Recent Advances in Evolutionary Multi-Criterion Optimization (EMO)

Kalyanmoy Deb
Koenig Endowed Chair Professor
Michigan State University
East Lansing, USA
kdeb@egr.msu.edu
http://www.egr.msu.edu/~kdeb

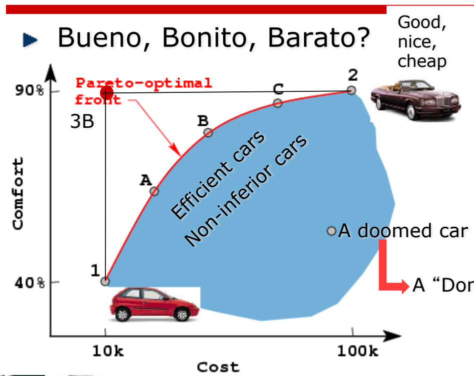
GECCO '17, Berlin, Germany
ACM ISBN 978-1-4503-4939-0/17/07.
http://dx.doi.org/10.1145/3067695.3067697

Outline of the Tutorial

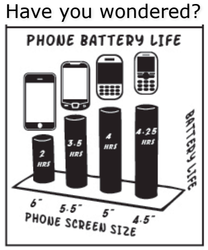
- Multi-criterion Optimization Basics
- Evol. Multi-criterion Optimization (EMO) Basics
  - Past 24 Years in a Flash (1993 – Today)
  - GECCO17-EMO Tutorial by Dimo Brockoff
- Advanced EMO Topics
  - Too many to cover, discuss main advanced topics
  - Many-objective and massive-objective optimization, Objective reduction, *Innovization*, Distributed computing, Visualization and decision-making, Problems with Uncertainty, Metamodel based EMO, Dynamic EMO, Bilevel EMO, Theoretical convergence measure, knee finding, Test problem construction, Extreme solutions

Multi-Criterion Optimization: *Basics*

- Bueno, Bonito, Barato?



Have you wondered?

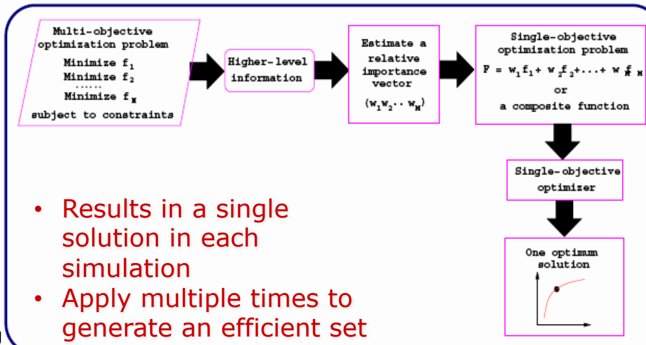


Formally, as a Mathematical Programming Problem

- Multiple objectives, constraints, and variables

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

Classical MO Principle



- Results in a single solution in each simulation
- Apply multiple times to generate an efficient set

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 5

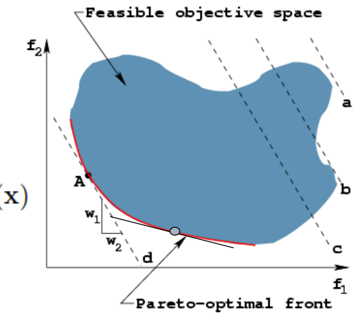
Classical Approach: *Weighted Sum Method*

- Construct a weighted sum of objectives and optimize

Minimize:

$$F_{w_1,w_2}(\mathbf{x}) = w_1 f_1(\mathbf{x}) + w_2 f_2(\mathbf{x})$$

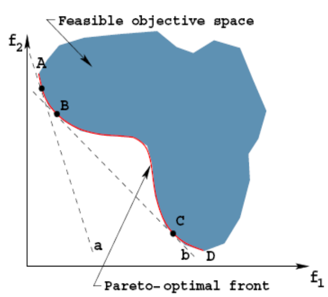
- User supplies weight vector *w*



GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 6

Difficulties Associated to 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 always Pareto-optimal

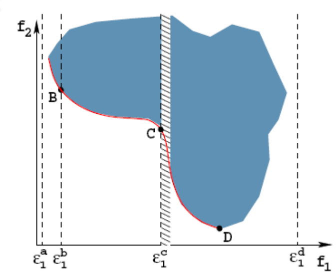


GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 7

ε-Constraint Method

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

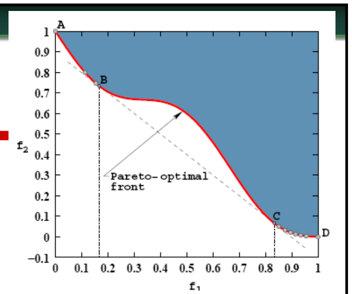
$$\text{Minimize} \quad f_\mu(\mathbf{x}),$$
$$\text{subject to} \quad f_m(\mathbf{x}) \le \epsilon_m, \; m \ne \mu;$$

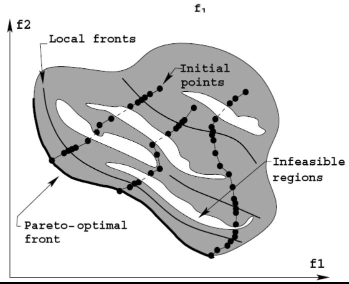


GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 8

Difficulties with Classical Methods

- Need to run a single-objective optimizer many times
- Expect a lot of problem knowledge
- Multi-objective optimization treated as an application of single-objective optimization
- Absence of any parallel search



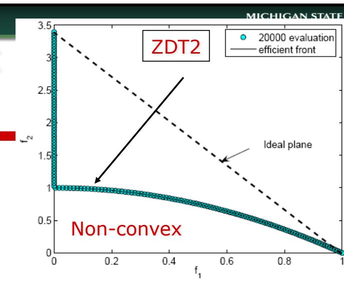


GECCO-2017 Tutc
kdeb@egr.m:

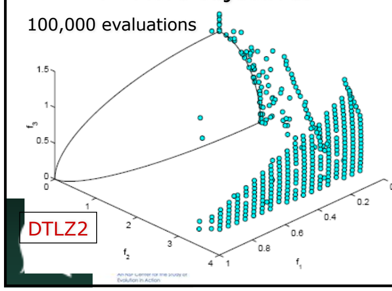
Normal Constraint Method Results

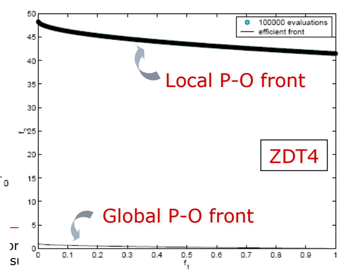
Difficulties:

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



100,000 evaluations





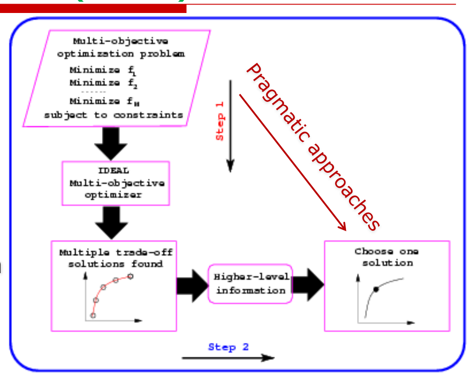
Evolutionary Multi-Criterion Optimization (EMO): *Basics*

Step 1 :

Find a set of Pareto-optimal solutions

Step 2 :

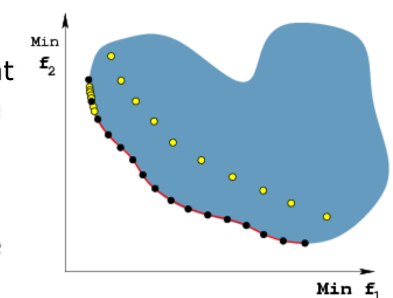
Choose one from the set



GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 11

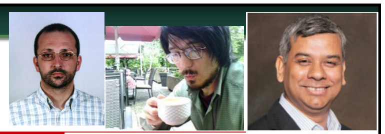
Three Goals in EMO

- Converge to the Pareto-optimal front
- Maintain as diverse a distribution as possible
- Preserve elites for better performance



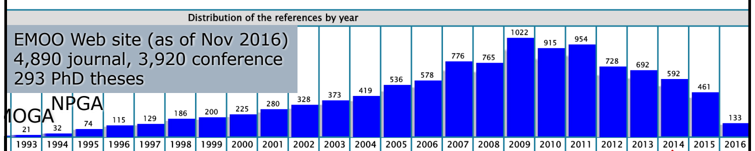
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 12

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)

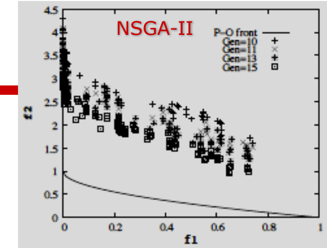


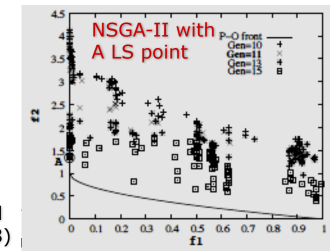
15

Why EMO May Do Well?

- 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

(Sindhya, Deb, and Miettinen, PPSN-08)

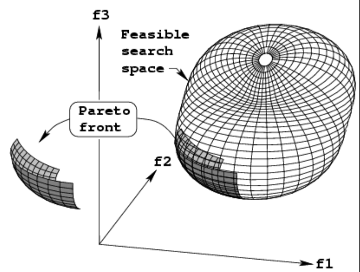




Elitist Non-dominated Sorting Genetic Algorithm (NSGA-II)

(Deb et al., IEEE TEC 2002)

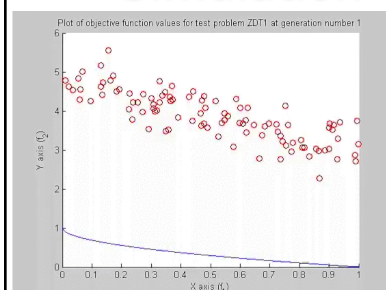
- NSGA-II
  - Modular
  - No additional parameter
  - Fast
- Commercialization:
  - ModeFrontier (Estico)
  - iSIGHT (Engeneous)
  - VisualDoc (Vanderplatts)
  - Heeds (NSGA-II like)



Fast-Breaking Paper in Engineering by ISI Web of Science (Feb'04), Thomson Citation Laureate Award 2006, Current Classic and Most Highly Cited Paper (24,000+ GS citations)

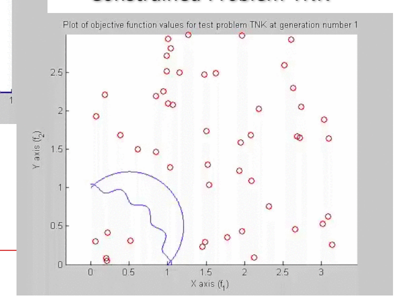
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 15

Simulation of NSGA-II



Unconstrained Problem ZDT1

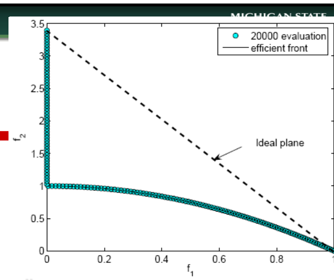
Constrained Problem TNK

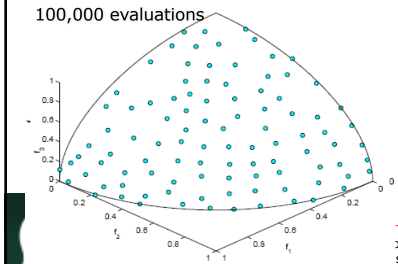


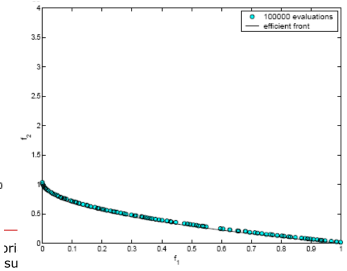
NSGA-II Head-to-head with NCM

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

100,000 evaluations







Achievements During 1993-2008

- Efficient EMO algorithms for 2-3 objectives demonstrated on test problems
  - Advantage of EMO on 2-3 obj is overwhelming compared to 1-obj
  - Commercialization and spread to non-EC areas
- Limited practical applications
- What else to do?
  - Advanced Topics in EMO

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 18

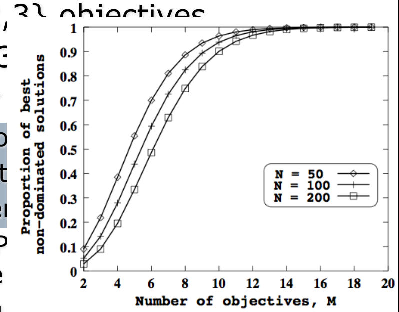
EMO Spread its Wings

- EMO in many-objective optimization
- EMO + MCDM, the whole story!
- Visualization of EMO Solutions
- EMO to aid other problem solving
  - Multiobjectivization (Corn/Knowles)
- EMO for handling practicalities
- More efficient EMO Algorithms
- Theoretical EMO

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 19

Evolutionary Many-Objective Optimization (EMO)

- Multi-objective: {2,3} objectives
- Many-objective: >3
- EMO difficulties for
  1. Large fraction of po
  2. Maintaining diversit
  3. Recombination ope
  4. Representation of P exponentially more
  5. Performance measures difficult to compute
  6. Visualization is difficult



GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 20

MICHIGAN STATE UNIVERSITY

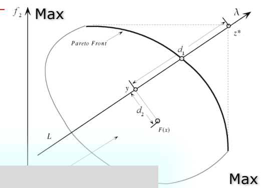
Decomposition Based EMO: MOEA/D

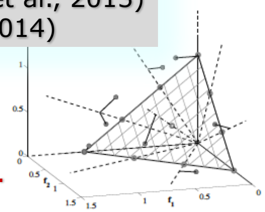
(Zhang and Li, 2008)

- Multiple reference lines
- Random association at first
- Points from neighboring (*T*) lines are mated

Recent Extensions:
- Decomposition based M2M (Liu et al., 2013)
- Stable matching EMO (Li et al, 2014)

- $PBI(\mathbf{x}, \mathbf{w}) = d_1 + \theta d_2$.
- $TCH(\mathbf{x}, \mathbf{w}, \mathbf{z}^*) = \max_{i=1}^{M} w_i |f_i(\mathbf{x}) - z_i^*|$
- No explicit selection oper.





