Multiobjective GAs

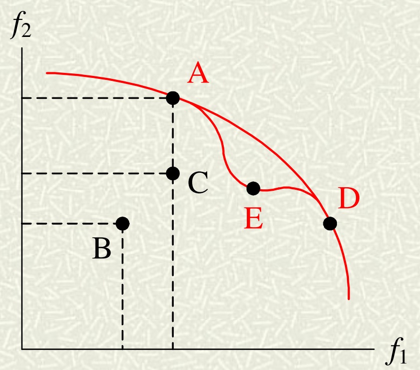
Deb et al., A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation* 6 (2002)

Multiobjective optimization

"max" or "min" $\{f_1(x), f_2(x),..., f_M(x)\}$
subject to $x \in X$

where x is the solution vector and X is the feasible solution space

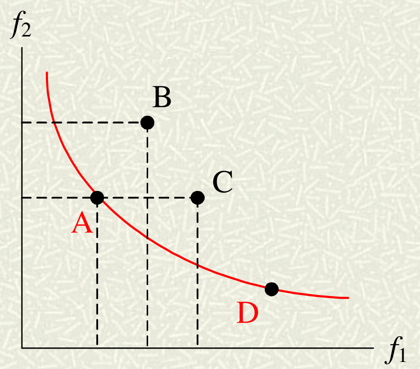
Multiobjective optimization (cont.)



f_1 and f_2 are to be maximized

- A dominates B (better in both f_1 and f_2)
- A dominates C (same in f_1 but better in f_2)
- A does not dominate D
- A and D are in the "pareto optimal frontier" (or efficient frontier or set)
- Frontier need not be concave (or convex), E is also nondominated

Multiobjective optimization (cont.)



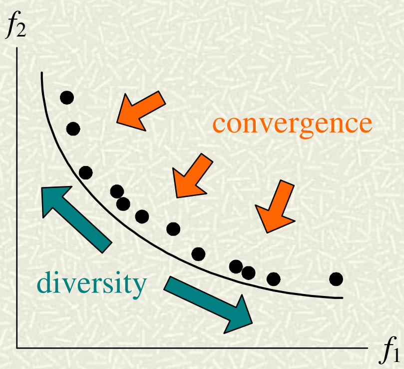
f_1 and f_2 are to be minimized

- A dominates B (better in both f_1 and f_2)
- A dominates C (same in f_2 but better in f_1)
- A does not dominate D
- A and D are in the "pareto optimal frontier" (or efficient frontier or set)

Pareto optimal frontier

- In the absence of weights for objectives, one of the pareto optimal (nondominated) solutions cannot be said to be better than the other; therefore it is desirable to find all
- Classical optimization methods including MCDM methods can find one such solution at a time
- With a population of solutions, GAs seem to be well-suited for approximating the pareto optimal frontier in a single run

Desirable MOEA features



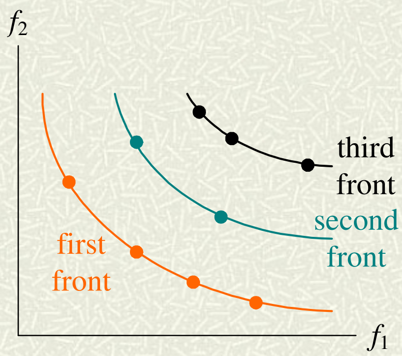
f_1 and f_2 are to be minimized

- Convergence to pareto optimal frontier
- Diversity (representation of the entire pareto optimal frontier)

Main MOEA design issues

- Solution representation, crossover and mutation are problem dependent, but are not affected by multiple objectives
- Main decisions in the presence of multiple objectives are:
  - Fitness assignment (major issue)
  - Parent selection, and
  - Replacement (forming population for next generation)

Nondominated sorting in MOEA



f_1 and f_2 are to be minimized

Fitness assignment:

- Solutions in the first nondominated front have the highest fitness (they are all ranked 1)
- Solutions in the same front have the same fitness (they all have the same rank)

Criticism of NSGA

- Nondominated sorting used in NSGA is expensive
  - Compare each solution in population with every other to find the first nondominated front
  - Temporarily leave out the solutions in the first front and repeat the comparison to find the second front, and then the third front, and so on
  - Runs in $O(MN^3)$ where M is the number of objectives and N is the population size (third N is for the maximum number of fronts)

Criticism of NSGA (cont.)

