zitzler, E., Deb, K., Thiele, L., Coello, C. A. C. and Corne, D. (Eds) (2001). *Evolutionary Multi-Criterion Optimization (Lecture Notes in Computer Science 1993).* Heidelberg: Springer.

EMO Resources (cont.)

Conference Proceedings (cont.)

- Fonseca, C., Zitzler, E., Deb, K., Fleming, P. and Thiele, L. (Eds) (2003). *Evolutionary Multi-Criterion Optimization (Lecture Notes in Computer Science 2632).* Heidelberg: Springer.

- EMO-2005 in Mexico (`http://www.cimat.mx/emo2005/`)

Mailing List

- emo-list@ualg.pt
- MCRIT-L@LISTSERV.UGA.EDU

Public-Domain Source Codes

- NSGA-II in C: `http://www.iitk.ac.in/kangal/soft.htm`
- SPEA2 and others: `http://www.tik.ee.ethz.ch/pisa`
  Java codes: University of Dortmund