Original study to 2-3 obj.
Param: *T*, *θ*

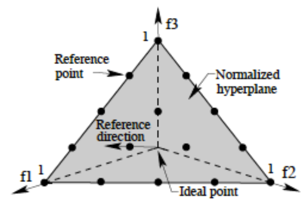
BEACON

MICHIGAN STATE UNIVERSITY

NSGA-III: (Deb and Jain, 2014)

Step 0: Supply of Reference Points

- If no preference, use Das and Dennis's approach
- # pts.: $H = \binom{M+p-1}{p}$
- Else, supply a preferred set of reference points
- Points are given on the normalized hyper-plane
- Any other structured set of points can also be supplied



p=4, M=3, H=15

BEACON

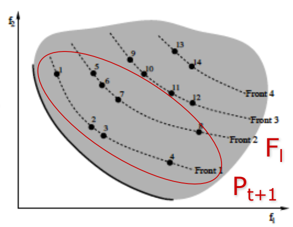
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 22

MICHIGAN STATE UNIVERSITY

NSGA-III:

Step 1: Identification of Non-dominated Fronts

- Parent and offspring population combined to R
- Non-dominated sorting of R
- Collect and save fronts 1 to l to S, delete other fronts
- If |S|=N, P_{t+1} = S, go to next iteration
- Else, $P_{t+1} = \{F_1,\ldots,F_{l-1}\}$, remaining pts chosen from F_l



N=7

BEACON

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 23

MICHIGAN STATE UNIVERSITY

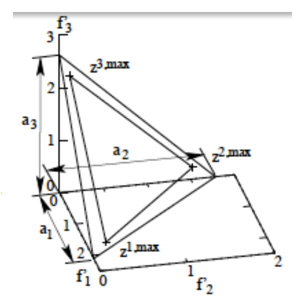
NSGA-III:

Step 2: Normalization of Population Members

- Identify ideal point of S
- Translate all members of S to make new ideal pt. as origin
- Identify extreme pts.:

$$ASF(\mathbf{x}, \mathbf{w}) = \max_{j=1}^{M} f_i'(\mathbf{x})/w_i, \quad \text{for } \mathbf{x} \in S_t$$

- Update with previous extreme pts.
- Normalize using ideal and extreme points obtained

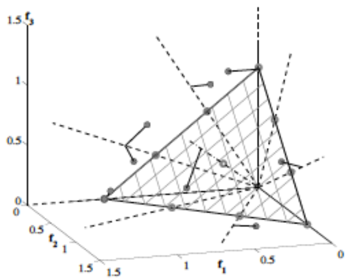


BEACON

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 24

NSGA-III:
Step 3: Association of Population Members

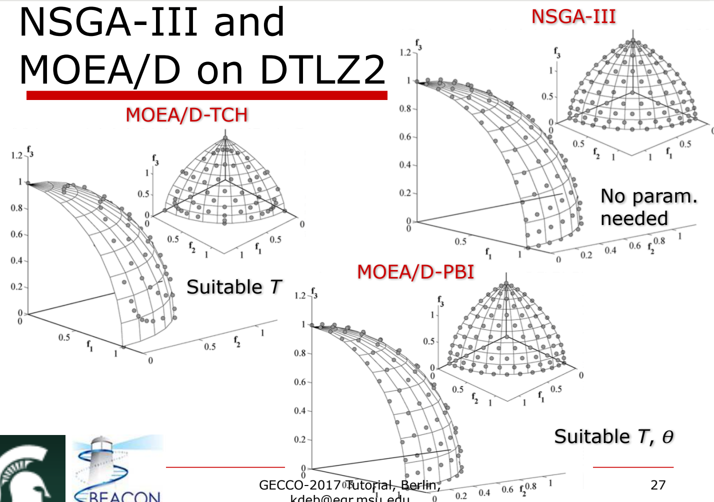
- Reference line for each ref. pt. is found
- The ref. pt. having shortest perpendicular dist. from a S member is associated
- A reference point may have zero, one, or more than one S members associated



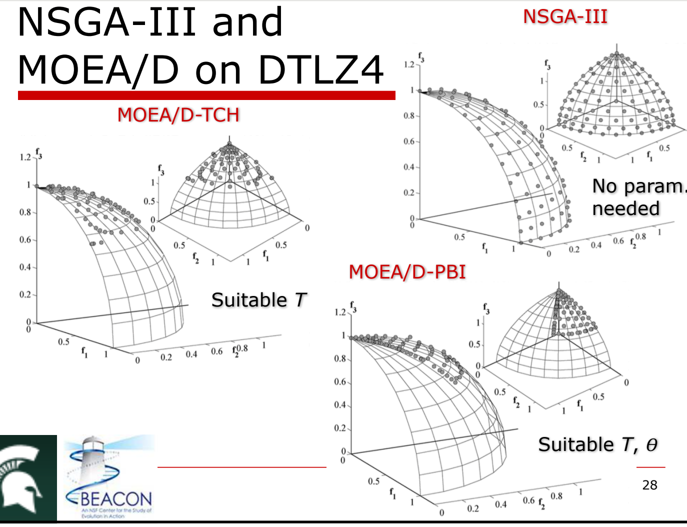
NSGA-III:
Genetic Operators

- Selection Operator is not used.
- Recombination and mutation operators as before
- Use large value of distribution index of SBX
  - To create meaningful offspring
- No additional parameter needed, like in NSGA-II, unlike in MOEA/D

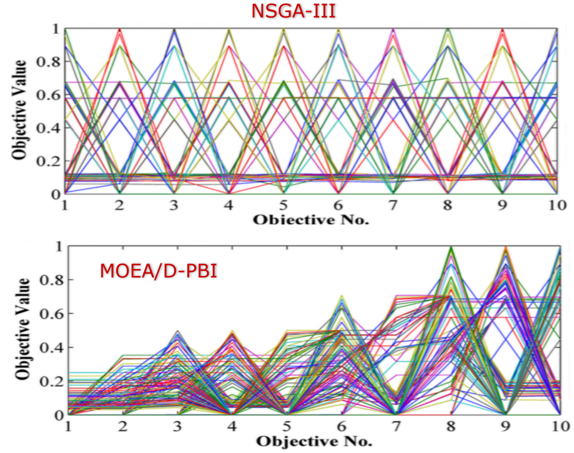
NSGA-III and MOEA/D on DTLZ2



NSGA-III and MOEA/D on DTLZ4



10-Objective DTLZ4



Classical ASF Method and NSGA-III

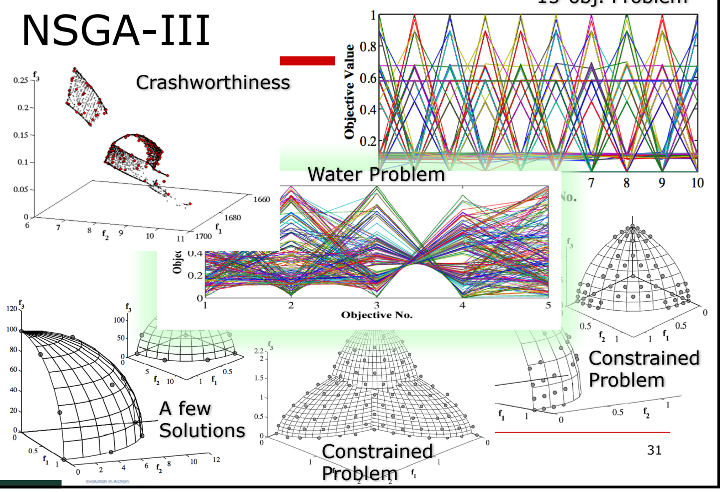
- Classical method works, but is expensive

$$\text{Minimize}_{\mathbf{x}} \quad d_1 + \theta d_2 = \mathbf{w}^T\mathbf{f}(\mathbf{x}) + \theta \left(\|\mathbf{f}(\mathbf{x}) - \mathbf{w}^T\mathbf{f}(\mathbf{x})\right) \mathbf{w}\|$$

Prob.	FE	NSGA-III IGD	NSGA-III GD	Generative Method IGD	Generative Method GD
DTLZ1	36,400	**4.880×10^{-4}**	**4.880×10^{-4}**	6.400×10^{-2}	1.702×10^{1}
		1.308×10^{-3}	**6.526×10^{-4}**	8.080×10^{-2}	1.808×10^{1}
		4.880×10^{-3}	**7.450×10^{-4}**	1.083×10^{-1}	1.848×10^{1}
DTLZ2	22,750	1.262×10^{-3}	1.264×10^{-3}	**1.113×10^{-3}**	9.678×10^{-5}
		1.357×10^{-3}	1.270×10^{-3}	6.597×10^{-3}	1.019×10^{-4}
		2.114×10^{-3}	1.274×10^{-3}	9.551×10^{-3}	1.082×10^{-4}

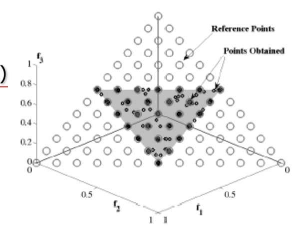
Some Results of NSGA-III

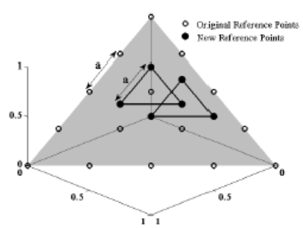
(IEEE TEC 2014)



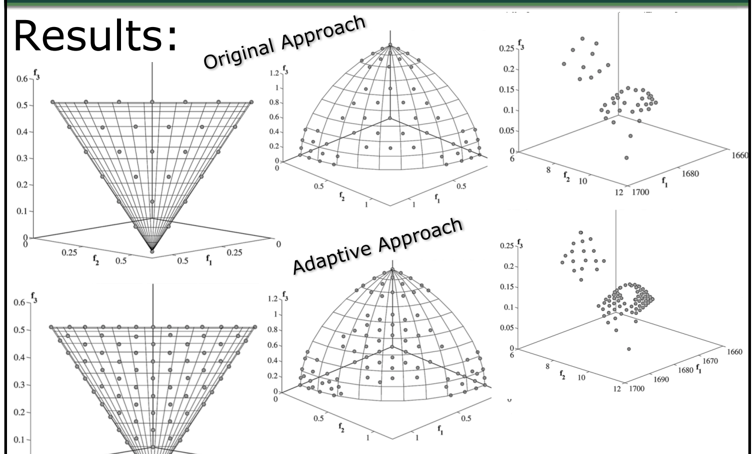
Adaptive NSGA-III (Jain and Deb, 2014)

- Every reference point may not have a associated point on the PO front
- Need to increase # of ref. pts. to more PO points
  - Not efficient
- Add/Delete with a strategy





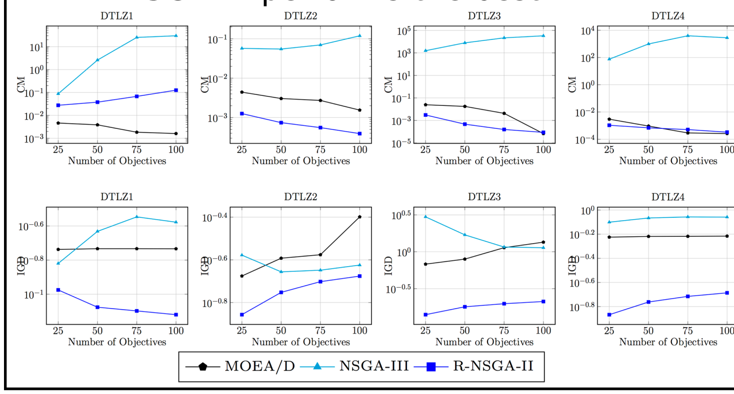
Results:



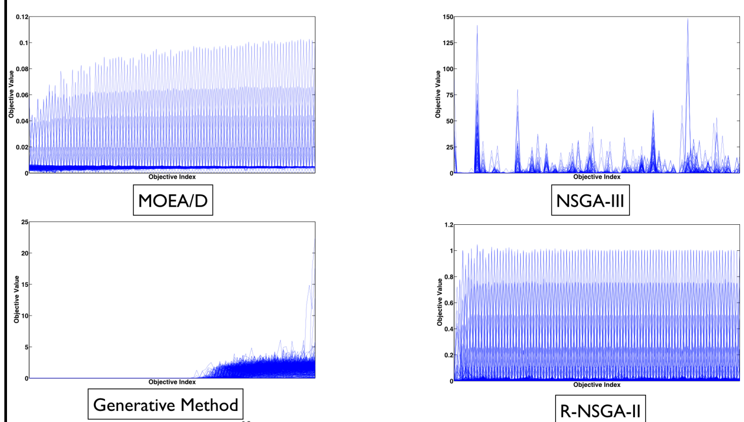
ECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 33

Evolutionary Massive Optimization (EMO)

- R-NSGA-II performs the best



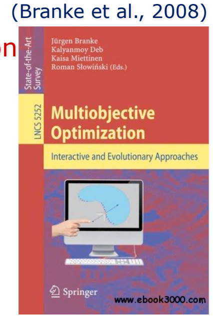
Parallel Coordinate Plots for 100-obj DTLZ1 Problem



EMO and Decision-Making

- Finding a P-O set (using EMO) is half the story
- How to choose one preferred solution (MCDM)
- A-priori approach:
  - First MCDM, then EMO
- A-posteriori approach:
  - First EMO, then MCDM
- Interactive approach:
  - MCDM during EMO

(Branke et al., 2008)

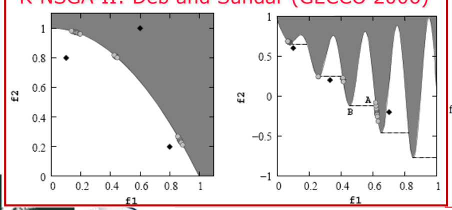


GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 36

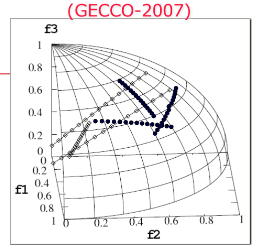
Making Decisions: *A Priori or A Posteriori*