- Specifying the "sharing" parameter σ_{share} for ensuring diversity is difficult (nonparametric diversity preservation is desirable)
- If the distance between two solutions (measured by a metric) < σ_{share} then they share each other's fitness (used to reduce the chance of keeping two close solutions in the next population)
- NSGA lacks elitism, which can speed up the performance and prevent loss of good solutions

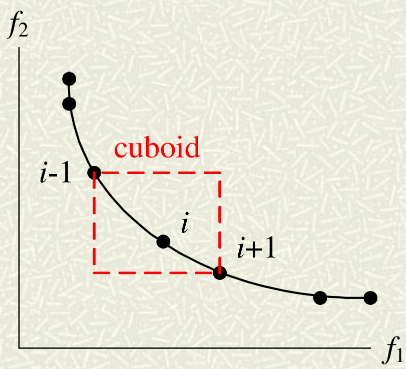
NSGA II features:
Fast nondominated sorting

1. For each solution p in population, find
   n_p: number of solutions that dominate p
   S_p : set of solutions that p dominates
2. Place all p with $n_p = 0$ in set F_1, the first front (R_p=1)
3. For each $p \in F_1$, visit each $q \in S_p$ and reduce n_q by one. In doing this, if n_q becomes 0 then place q in set F_2 (q belongs to the second front, R_q=2)
4. Repeat Step 3 with each member of F_2 to find the third front, and so on

Runs in $O(MN^2)$

NSGA II features:
Diversity preservation

- "Sharing" is replaced with "crowded comparison"
- "Crowding distance" of solution i in a front is the average side length of the cuboid



f_1 and f_2 are to be minimized

NSGA II features:
Diversity preservation

1. Sort all l solutions in a front in ascending order of f_m and compute

$$CD_{im} = \frac{f_m(x_{i+1}) - f_m(x_{i-1})}{f_m(x_{\max}) - f_m(x_{\min})}, \; i = 2, \ldots l-1$$

2. Repeat Step 1 for each objective and find the crowding distance of solution i as

$$CD_i = \sum_{m=1}^{M} CD_{im}$$

Runs in $O(MN \log N)$

NSGA II features:
Crowded comparison operator

- Given two solutions i and j, solution i is preferred to solution j if $R_i < R_j$ or ($R_i = R_j$ and $CD_i > CD_j$)
- Between two solutions with different nondomination ranks, the one with the lower (better) rank is preferred
- When two solutions have the same nondomination rank (belong to the same front), the one located in a less crowded region of the front is preferred

NSGA II algorithm

For minimization and in generation t:

1. Using binary tournament selection and problem dependent crossover and mutation operators, generate child population Q_t from parent population P_t
2. Let $R_t = P_t \cup Q_t$ and sort R_t based on nondomination (selection from combined parent and child population ensures elitism)
3. From $2N$ solutions in R_t, select N best solutions by using the crowded comparison operator to form P_{t+1}

NSGA II algorithm (cont.)

- If $|F_1|<N$ then solutions from F_2 and then F_3 and so on are selected to form P_{t+1}
- Only for the last front included in P_{t+1}, selection is based on the crowding distance
- Overall complexity is $O(MN^2)$, which is governed by nondominated sorting

Comparable MOEAs

- SPEA (Zitzler and Thiele 1998)
  - Keeps all nondominated (elite) solutions discovered so far and lets them participate in all genetic operators in every generation
  - Fitness is based on the number of dominated solutions (a dominated solution has lower fitness than the worst nondominated solution)
  - Uses clustering of similar solutions to preserve diversity
  - Runs in $O(MN^3)$, can be reduced to $O(MN^2)$

Comparable MOEAs (cont.)