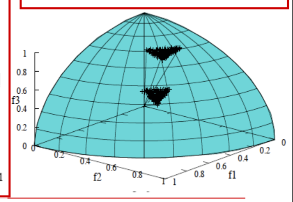
- Ranking based on closeness to each reference point or a reference direction

R-NSGA-II: Deb and Sundar (GECCO 2006)



RD-NSGA-II: Deb and Kumar (GECCO-2007)



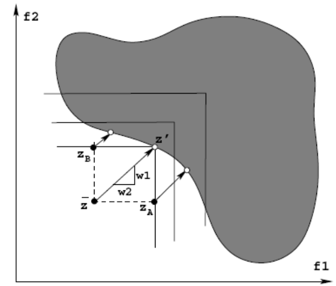


'Light Beam' Approach in CEC-07

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu

Reference Point Based MCDM

- Wierzbicki, 1980
- Start with a reference point (RP)
- A P-O solution closer to RP using L_∞
- Construct other RPs from P-O solution
- Obtain more P-O solutions
- Continue till satisfied



A way to find a preferred *region*, not a preferred point

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 38

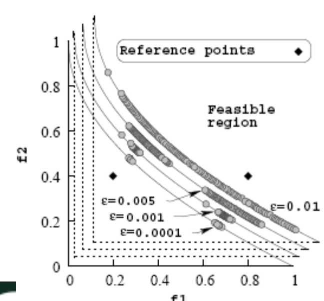
Reference Point Based NSGA-II

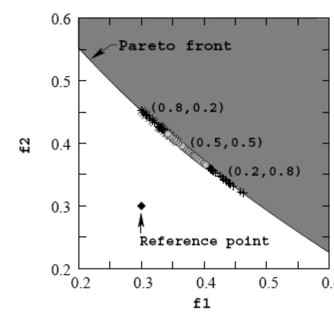
- Modify crowding distance operator
- Multiple reference points (RPs) given
- For each RP, rank members in increasing Euclidean distance in objective space
- Prefer members with smaller ranks
- To maintain a spread, use a clearing idea
  - Pick a solution randomly
  - Clear all ε-neighboring solutions (high rank)
  - Continue with un-cleared solutions

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 39

Simulation Studies: Effect of ε and weights on ZDT1

- Larger ε, more spread
- Region depends on weights

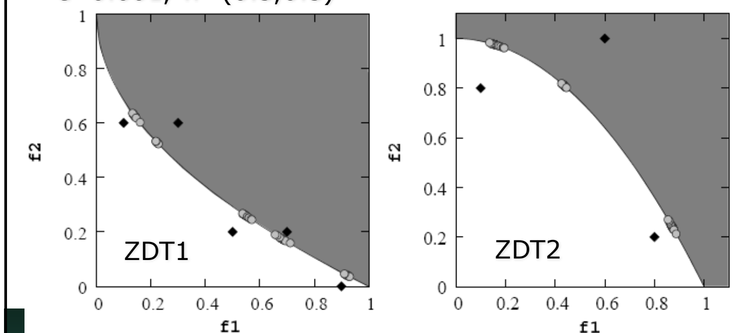




GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 40

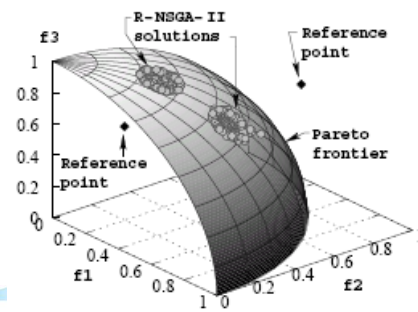
Effect of Convexity (ZDT1 vs. ZDT2)

ε=0.001, w=(0.5,0.5)



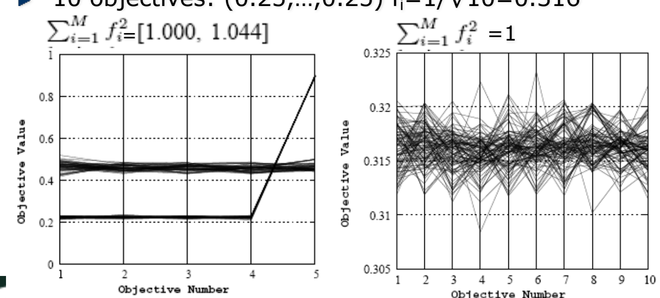
Three-Objective Problems

- DTLZ2 with two RPs (ε=0.01)



Higher Objectives

- Five objectives: (0.5,…,0.5) and (0.2,…0.2,0.8)
- 10 objectives: (0.25,…,0.25) $f_i=1/\sqrt{10}=0.316$



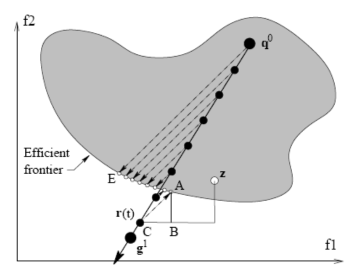
Reference Direction Based MCDM and RD-NSGA-II

(Korhonen and Laakso, 1986 and Deb and Kumar, 2007)

- Choose two points q^0 and g^1
- Get direction $d^k=g^k-q^{k-1}$
- Solve for various t

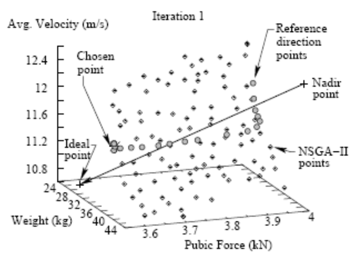
$$\text{Minimize}_{\mathbf{z}} \quad s(\mathbf{z},\mathbf{r},\mathbf{w}) = \max_{i\in I}(z_i - r_i)/w_i,$$
$$\text{subject to} \quad \mathbf{r}(t) = \mathbf{q}^{k-1} + t\mathbf{d}^k.$$

- Choose most preferred solution (q^k) and another point g^{k+1}
- RD-NSGA-II developed



Example: Car Side Impact Problem

- Three objectives, 10 constraints, seven variables
- Utility: Min. avg. force at abdomen and pubic area
- q^0=nadir point=(42.69,4.00,12.44)
- g^1=ideal point=(24.37,3.59,10.61)
- 25 points -> q^1=(35.95,3.56,11.53)

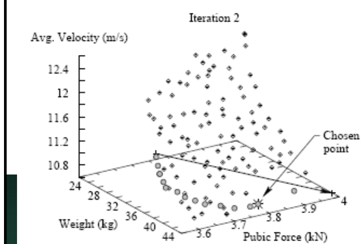


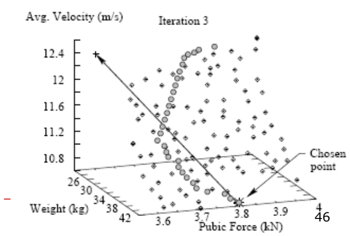
- Avg. vel (f_3) smaller
- g^2=(42.69,4.00,10.61)
- RD: q^2 to g^2
- 15 points

45

Second and Third Iterations

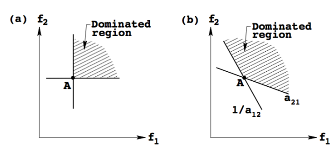
- q^2=(40.98,3.81,10.61), f_3 is reduced
- Reduce f_1 and f_2: g^3=(24.37,3.59,12.44)
- q^3=(40.92,3.81,10.61)
- Consider q^2 and q^3 close and terminate
- Declare x-vector





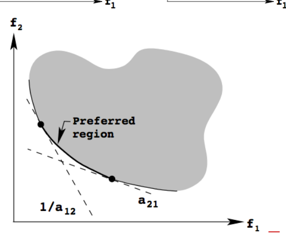
Cone Domination

- Using a DM's preference (not a solution but a region)
- Cone/guided domination principle: Biased niching approach



$$\Omega_i(\mathbf{f}(\mathbf{x})) = f_i(\mathbf{x}) + \sum_{j=1, j\neq i}^{M} a_{ij} f_j(\mathbf{x}), \quad i = 1, 2, \ldots, M.$$

$$\Omega_1(f_1, f_2) = f_1 + a_{12} f_2,$$
$$\Omega_2(f_1, f_2) = a_{21} f_1 + f_2.$$

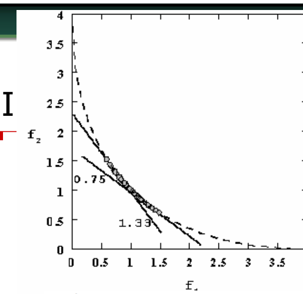


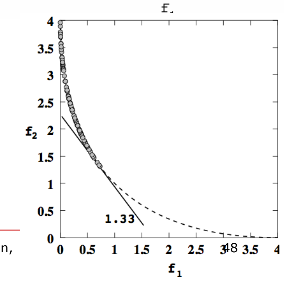
Results Using Cone-Dominated NSGA-II

Two Approaches:

1. Use cone domination in NSGA-II
2. Use a modified set of objectives

Can be extended to more than two objectives



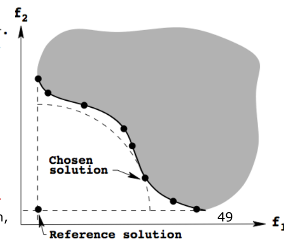


A-Posteriori Approaches

- Any of the a priori approaches can be applied after a set of non-dominated points are found
- Compromise Programming:

l_p-metric: $$d(\mathbf{f},\mathbf{z}) = \left(\sum_{m=1}^{M} |f_m(\mathbf{x}) - z_m|^p\right)^{1/p},$$

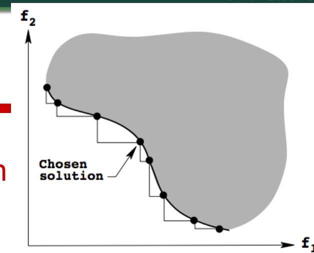
Tchebycheff metric: $$d(\mathbf{f},\mathbf{z}) = \max_{m=1}^{M} \frac{|f_m(\mathbf{x}) - z_m|}{\max_{\mathbf{x}\in\mathcal{S}} f_m(\mathbf{x}) - z_m}.$$



- Need a reference point
  - Usually the ideal point
- Quite common in practice

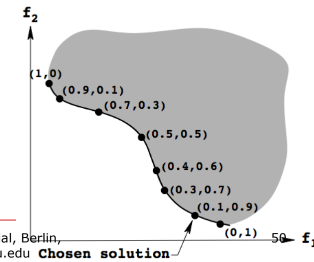
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 49

A Posteriori Approaches (Cont.)



- Marginal Rate of Substitution Approach
- Pseudo-weight Approach:

$$w_i = \frac{(f_i^{max} - f_i(\mathbf{x}))\,/\,(f_i^{max} - f_i^{min})}{\sum_{m=1}^{M} (f_m^{max} - f_m(\mathbf{x}))\,/\,(f_m^{max} - f_m^{min})}$$



  - Choose a solution close to desired weight combination

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 50

Progressively Interactive EMO (PI-EMO)

- Deb, Sinha, Korhonen and Wallenius, 2010 (IEEE TEC)
- Preference information during an EMO run
  - Ask DM after a few generations
  - Modify search thereafter
  - Continue till convergence
- Future of preference based EMO
- Branke et al. (2009) and others

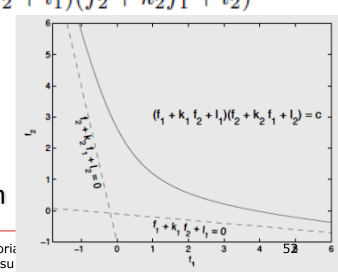
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 51

Utility Function

- Choose k well-distributed non-dominated points
- Ask DM for pair-wise information
- Form a utility function:

$$V(f_1, f_2) = (f_1 + k_1 f_2 + l_1)(f_2 + k_2 f_1 + l_2)$$

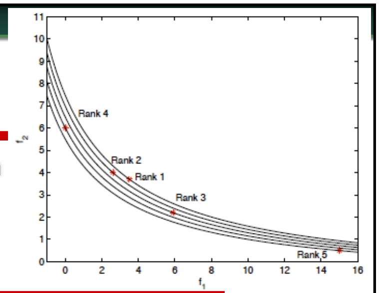
- Can be generalized to any number of objectives
- Parameters k_i, l_i are to determined by solving an optimization problem



GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 52

Forming Utility Function

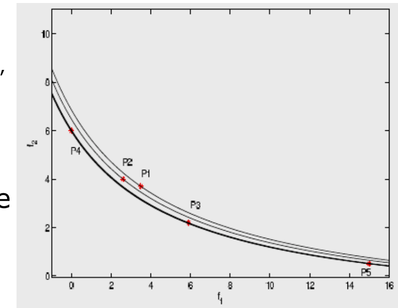
- Solve following problem with k_i, l_i, and ε as variables:



$$\begin{array}{ll}\text{Maximize} & \epsilon,\\ \text{subject to} & V \text{ is non-negative at every point } P_i,\\ & V \text{ is strictly increasing at every point } P_i,\\ & V(P_i) - V(P_j) \geq \epsilon, \quad \text{for all } (i,j) \text{ pairs}\\ & \quad \text{satisfying } P_i \succ P_j,\\ & |V(P_i) - V(P_j)| \leq \delta_V, \quad \text{for all } (i,j) \text{ pairs}\\ & \quad \text{satisfying } P_i \equiv P_j.\end{array}$$

Another Example

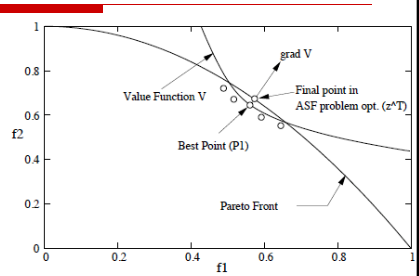
- P_1 better than P_2 better than (P_3, P_4, P_5)
- P_3, P_4 and P_5 are incomparable
- Absolute difference in V values are kept within 0.1ε



$$V(f_1, f_2) = (f_1 + 5.9355)(f_2 + 1.6613).$$

Termination Check

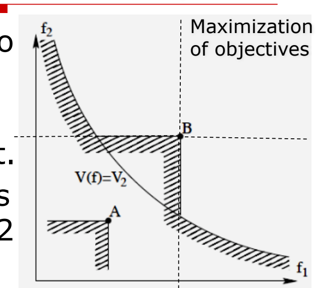
- Best point z=P_1 is identified
- Extent of improvement is determined
- Along normal of V
- If the extent is less than a threshold, quit



$$\begin{array}{ll}\text{Maximize} & \left(\min\limits_{i=1}^{M} \frac{f_i(\mathbf{x}) - z_i^b}{\frac{\partial V}{\partial f_i}}\right) + \rho \sum_{j=1}^{M} \frac{f_j(\mathbf{x}) - z_j^b}{\frac{\partial V}{\partial f_j}}.\\ \text{subject to} & \mathbf{x} \in \mathcal{S}.\end{array}$$

Revised Domination Principle

- Utility function is used to revise domination principle
- V2: V for second best pt.
- If two points have V less than V2 or more than V2
  - Usual domination rule
- Else the one having >V2 dominates other

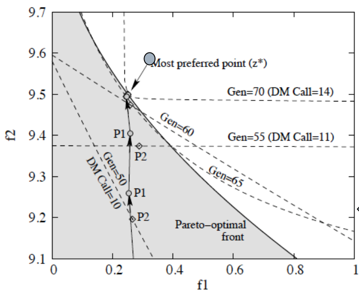


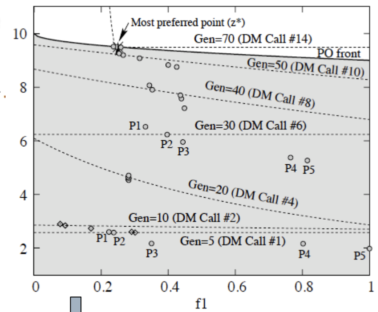
MICHIGAN STATE UNIVERSITY

PI-EMO Results

- ZDT1 with a preference function:

$$V(f_1, f_2) = \frac{1}{(f_1 - 0.35)^2 + (f_2 - 9.6)^2}.$$





MICHIGAN STATE UNIVERSITY

Visualization Methods in EMO

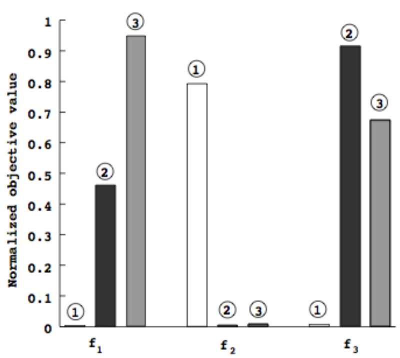
- Bar Charts
- Scatter Plots
- Multi-way Dot Plots
- Table Lens Plots
- Heat Maps
- Parallel Coordinate Plots (PCP)
- Level Diagrams and Hyper Radial Visualization (HRV)

MICHIGAN STATE UNIVERSITY

Bar Chart Method

- Cannot get a good idea of trade-off
- Not very popular

(Jarvenpaa, 1989)



MICHIGAN STATE UNIVERSITY

Scatter Plot

- M choose 2 pair-wise combinations

(Chambers and Kleiner, 1982)

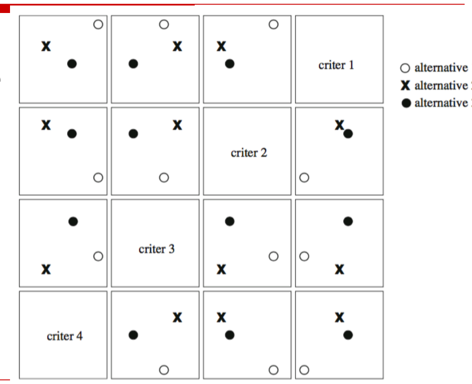
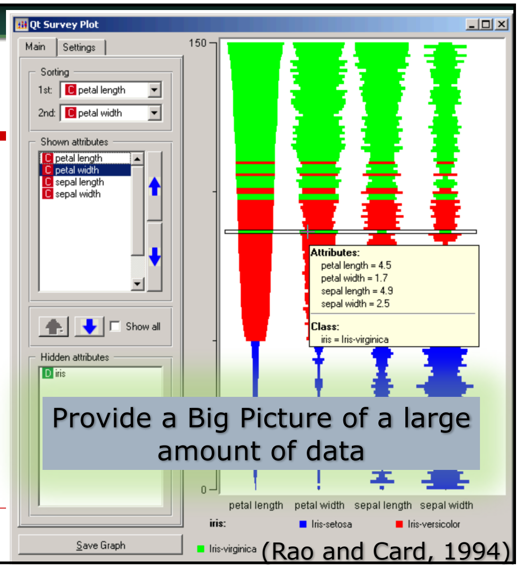


Table Lens Plot

- Each row is a data point with M obj.
- Color indicates its class
- Arranged in sorted order of one obj.

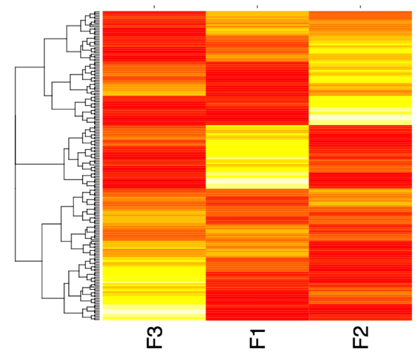


(Rao and Card, 1994)

Heat Maps

- Objective values in different colors
- Difficult to comprehend trade-off

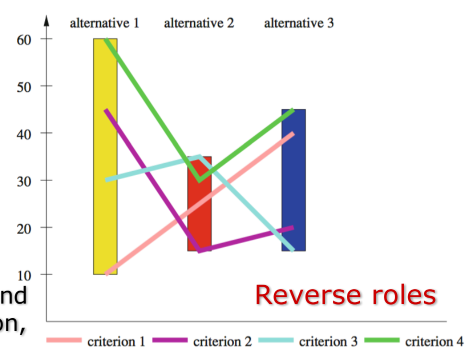
(Pryke et al., 2007)



Parallel Coordinate Plot (PCP)

- Value path method
- Most popular
- 3D PCPs
- Permutation of objectives
- Many variations exist

(Mezzich and Worthington, 1978)

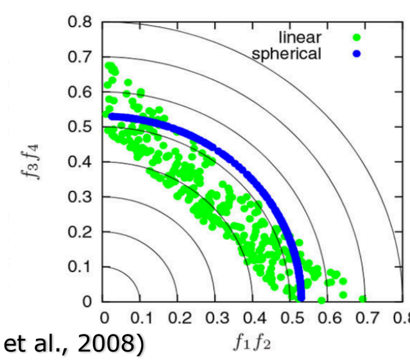


Level Diagrams

Hyper Radial Visualization: Pair-wise Level Diagram (Chiu and Bloebaum, 2010)

- Distance from the ideal point
  - Two PO sets: Linear and spherical
- Traditionally, smallest distance solution is chosen

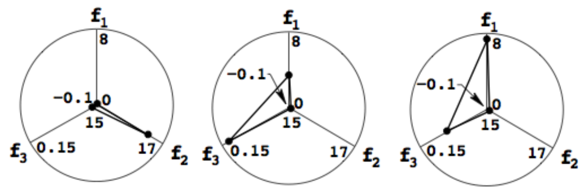
(Blasco et al., 2008)



Star-Coordinate Method

- Used in some softwares
- Still cannot get a good idea of trade-off

(Manas, 1982)

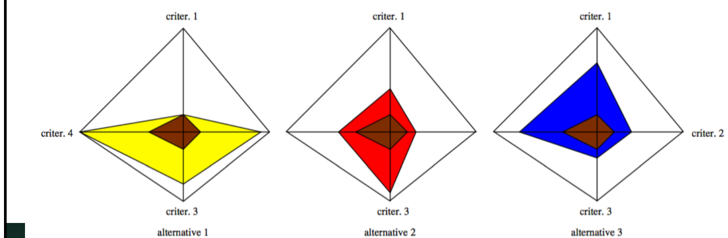


GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 65

Spider-Web Chart

- Value of each objective along with their minimum and maximum

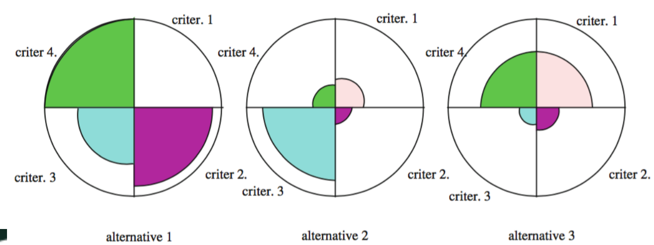
(Kasanen et al., 1991)



GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 66

Petal Diagram (Tan and Fraser, 1998)

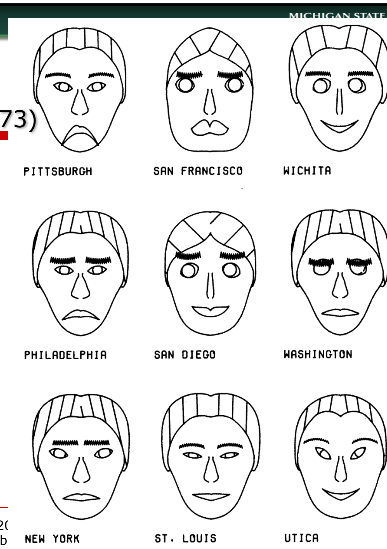
- Area in each sector indicates the value of respective objective



GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 67

Chernoff Faces (Chernoff, 1973)

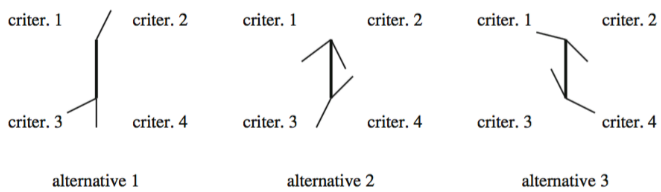
- 16 features represent 16 objectives
- Pittsburgh: High pollution (mouth)
- Wash: High prop of non-white pop. With high income
- San Fransisco: Low density of pop (eye separated)



GECCO-2(kdeb

Stick Figure

- Angles and length of sticks represent objectives

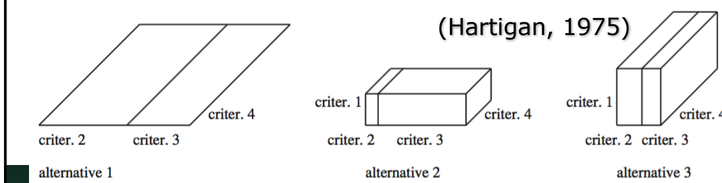


(Pickett and Grinstein, 1988)

Boxes

- Easily extendable to any number of objectives, but may not portray trade-off
- Order of criteria is important

(Hartigan, 1975)

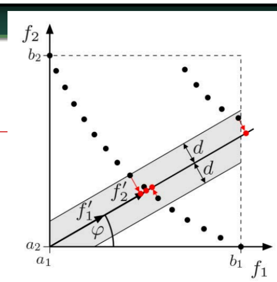


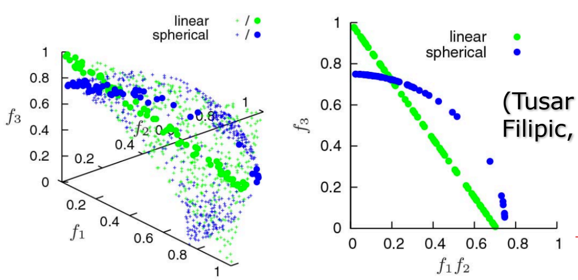
Prosection Method

(Tusar and Filipic, 2014)

- Make a slice along M-dim PO front and project on a hyperplane

$m\mathrm{D}(\boldsymbol{a}, f_i f_j, \varphi, d)$



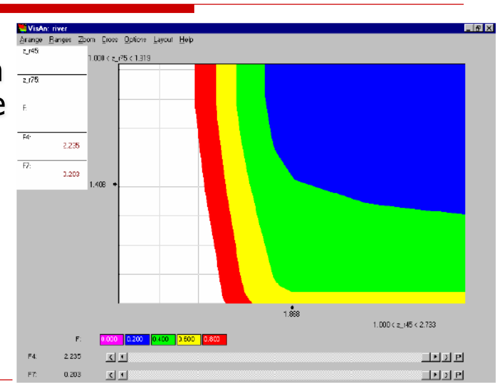


(Tusar and Filipic, 2014)

Decision Maps

- Three objective in prominence
- Others as scroll bars
- Requires pre-processing

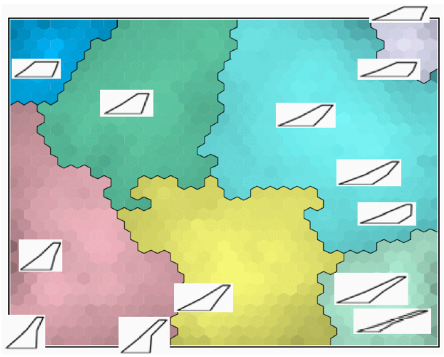
(Lotov, 2004)



Self-Organizing Maps

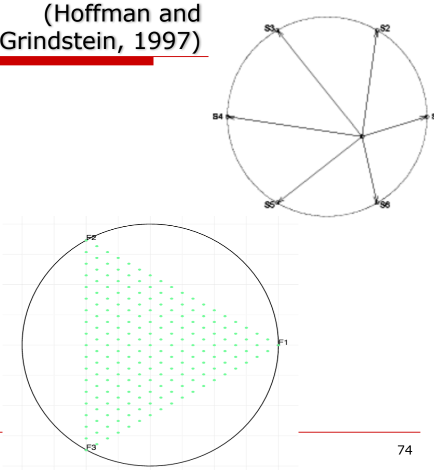
- Design space is divided based on similarity in variables

(Kohonen, 2001)



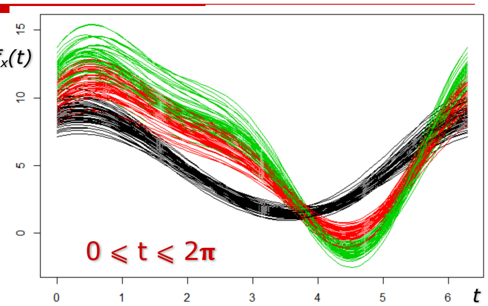
Radial Coordinate Visualization (RadViz) (Hoffman and Grindstein, 1997)