- PAES (Knowles and Corne 1999)
  - Single parent, single offspring EA
  - If offspring dominates parent, accept offspring as the next parent
  - If parent dominates offspring, discard offspring and generate a new offspring
  - If neither dominate, compare them with best solutions in archive in terms of domination and nearness
  - Accept offspring and put it in archive if it is in a less crowded region of solution space (far from others)
  - Runs in $O(MN^2)$

Comparable MOEAs (cont.)

- Elitist MOEA (Rudolph 1999)
  - Compares all nondominated offspring with all parents to form an overall nondominated population for the next generation
  - Convergence to pareto optimal frontier is proved with this strategy
  - Has no explicit diversity preservation mechanism
  - Runs in $O(MN^2)$
  - Not used for comparison but inspired the elitism in NSGA II

Operators and parameter settings

- Single point crossover and bit mutation for binary coded SPEA, PAES and NSGA II
- Simulated binary crossover (SBX) and polynomial mutation for real coded NSGA II
- Population size is 100
- Run for maximum 250 generations
- 25,000 function evaluations for all algorithms
- Run for 500 generations works better

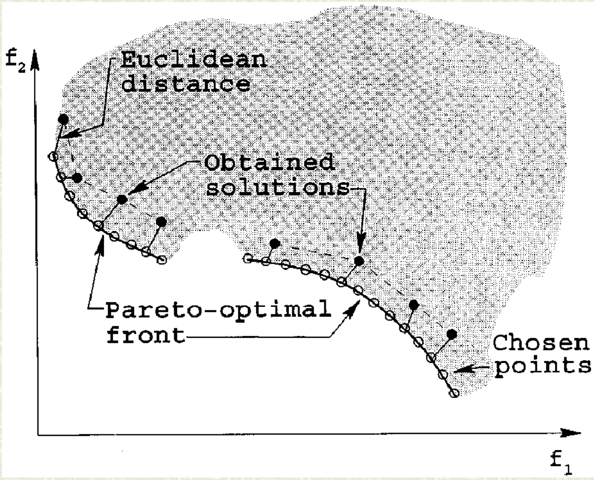
Test problems

(All functions are to be minimized)