- A puck is in equilibrium from M points on the circle with stiffness prop. to obj values



Andrews Plot (Andrews, 1972)

- Generalized PCP
- Grand Tour to find outliers



$$f_x(t) = \frac{x_1}{\sqrt{2}} + x_2 \sin(t) + x_3 \cos(t) + x_4 \sin(2t) + x_5 \cos(2t) + \ldots$$

Multiobjectivization:

Solving Problems Using MO Principle

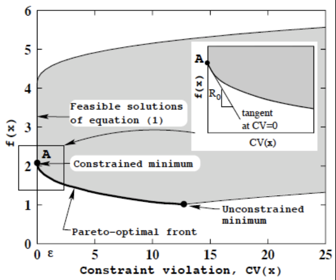
- Constrained handling
  - Constraint violations as additional objectives
- Multimodal problems
- Bloating in Genetic Programming (Blueler et al, 2001)
- Diversity preservation in EAs (Jensen, 2003, Abbas and Deb, 2003)
- Fuzzy clustering methods
- Goal programming

(Knowles, Corne & Deb, 2008)

1-obj. Constraint Handling

minimize
subject to
minimize
minimize

- Pose as a two-obj problem
- Compare penalty based approach with weighted-sum approach



Penalty Function Approach:

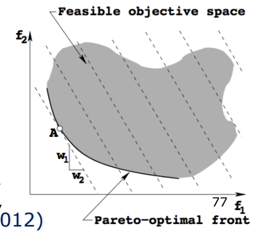
$$P(\mathbf{x}, R) = f(\mathbf{x}) + R \cdot \mathrm{CV}(\mathbf{x}),$$
$$= f_2(\mathbf{x}) + R f_1(\mathbf{x}),$$

Bi-objective Approach:

$$\text{minimize} \quad F_{w_1,w_2}(\mathbf{x}) = w_1 f_1(\mathbf{x}) + w_2 f_2(\mathbf{x}).$$

Equating the two:

$$w_1 = R \text{ and } w_2 = 1.$$



(Deb and Datta, EO, 2012)

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 77

EMO + Classical Penalty Based Approach

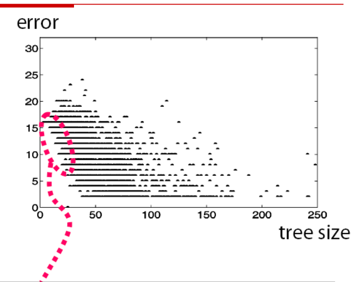
- EMO to get R
- Classical penalized approach to find a local solution
- Improvements of one or two-order in standard test problems

Prob.	Zavala, Aguirre & Diharce [52]			Takahama & Sakai [45]			Brest [6]			Proposed Hybrid Approach		
	Best	Median	Worst	Best	Median	Worst	Best	Median	Worst	Best	Median	Worst
g01	80,776	90,343	96,669	18,594	19,502	19,917	51,685	55,211	57,151	**2,630**	**3,722**	**4,857**
g02	87,419	93,359	99,654	1,08,303	114347	1,29,255	1,75,090	2,26,789	2,53,197	**26,156**	**50,048**	**63,536**
g04	93,147	1,03,308	1,109,15	12,771	13,719	14,466	56,730	62,506	67,383	**1,210**	**1,449**	**2,295**
g06	95,944	1,09,795	1,30,293	5,037	5,733	**6,243**	31,410	34,586	37,033	**1,514**	**4,149**	11,735
g07	1,14,709	1,38,767	2,08,751	60,873	67,946	75,569	1,84,927	1,97,901	2,21,866	**15,645**	**30,409**	**64,732**
g08	2,270	4,282	5,433	**621**	**881**	**1,173**	1,905	4,044	4,777	822	1,226	2,008
g09	94,593	1,03,857	1,19,718	19,234	21,080	21,987	79,296	89,372	98,062	**2,732**	**4,850**	**5,864**
g10	1,09,243	1,35,735	1,93,426	87,848	92,807	**1,07,794**	2,03,851	2,20,676	2,64,575	**7,905**	**49,102**	1,80,446
g12	482	6,158	9,928	2,901	4,269	5,620	**364**	6,899	10,424	496	**504**	**504**
g18	97,157	1,07,690	1,24,217	46,856	57,910	60,108	1,39,131	1,69,638	1,91,345	**4,493**	**7,267**	**10,219**
g24	11,081	18,278	6,33,378	1,959	2,451	**2,739**	9,359	12,844	14,827	**1,092**	**1,716**	2,890

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 78

Reducing Bloating in GP

- Bleuler et al., (2001)
- Find small-sized programs with small error
- Minimization of Size of Program as second objective



Keep and optimize small trees (potential building blocks)

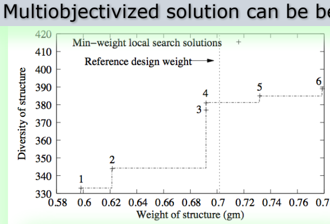
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 79

EMO to Explore Design Space

- Use of additional objectives for a reason
- Bi-objective optimization

1-obj Weight Minimization

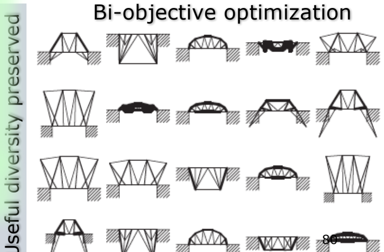
Multiobjectivized solution can be better



(Sharma, Deb, Kishore, 2013)

Useful diversity preserved

Bi-objective optimization

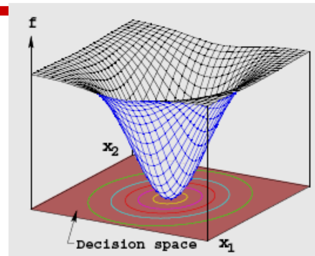


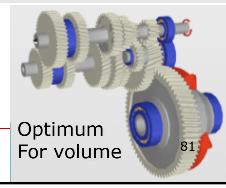
(Tamara and Thiele, 2012)

80

Innovization: Learning from Trade-off Solutions

- Often, one optimum x*
- x* minimizes f(x) subject to satisfaction of some constraints
- *Sensitivity analysis* provides neighborhood information
- Not much can be learned from one solution



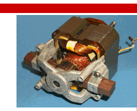


Optimum For volume

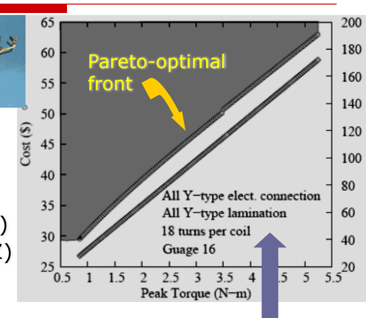
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 81

Case Studies of Manual *Innovization*: *Brushless DC Permanent Magnet Motor Design for Cost and Peak Torque*

- Five variables (all discrete), three constraints
- Non-convex, disconnected P-O fronts



- *Innovization*:
  - Connection: Y (betn. Y & Δ)
  - Lamination Type: Y (X, Y, Z)
  - 1 out of 16 wire guages
  - 18 turns per coil (10,80)
  - More peak torque by adding linearly more laminations

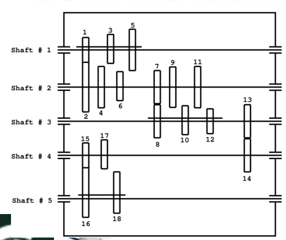


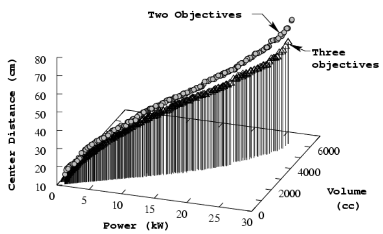
Design Innovation

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 82

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)

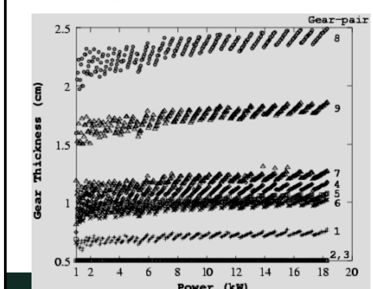


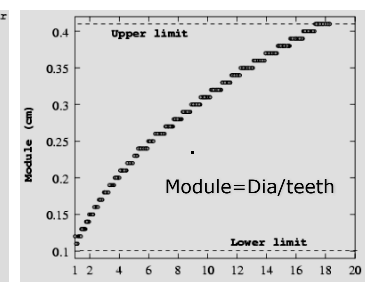


GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 83

Innovized Principles

- Module (discrete) varies proportional to square-root of power
- Keep other 28 variables more or less the same

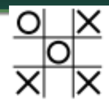




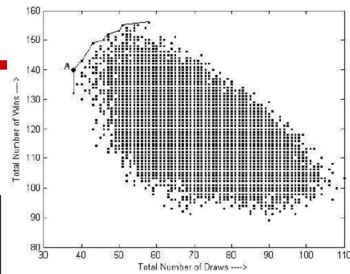
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 84

(Bhatt, Varshney, and Deb, 2008)

Tic-Tac-Toe Game



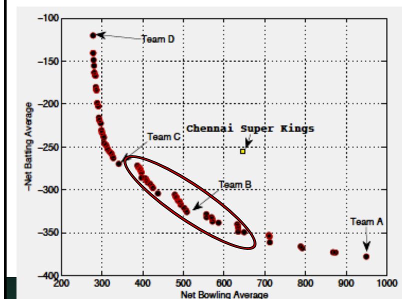
- Trade-off solutions are found and analyzed
- Following principles are discovered:
  - If opponent is one short of winning, block it
  - If center is empty, occupy it
  - If center is filled, occupy corner and edge-center, in this order

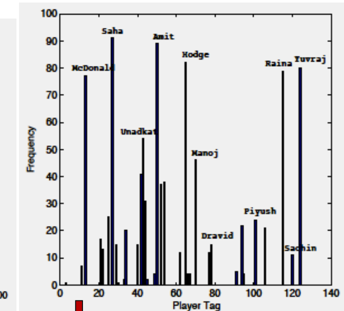


- All 72,657 solutions are split in #draws and #wins

Player Selection in the Game of Twenty20 Cricket

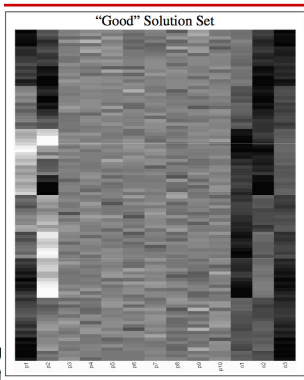
- Compute frequency of players and choose from them

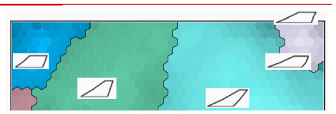


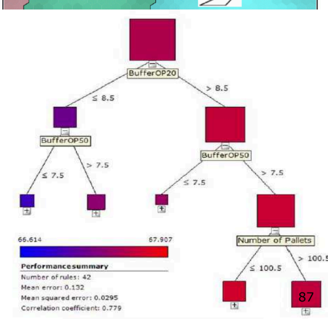


29 of 129 players Important

Heatmap, SOM and Decision Tree From Pareto-Optimal Set

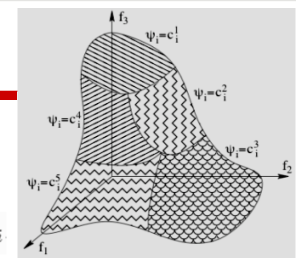






Automated Innovization

- Rules of type:
  $\psi_i(\mathbf{x}, \mathbf{f}(\mathbf{x}), \mathbf{g}(\mathbf{x})) = c_i.$
- Currently, limited to
  $\psi_i(\boldsymbol{\phi}(\mathbf{x})) \equiv \Pi_{j=1}^{N} \phi_j(\mathbf{x})^{b_{ij}} = c_i$
- Solve (m data points):

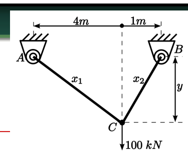


(Bandaru and Deb, 2010, EO)

$$\begin{aligned} \text{Minimize} \quad & \left(number\ of\ clusters + unclustered\ points + \sum_{clusters} c_v \right), \\ \text{Subject to} \quad & 1 \le d_i \le m, \\ & unclustered\ points \le 0, \\ & -1 \le b_{ij} \le 1\ \forall j, \\ & |b_{ij}| \ge 0.1\ \forall j, \\ & d_i \text{ is an integer and } b_{ij}\text{'s are real.} \end{aligned}$$

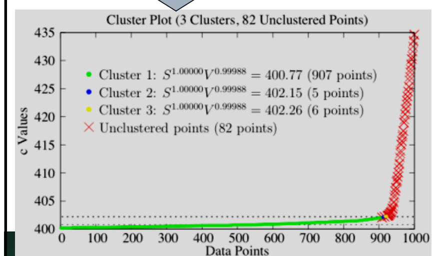
Coeff. of variance

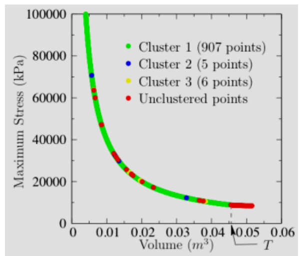
Truss Design



- Obtained Rules (independent applications):

$$SV = 400.77, \quad \frac{x_1}{V} = 0.1105, \quad \frac{x_2}{V} = 0.2236, \quad \frac{x_2}{x_1} = 1.9838$$





Multiple Rules Simultaneously

(Bandaru and Deb, 2010, Engg. Optimization)

- Use of a *niching* operator to find multiple rules
- Row-echelon operation to get a minimal set

i	$a^*_{i1}b^*_{i1}$	$a^*_{i2}b^*_{i2}$	$a^*_{i3}b^*_{i3}$	$a^*_{i4}b^*_{i4}$	$a^*_{i5}b^*_{i5}$	d^*_i	S_i
DR1	1.0000000	0.0000000	0.0000000	-1.0006158	0.0000000	520	88.2%
DR2	0.0000000	1.0000000	0.0000000	1.0005781	0.0000000	508	80.8%
DR3	0.0000000	0.0000000	1.0000000	-1.0009661	0.0000000	507	86.8%
DR4	0.0000000	0.0000000	0.0000000	0.0000000	1.0000000	511	87.2%