Problem	n	Variable bounds	Objective functions	Optimal solutions	Comments
SCH	1	$[-10^3, 10^3]$	$f_1(x) = x^2$ $f_2(x) = (x-2)^2$	$x \in [0, 2]$	convex
FON	3	$[-4, 4]$	$f_1(\mathbf{x}) = 1 - \exp\left(-\sum_{i=1}^{3}\left(x_i - \frac{1}{\sqrt{3}}\right)^2\right)$ $f_2(\mathbf{x}) = 1 - \exp\left(-\sum_{i=1}^{3}\left(x_i + \frac{1}{\sqrt{3}}\right)^2\right)$	$x_1 = x_2 = x_3$ $\in [-1/\sqrt{3}, 1/\sqrt{3}]$	nonconvex
POL	2	$[-\pi, \pi]$	$f_1(\mathbf{x}) = \left[1 + (A_1 - B_1)^2 + (A_2 - B_2)^2\right]$ $f_2(\mathbf{x}) = \left[(x_1 + 3)^2 + (x_2 + 1)^2\right]$ $A_1 = 0.5 \sin 1 - 2 \cos 1 + \sin 2 - 1.5 \cos 2$ $A_2 = 1.5 \sin 1 - \cos 1 + 2 \sin 2 - 0.5 \cos 2$ $B_1 = 0.5 \sin x_1 - 2 \cos x_1 + \sin x_2 - 1.5 \cos x_2$ $B_2 = 1.5 \sin x_1 - \cos x_1 + 2 \sin x_2 - 0.5 \cos x_2$	(refer [1])	nonconvex, disconnected
KUR	3	$[-5, 5]$	$f_1(\mathbf{x}) = \sum_{i=1}^{n-1}\left(-10 \exp\left(-0.2\sqrt{x_i^2 + x_{i+1}^2}\right)\right)$ $f_2(\mathbf{x}) = \sum_{i=1}^{n}\left(\lvert x_i \rvert^{0.8} + 5 \sin x_i^3\right)$	(refer [1])	nonconvex
ZDT1	30	$[0, 1]$	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x})\left[1 - \sqrt{x_1/g(\mathbf{x})}\right]$ $g(\mathbf{x}) = 1 + 9\left(\sum_{i=2}^{n} x_i\right)/(n-1)$	$x_1 \in [0, 1]$ $x_i = 0$, $i = 2, \ldots, n$	convex
ZDT2	30	$[0, 1]$	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x})\left[1 - (x_1/g(\mathbf{x}))^2\right]$ $g(\mathbf{x}) = 1 + 9\left(\sum_{i=2}^{n} x_i\right)/(n-1)$	$x_1 \in [0, 1]$ $x_i = 0$, $i = 2, \ldots, n$	nonconvex
ZDT3	30	$[0, 1]$	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x})\left[1 - \sqrt{x_1/g(\mathbf{x})} - \frac{x_1}{g(\mathbf{x})} \sin(10\pi x_1)\right]$ $g(\mathbf{x}) = 1 + 9\left(\sum_{i=2}^{n} x_i\right)/(n-1)$	$x_1 \in [0, 1]$ $x_i = 0$, $i = 2, \ldots, n$	convex, disconnected
ZDT4	10	$x_1 \in [0, 1]$ $x_i \in [-5, 5]$, $i = 2, \ldots, n$	$f_1(\mathbf{x}) = x_1$ $f_2(\mathbf{x}) = g(\mathbf{x})\left[1 - \sqrt{x_1/g(\mathbf{x})}\right]$ $g(\mathbf{x}) = 1 + 10(n-1) + \sum_{i=2}^{n}\left[x_i^2 - 10 \cos(4\pi x_i)\right]$	$x_1 \in [0, 1]$ $x_i = 0$, $i = 2, \ldots, n$	nonconvex
ZDT6	10	$[0, 1]$	$f_1(\mathbf{x}) = 1 - \exp(-4x_1) \sin^6(6\pi x_1)$ $f_2(\mathbf{x}) = g(\mathbf{x})\left[1 - (f_1(\mathbf{x})/g(\mathbf{x}))^2\right]$ $g(\mathbf{x}) = 1 + 9\left[\left(\sum_{i=2}^{n} x_i\right)/(n-1)\right]^{0.25}$	$x_1 \in [0, 1]$ $x_i = 0$, $i = 2, \ldots, n$	nonconvex, nonuniformly spaced

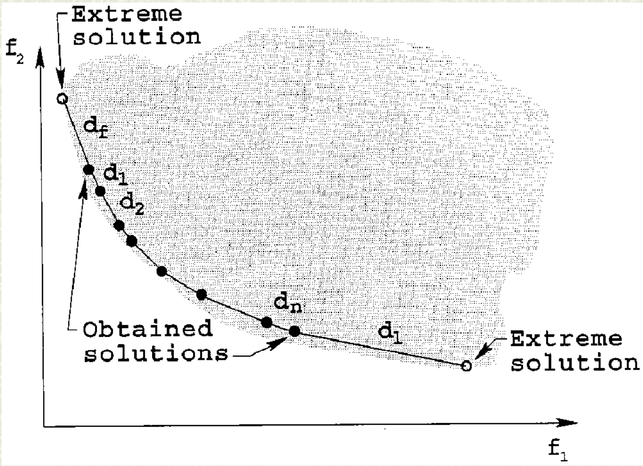
Performance measures

Distance metric



Performance measures

Diversity metric



Computational results

Convergence

Table II

Mean (first rows) and variance (second rows) of the convergence metric (smaller is better)