10	10101	0.8388214	0.0000000	[illegible]1667	0.0000000	1.0000000	869
11	10111	0.8813441	[illegible]	[illegible]007	-0.3013669	1.0000000	869
18	10110	1.0000000	0.0000000	-0.7353538	-0.2647701	0.0000000	869
19	11010	1.0000000	0.6702390	0.0000000	-0.3300355	0.0000000	869
20	11100	1.0000000	0.7911332	-0.2102126	0.0000000	0.0000000	860

Truss Design

DR1: $\frac{V}{x_1} = c_{\text{DR1}}$, DR2: $Sx_2 = c_{\text{DR2}}$, DR3: $\frac{x_1}{x_2} = c_{\text{DR3}}$, DR4: $y = c_{\text{DR4}}$

Dimensionally Aware Genetic Programming (GP) (Bandaru and Deb, 2013 EMO)

- Parse tree representation
- Physically meaningful DPs through constraints
- Subtree crossover
- Generic rules possible

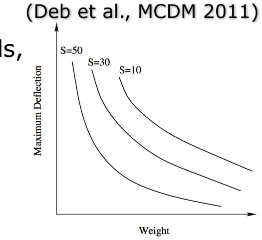
Welded Beam Design Problem

Notation	Design Principle (DP) $\psi(\mathbf{x}) = constant$	Significance
DP1	$(D+t) = constant$	95.20%
DP2	$t = constant$	95.60%
DP3	$D \times b = constant$	95.00%
DP4	$D \times b \times t = constant$	95.60%
DP5	$\sigma \times b = constant$	94.80%
DP6	$\sigma \times b \times t = constant$	95.60%
DP7	$D/\sigma = constant$	95.60%
DP8	$D/(\sigma \times t) = constant$	95.60%

Higher-Level *Innovization*

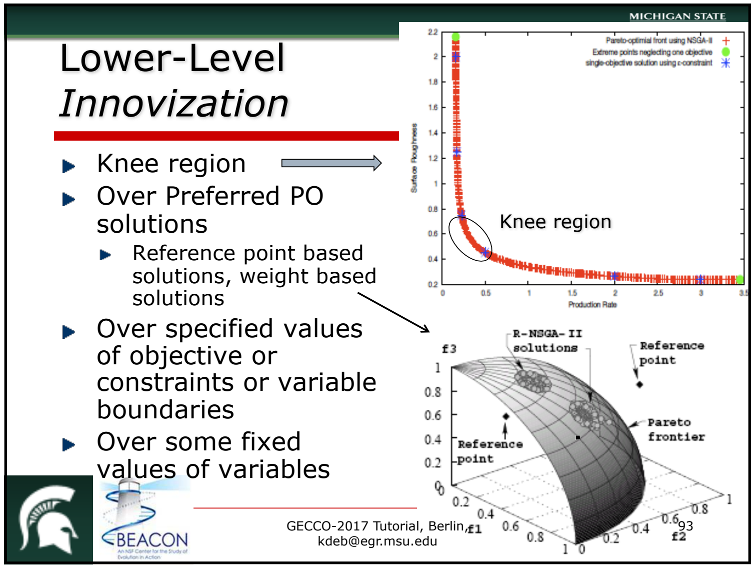
(Deb et al., MCDM 2011)

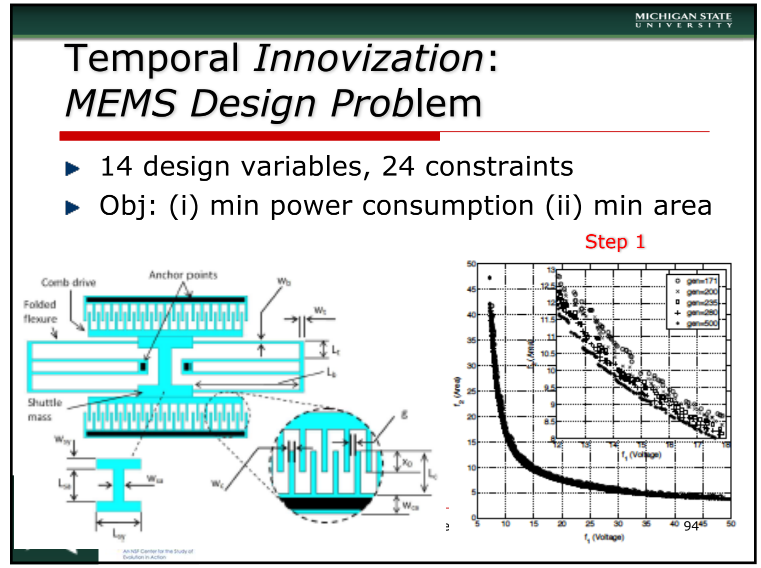
- Over different parameter settings
  - Material properties, load, bounds, resources
- Over different variable sizes and types
  - Continuous to discrete
- Over multi-modal solutions
- Over different constraint combinations
- More possibilities

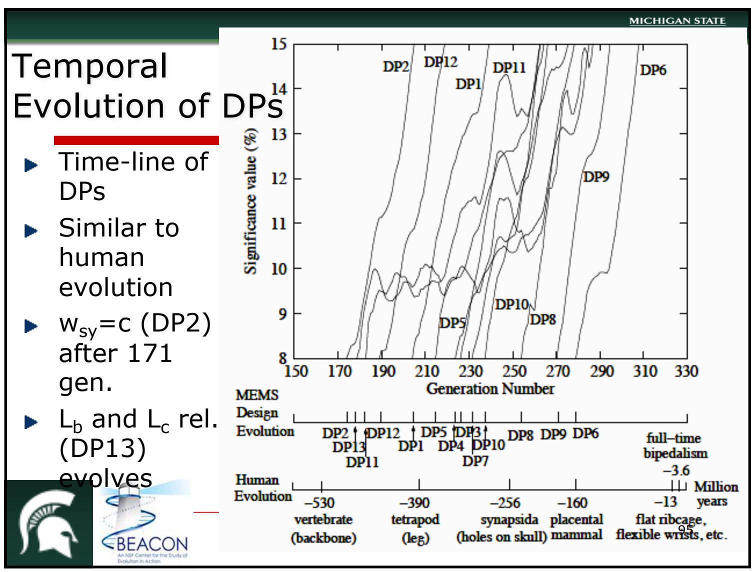


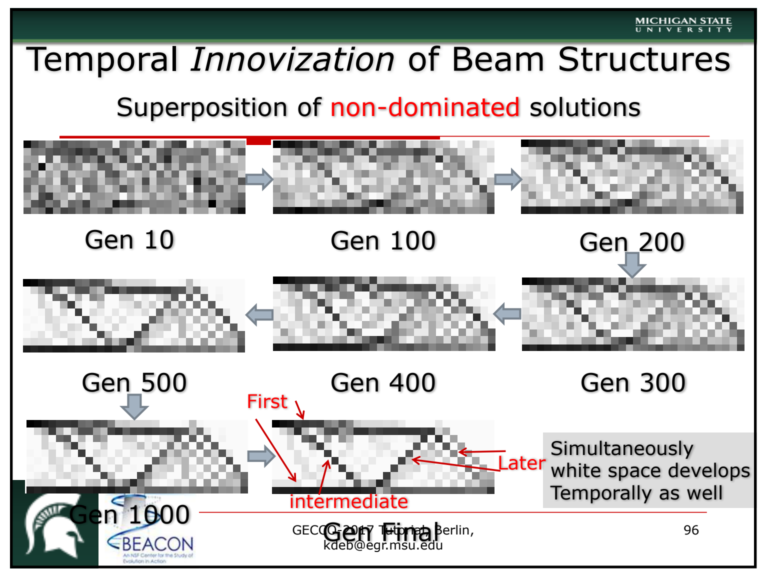
Procedure:

- Multiple fronts put together -> straightforward
- Perform *innovization* task


Lower-Level Innovization
Knee region
Over Preferred PO solutions
Reference point based solutions, weight based solutions
Over specified values of objective or constraints or variable boundaries
Over some fixed values of variables
Knee region
R-NSGA-II solutions
Reference point
Pareto frontier
Reference point
GECCO-2017 Tutorial, Berlin,
kdeb@egr.msu.edu
93


Temporal Innovization: MEMS Design Problem
14 design variables, 24 constraints
Obj: (i) min power consumption (ii) min area
Step 1
Comb drive
Anchor points
Folded flexure
Shuttle mass
94

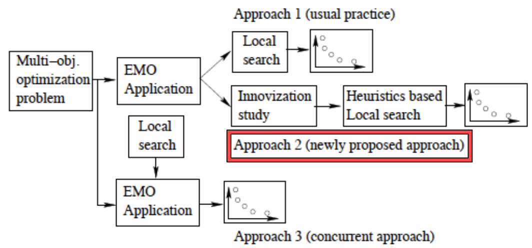

Temporal Evolution of DPs
Time-line of DPs
Similar to human evolution
w_{sy}=c (DP2) after 171 gen.
L_b and L_c rel. (DP13) evolves
Significance value (%)
Generation Number
MEMS Design Evolution
Human Evolution
Million years
vertebrate (backbone)
tetrapod (leg)
synapsida (holes on skull)
placental mammal
flat ribcage, flexible wrists, etc.
full-time bipedalism
95


Temporal Innovization of Beam Structures
Superposition of non-dominated solutions
Gen 10
Gen 100
Gen 200
Gen 500
Gen 400
Gen 300
First
Later
intermediate
Gen 1000
Gen Final
Simultaneously white space develops Temporally as well
GECCO-2017 Tutorial, Berlin,
kdeb@egr.msu.edu
96

Innovization to Speed-up Optimization

(Deb and Datta, 2013 EO)

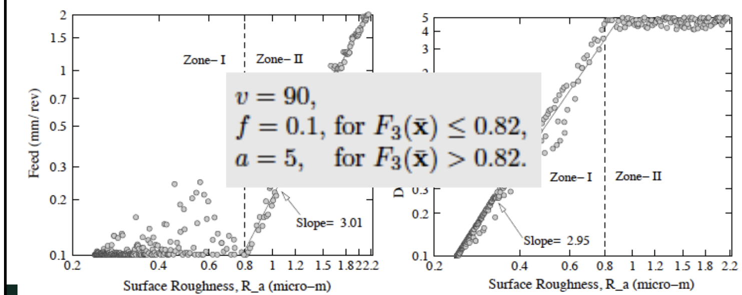
- Innovized principles as heuristics for local searches for a further EMO run
- A metal-cutting problem



GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 97

Feed and Depth of Cut

After a few generations of NSGA-II:

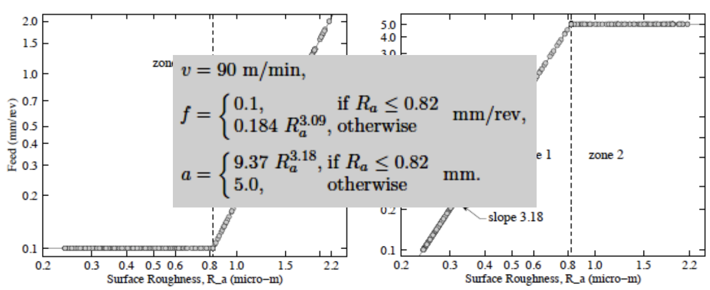


$$v = 90,$$
$$f = 0.1, \text{ for } F_3(\bar{\mathbf{x}}) \leq 0.82,$$
$$a = 5, \quad \text{for } F_3(\bar{\mathbf{x}}) > 0.82.$$

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 98

Feed and Depth of Cut (cont.)

After Heuristics Based EMO:



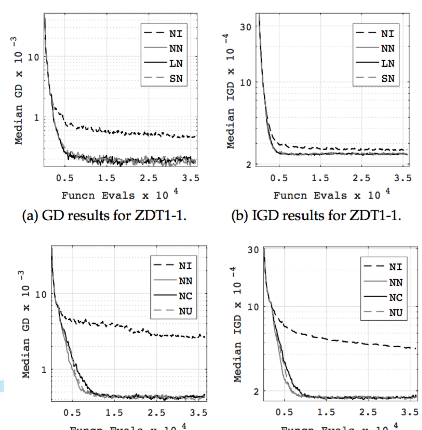
$$v = 90 \text{ m/min},$$
$$f = \begin{cases} 0.1, & \text{if } R_a \leq 0.82 \\ 0.184\ R_a^{3.09}, & \text{otherwise} \end{cases} \text{ mm/rev},$$
$$a = \begin{cases} 9.37\ R_a^{3.18}, & \text{if } R_a \leq 0.82 \\ 5.0, & \text{otherwise} \end{cases} \text{ mm.}$$

Currently under study

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 99

New Results on Innovization Based EMO Convergence

(Gaur and Deb, CEC-2017)



(a) GD results for ZDT1-1. (b) IGD results for ZDT1-1.

(a) GD results for ZDT1-2. (b) IGD results for ZDT1-2.

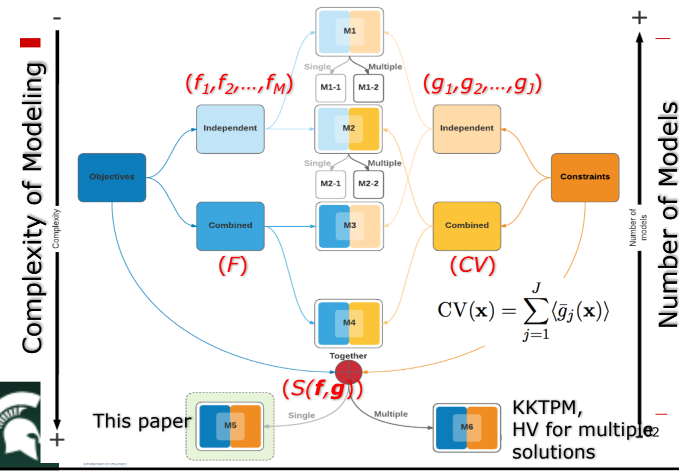
100

EMO for Handling Practicalities

- Metamodeling based EMO
- Uncertainty handling EMO
- Distributed computing in EMO
- Objective reduction in EMO
- Dynamic EMO
- Bilevel EMO
- Etc.

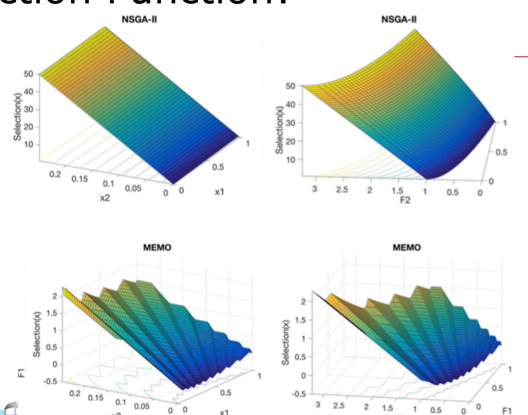
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 101

Metamodeling in Multi-objective Optimization



102

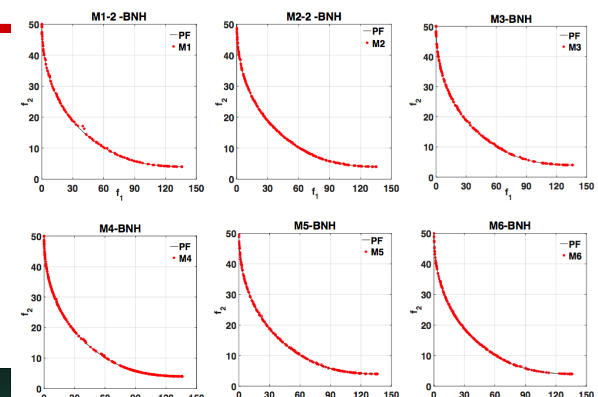
Selection Function:



103

First Results

(Deb et al., EMO-17)

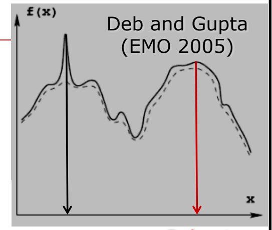


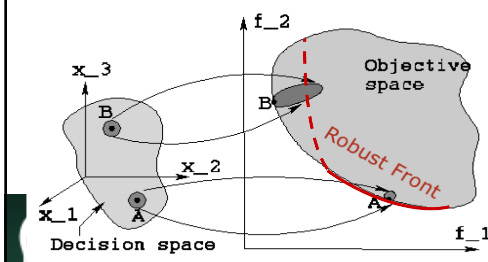
104

Robust EMO:

Handling uncertainties in variables and parameters

- Parameters are uncertain and sensitive to implementation
  - Find robust, instead of optimal, solution





- Robustness can be used as DM tool
- Robust front can be identified

105

Multi-Objective Robust Solutions of Type I and II

- Similar to single-objective robust solution of type I

$$\begin{aligned} &\textit{Minimize} \quad (f_1^{\text{eff}}(\mathbf{x}), f_2^{\text{eff}}(\mathbf{x}), \ldots, f_M^{\text{eff}}(\mathbf{x})), \\ &\textit{subject to} \quad \mathbf{x} \in \mathcal{S}, \end{aligned}$$

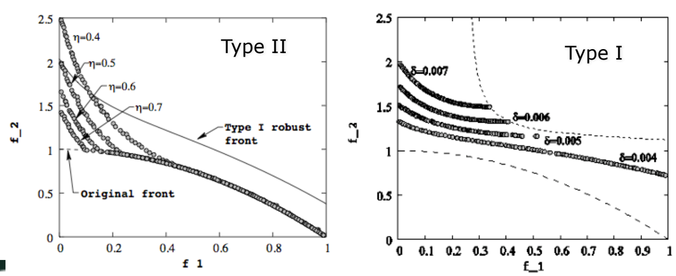
- Type II

$$\begin{aligned} &\textit{Minimize} \quad \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x})), \\ &\textit{subject to} \quad \frac{\|\mathbf{f}^p(\mathbf{x}) - \mathbf{f}(\mathbf{x})\|}{\|\mathbf{f}(\mathbf{x})\|} \le \eta, \\ &\qquad\qquad\quad \mathbf{x} \in \mathcal{S}. \end{aligned}$$

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 106

Robust Frontier for Two Objectives

- Identify robust region
- Allows a control on desired robustness



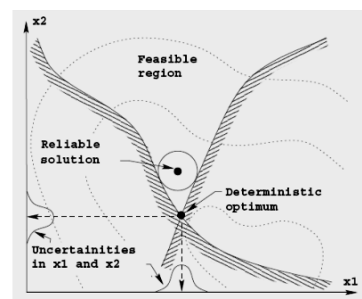
GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 107

Reliability-Based Optimization:

Making designs safe against failures

- Deterministic optimum is not usually reliable
- Reliable solution is an interior point
- Chance constraints with a given reliability

Minimize $\mu_f + k\sigma_f$
Subject to $Pr(g_j(x)\geq 0) \geq \beta_j$
β_j is user-supplied

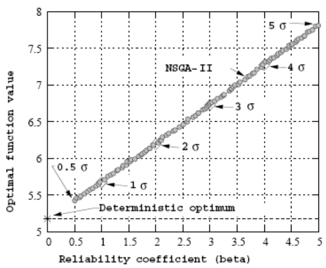


Deb et al. (EMO 2005)

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 108

MICHIGAN STATE UNIVERSITY

Multiple Reliability Solutions: *Get a better insight*



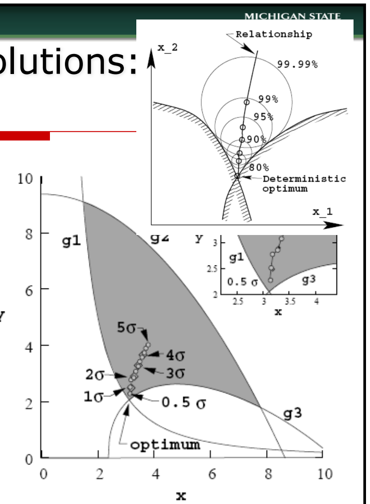
Maximize $x+y$,

Subject to $g_1(x,y)\equiv \frac{1}{20}x^2y-1\ge 0$,

$g_2(x,y)\equiv\frac{1}{30}(x+y-5)^2+\frac{1}{120}(x-y-12)^2-1\ge 0$,

$g_3(x,y)\equiv\frac{80}{x^2+8y+5}-1\ge 0$,

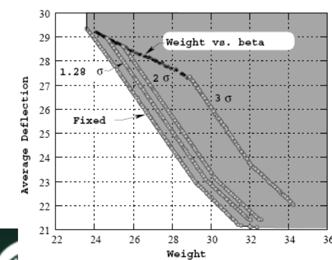
$0\le x,y\le 10$.

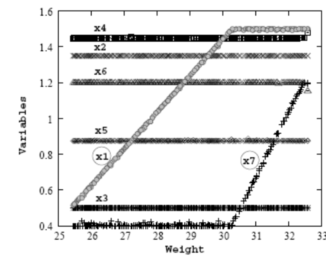


MICHIGAN STATE UNIVERSITY

Multi-Objective Reliability-Based Optimization

- Reliable fronts show rate of movement
- What remains unchanged and what gets changed!

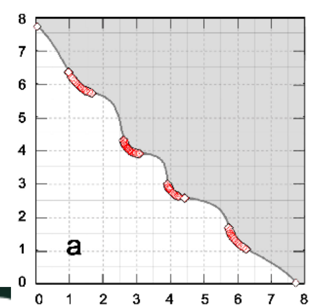


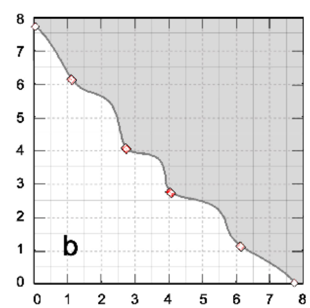


MICHIGAN STATE UNIVERSITY

Finding Knee Solutions

Branke et al. (2004) for more details



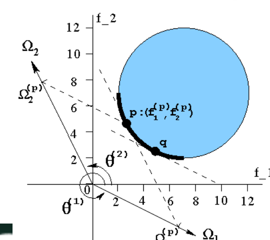


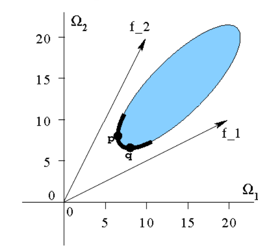
MICHIGAN STATE UNIVERSITY

Distributed Computing of Pareto-Optimal Set

Deb, Zope & Jain (EMO-2003)

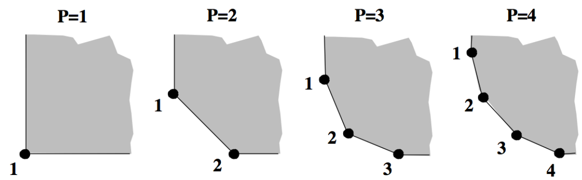
- Guided domination concept to search different parts of Pareto-optimal region
- Distributed computing of different parts





Distributed Computing Idea Using Cone-Domination

- Use non-overlapping cones

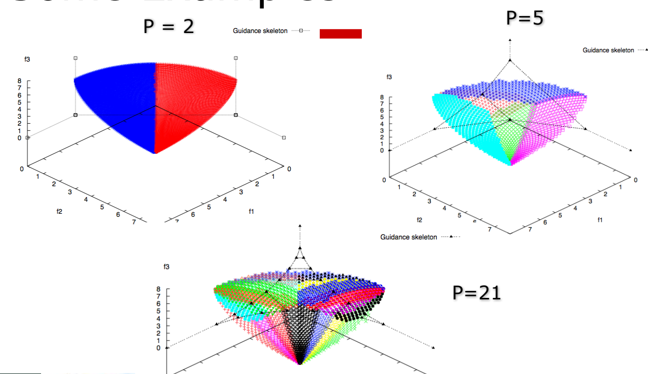


- Inclusive angle:

$$\phi = \pi \left(1 - \frac{1}{2P}\right)$$

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 113

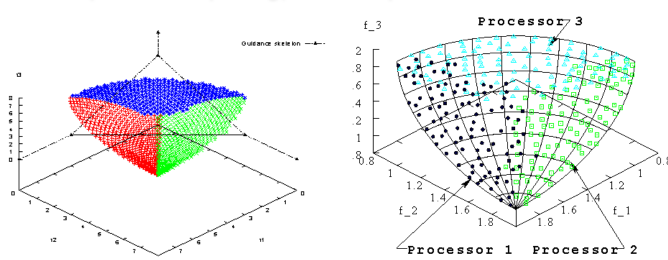
Some Examples



GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 114

Distributed Computing: A Three-Objective Problem

- Spatial computing, not temporal

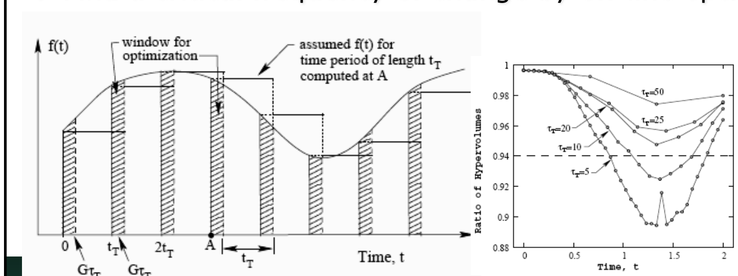


Theory NSGA-II Simulations

GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 115

Dynamic Multi-Objective Optimization

- Assume a *statis* in problem for a time step
- Find a critical frequency of change by off-line opt.

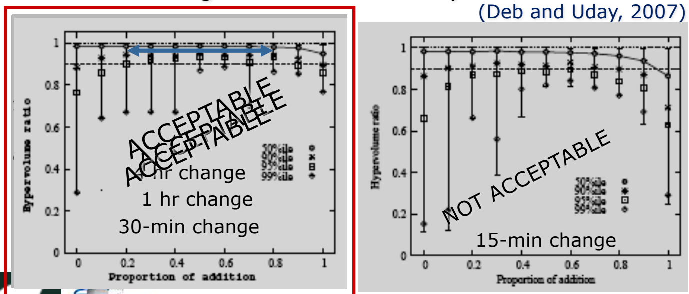


GECCO-2017 Tutorial, Berlin, kdeb@egr.msu.edu 116

Dynamic Hydro-Thermal Power Scheduling

- Addition of random or mutated points at changes
- 30-min change found satisfactory

(Deb and Uday, 2007)



Objective Reduction

- Identify redundant objectives
- EMO+PCA in iterations

10-objective DTLZ5 problem

Iter.1

Iter. 1 : PCA-1 (58.83 % variance)	f_7	f_{10}
PCA-2 (28.26 % variance)	f_1	
PCA-3 (06.53 % variance)		f_8
PCA-4 (03.27 % variance)		f_8

	f_1	f_7	f_8	f_{10}
f_1	+	-	+	-
f_7	-	+	+	-
f_8	+	+	+	-
f_{10}	-	-	-	+

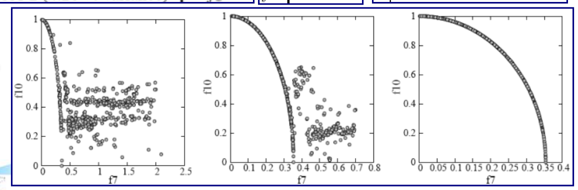
Iter.2

Iter. 2 : PCA-1 (94.58 % variance)	f_7	f_{10}
PCA-2 (4.28 % variance)	f_8	

	f_7	f_8	f_{10}
f_7	+	+	-
f_8	+	+	-
f_{10}	-	-	+

	e1:0.9458	e2:0.0428	
f_7	+0.543	-0.275	c7=0.5253
f_8	+0.457	+0.672	c8=0.4610
	PCA1	PCA2	

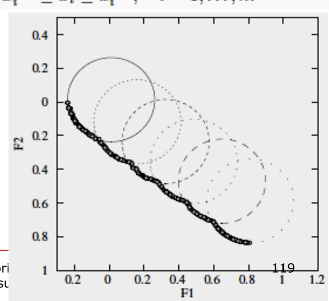
Saxena and Deb (CEC-2006, EMO-2007, CEC-2007)



Multi-Objective Bilevel Programming (Deb and Sinha, EMO-2009)