Algorithm	SCH	FON	POL	KUR	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
NSGA-II	0.003391	0.001931	0.015553	0.028964	0.033482	0.072391	0.114500	0.513053	0.296564
Real-coded	0	0	0.000001	0.000018	0.004750	0.031689	0.007940	0.118460	0.013135
NSGA-II	0.002833	0.002571	0.017029	0.028951	0.000894	0.000824	0.043411	3.227636	7.806798
Binary-coded	0.000001	0	0.000003	0.000016	0	0	0.000042	7.30763	0.001667
SPEA	0.003403	0.125692	0.037812	0.045617	0.001799	0.001339	0.047517	7.340299	0.221138
	0	0.000038	0.000088	0.00005	0.000001	0	0.000047	6.572516	0.000449
PAES	0.001313	0.151263	0.030864	0.057323	0.082085	0.126276	0.023872	0.854816	0.085469
	0.000003	0.000905	0.000431	0.011989	0.008679	0.036877	0.00001	0.527238	0.006664

Computational results
Diversity

Table III

Mean (first rows) and variance (second rows) of the diversity metric

(smaller is better)

Algorithm	SCH	FON	POL	KUR	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
NSGA2R	0.477899	0.378065	0.452150	0.411477	0.390307	0.430776	0.738540	0.702612	0.668025
Real-coded	0.003471	0.000639	0.002868	0.000992	0.001876	0.004721	0.019706	0.064648	0.009923
NSGA-II	0.449265	0.395131	0.503721	0.442195	0.463292	0.435112	0.575606	0.479475	0.644477
Binary-coded	0.002062	0.001314	0.004656	0.001498	0.041622	0.024607	0.005078	0.009841	0.035042
SPEA	1.021110	0.792352	0.972783	0.852990	0.784525	0.755148	0.672938	0.798463	0.849389
	0.004372	0.005546	0.008475	0.002619	0.004440	0.004521	0.003587	0.014616	0.002713
PAES	1.063288	1.162528	1.020007	1.079838	1.229794	1.165942	0.789920	0.870458	1.153052
	0.002868	0.008945	0	0.013772	0.004839	0.007682	0.001653	0.101399	0.003916

Sample plots

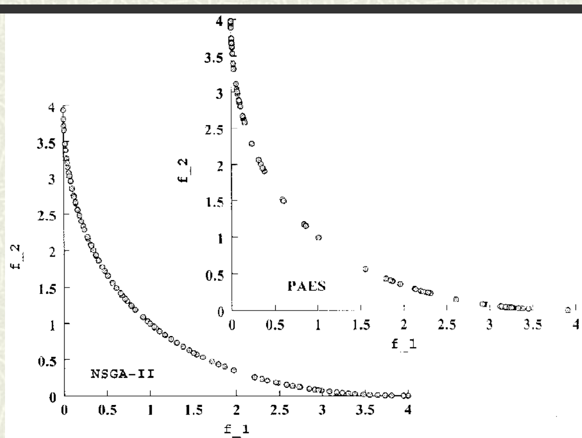


Fig. 5. NSGA-II finds better spread of solutions than PAES on SCH.

Sample plots (cont.)

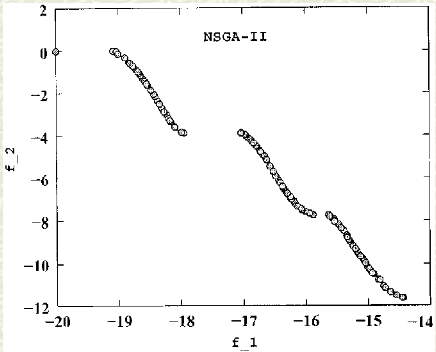


Fig. 6. Nondominated solutions with NSGA-II (real-coded) on KUR.

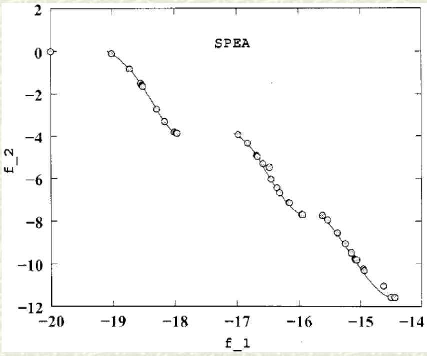


Fig. 7. Nondominated solutions with SPEA on KUR.