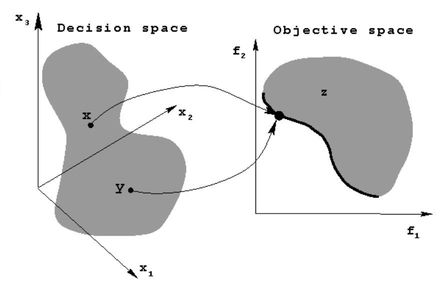
- Upper level solution is feasible only if it is a lower level PO solution
- Often appears in engineering problems to deal with stability, equilibrium etc.
- NSGA-II in both levels

$$\min_{(\mathbf{x}_u,\mathbf{x}_l)} \mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}),\ldots,F_M(\mathbf{x})),$$
$$\text{s.t.} \quad \mathbf{x}_l \in \operatorname{argmin}_{(\mathbf{x}_l)} \{\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}),\ldots,f_m(\mathbf{x})) \mid \mathbf{g}(\mathbf{x}) \geq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\},$$
$$\mathbf{G}(\mathbf{x}) \geq \mathbf{0}, \mathbf{H}(\mathbf{x}) = \mathbf{0},$$
$$x_i^{(L)} \leq x_i \leq x_i^{(U)}, \quad i = 1,\ldots,n.$$

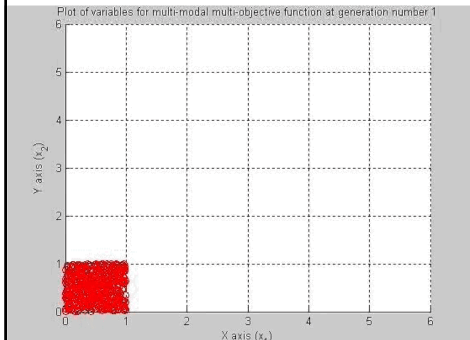


Multi-Modal EMOs

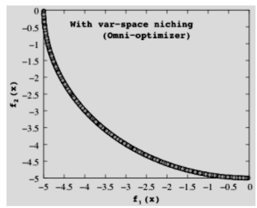
- Different solutions having identical objective values
- Multi-modal Pareto-optimal solutions: Design, Bioinformatics
- Find multiple solutions having identical objective values
- Modified crowding approach in NSGA-II



Multi-Modal Multi-Objective Problem (Omni-Optimizer, EJOR-2008)



Each part produces
An identical front

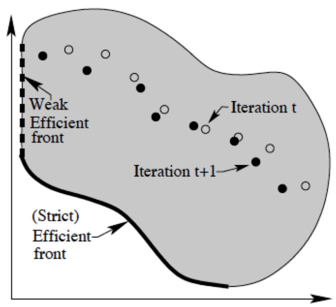


Theoretical Developments

- KKT Proximity Measure (KKTPM) for convergence
- Performance Measures
  - Hypervolume
  - Attainment surfaces
  - R-HV for reference point based EMO
- Other theoretical studies

KKT Proximity in Evolutionary Multi-objective Optimization

- As points move towards efficient front, KKTPM must reduce
- KKTPM must have similar values for points equidistant from front



KKT Optimality Conditions for MO

$$\sum_{k=1}^{M} \lambda_k \nabla f_k(\mathbf{x}^k) + \sum_{j=1}^{m} u_j \nabla g_j(\mathbf{x}^k) = \mathbf{0}, \quad \text{Equilibrium condition}$$

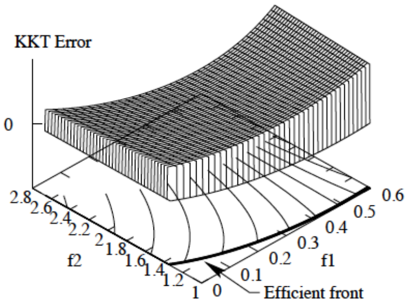
x^k is supplied

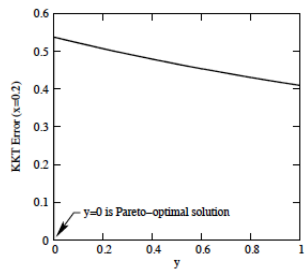
$$\begin{aligned} g_j(\mathbf{x}^k) &\leq 0, \quad j = 1,2,\ldots,J, \quad \text{Constraint satisf.} \\ u_j g_j(\mathbf{x}^k) &= 0, \quad j = 1,2,\ldots,J, \quad \text{Compl. slackness} \\ u_j &\geq 0, \quad j = 1,2,\ldots,J, \quad \text{Non-neg. of mult.} \\ \lambda_k &\geq 0, \quad k = 1,2,\ldots,M, \text{ and } \boldsymbol{\lambda} \neq \mathbf{0}. \end{aligned}$$

- Find $\boldsymbol{\lambda}^*, \mathbf{u}^*$ for minimum KKT Error:

$$\begin{aligned} \text{Minimize}_{(\boldsymbol{\lambda},\mathbf{u})} \quad & \|\textstyle\sum_{k=1}^{M} \lambda_k \nabla f_k(\mathbf{x}^k) + \sum_{j=1}^{m} u_j \nabla g_j(\mathbf{x}^k)\|, \\ \text{Subject to} \quad & g_j(\mathbf{x}^k) \leq 0, \quad j = 1,2,\ldots,J, \\ & u_j g_j(\mathbf{x}^k) = 0, \quad j = 1,2,\ldots,J, \\ & u_j \geq 0, \quad j = 1,2,\ldots,J, \\ & \lambda_k \geq 0, \quad k = 1,2,\ldots,M, \text{ and } \boldsymbol{\lambda} \neq \mathbf{0}. \end{aligned}$$

KKT Error on P1

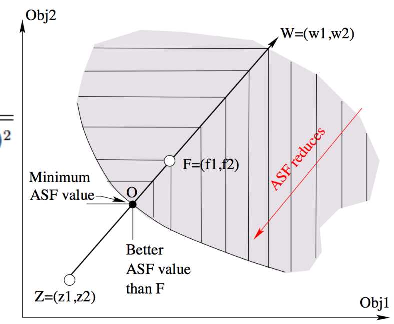




- KKT Error increases towards efficient front
- KKT Error cannot be a metric

KKTPM Concept

- 1-obj KKTM extended
- ***z***: utopian point
- ***w***: $w_i = \frac{f_i(\mathbf{x}) - z_i}{\sqrt{\sum_{k=1}^{M}(f_k(\mathbf{x}) - z_k)^2}}$
- If *F* is on efficient front, it is minimum ASF



KKT Proximity Metric (cont.)

Treat ASF as single-obj. problem

- Solve and find optimal ϵ_k*

$$\begin{aligned} \text{Minimize}_{(\epsilon_k, x_{n+1}, \mathbf{u})} \quad & \epsilon_k + \sum_{j=1}^{J}\left(u_j g_j(\mathbf{x}^k)\right)^2, \\ \text{Subject to} \quad & \|\nabla F(\mathbf{y}) + \sum_{j=1}^{J+M} u_j \nabla G_j(\mathbf{y})\|^2 \le \epsilon_k, \\ & \sum_{j=1}^{M+J} u_j G_j(\mathbf{y}) \ge -\epsilon_k, \\ & u_j \ge 0, \quad j = 1, 2, \ldots, (M+J), \\ & -\epsilon_k \le 0, \\ & -x_{n+1} \le 0. \end{aligned}$$

1. Relax compl. slackness cond.
2. Add a penalty

- Define KKTPM:

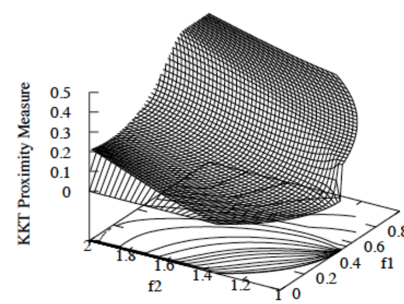
Use Matlab's fmincon() to solve it

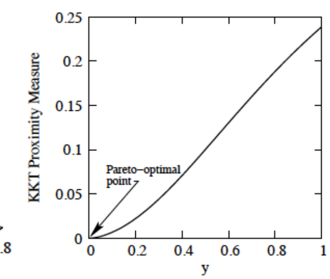
- 1 linear and 1 quadratic constraints

$$\text{KKT Proximity Measure}(\mathbf{x}^k) = \begin{cases} \epsilon_k^*, & \text{if } \mathbf{x}^k \text{ is feasible,} \\ 1 + \sum_{j=1}^{J}\langle g_j(\mathbf{x}^k)\rangle^2, & \text{otherwise.} \end{cases}$$

KKT Proximity Metric on P1

- Smooth reduction to zero

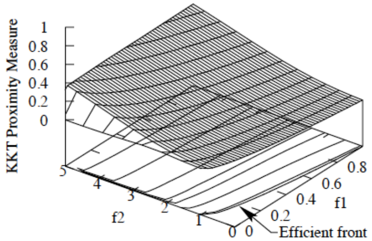




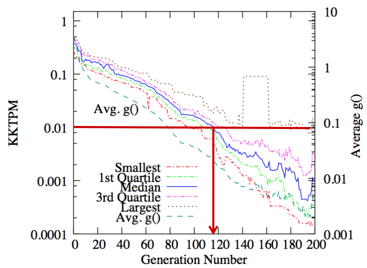
- Contour parallel to PO front

ZDT1 Test Problem with NSGA-II

KKTPM Surface



NSGA-II Populations
N=100

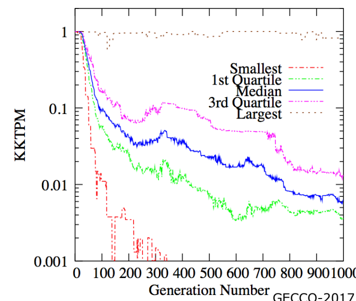


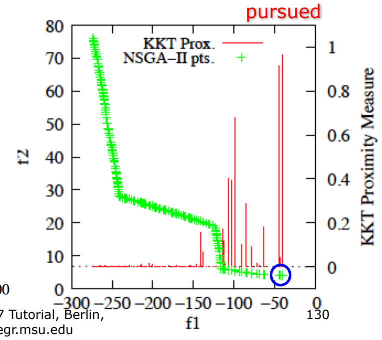
- KKTPM parallel to PO front
- Population converges as a front

Problem OSY with NSGA-II

- 25% points did not converge until 250 gen.
- Local search to speed up EMO runs

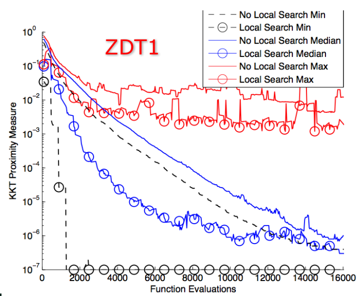
Local search method is bein pursued

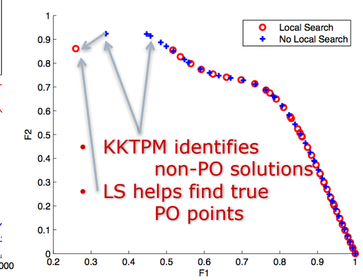




Results on Local Search Based EMO

- With and without local search

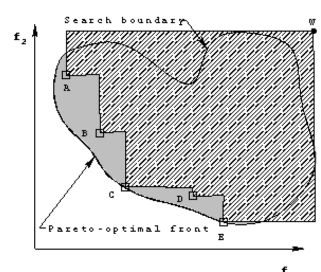




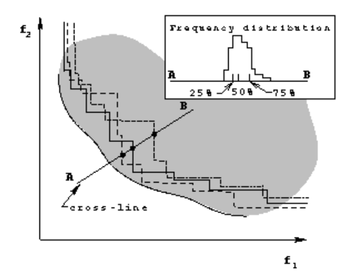
KKTPM allows differential treatment of ND solutions

Metrics for Convergence and Diversity

Hypervolume



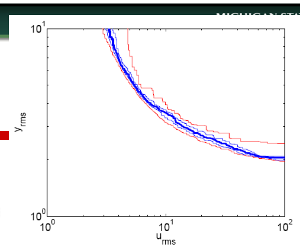
Attainment Surface Method

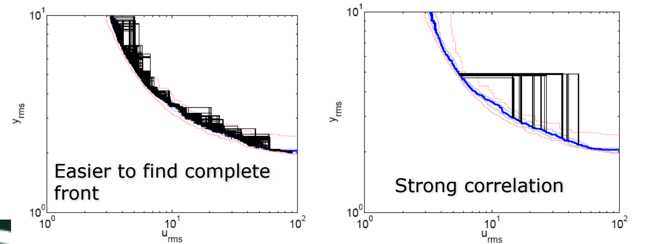


Attainment Surface Methods

- Positive and negative relationships

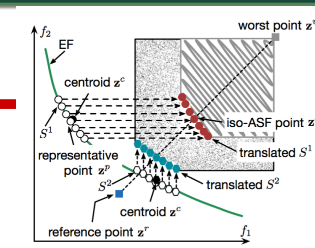
Fonseca (2005; 2007)

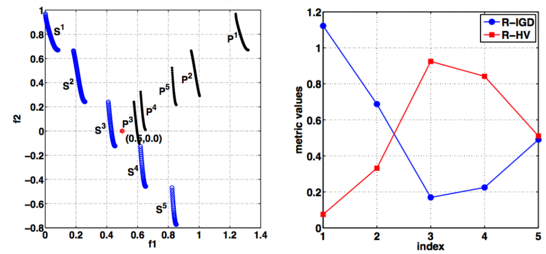




Reference Point Based HV

- Move points on the reference line and compute HV
- Results on ZDT3





(a) Population distribution. (b) $\mathbf{z}^r = (0.5, 0.0)^T$.

Further Theoretical Developments

- Hypervolume estimator (While et al., 2006, Bader et al., 2008)
  - Sample objective space to determine HV (WFG, Perth; Dortmund)
- Set-based search methods (SMS-MOEA, Beume, et al., 2007; HypE, Bader and Zitzler, 2011)
  - Single-objective search that maximizes HV or select points based in incremental HV
- EMOs with bounded computational complexity (LOTZ Boolean problem)
  - SEMO, FEMO methods (Laumanns et al., 2002)

Conclusions

- EMO is a fast-growing field of research and application
  - Exciting field within Computational Intelligence
- Practical applications and challenges were continuously addressed
- EMO+MCDM, EMO+Math optimization
- EMO is diversifying into new areas
- Commercial softwares available
  - ModeFrontier, iSight, VisualDoc
- Computer codes freely downloadable
  - Jmetal, PISA, MOEAFramework, EMOO websites