Sample plots (cont.)

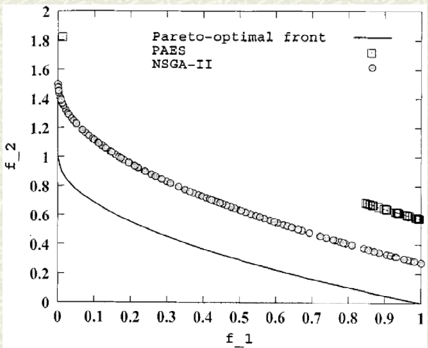


Fig. 10. NSGA-II finds better convergence and spread of solutions than PAES on ZDT4.

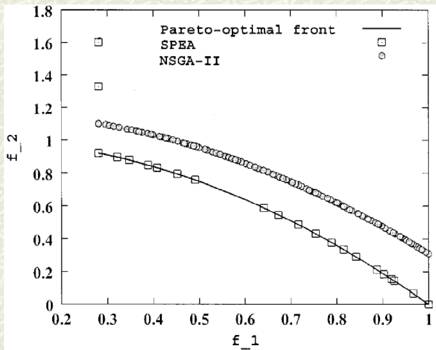


Fig. 11. Real-coded NSGA-II finds better spread of solutions than SPEA on ZDT6, but SPEA has a better convergence.

Constrained optimization

When two offspring are generated from two parents:

- If both offspring are feasible, leave the choice to the crowded comparison operator
- If one offspring is feasible and the other is infeasible, choose the feasible one
- If both are infeasible, choose the one with smaller overall constraint violation

Sample plots

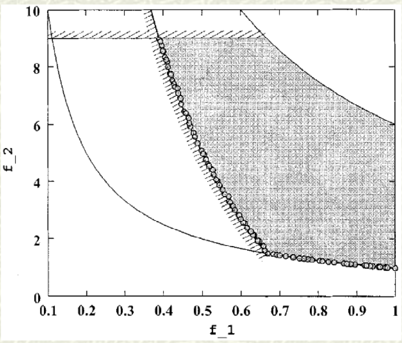


Fig. 14. Obtained nondominated solutions with NSGA-II on the constrained problem CONSTR.

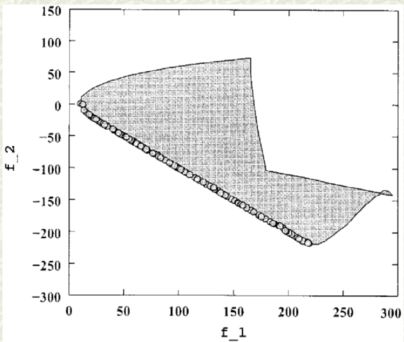


Fig. 16. Obtained nondominated solutions with NSGA-II on the constrained problem SRN.

Sample plots (cont.)

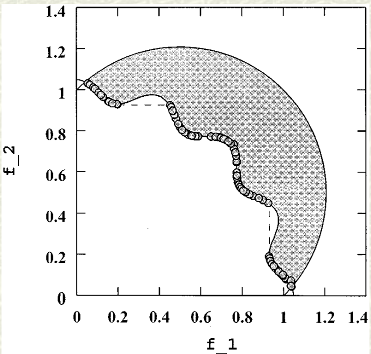


Fig. 18. Obtained nondominated solutions with NSGA-II on the constrained problem TNK.

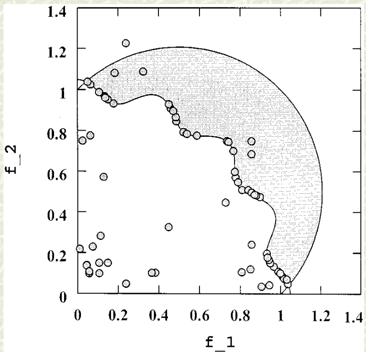


Fig. 19. Obtained nondominated solutions with Ray–Tai–Seow's algorithm on TNK (many are infeasible).