EVOLUTIONARY ALGORITHMS: What are EAs? Mathematical Formulation & Computer Implementation Multi-objective Optimization, Constraints Computing cost reduction

Kyriakos C. Giannakoglou

Associate Professor,

Lab. of Thermal Turbomachines (LTT),

National Technical University of Athens (NTUA),



GREECE

Outline

- From traditional problem solving techniques to EAs.
- Generalized EA: Basic and advanced operators.
- Mathematical foundations of EAs.
- EAs for multi-objective optimization.
- Distributed Evolutionary Algorithms (DGAs).
- Hierarchical Evolutionary Algorithms (HGAs).
- Constraints' handling.
- Efficient ways for reducing the computing cost of EAs.
- Applications in the field of aeronautics, turbomachinery, energy production, logistics.

Introductory Course to Design Optimization

Effective/Efficient Problem Solving Techniques

- The number of possible solutions in the <u>search space</u> is so large as to forbid exhaustive search
- Seeking the best combination of approaches that addresses the purpose to be achieved
- Finding the solution using the available computing resources
- Finding the solution within the available time
- One or more (contradictory) targets
- Solving the problem under a number of (hard/soft) constraints

Introductory Course to Design Optimization

Basic Concepts of Problem Solving Techniques

- <u>Representation</u>: how to encode alternative candidate solutions for manipulation
- Objective: describes the purpose to be fulfilled
- <u>Evaluation function</u>: returns a value that indicates the (numeric of ordinal) quality of any particular solution, given the representation

Introductory Course to Design Optimization

Hill-Climbing: A Traditional (Deterministic) PST

Useful Definitions:

- Neighborhood of a solution
- Local Optimum

Requirements:

- A starting point
- Computation of gradient
- Termination criteria



 \mathcal{X}_{1}

Ideas for creating Hill-Climbing Method variants

- How to select the new solution for comparison with the current solution (how to compute the gradient) ...
- To use more than one starting solutions, if necessary ...
- To be "less deterministic" ...

Introductory Course to Design Optimization

Algorithms Relying on Analogies to Natural Processes

- Evolutionary Programming
- Genetic Algorithms
- Evolution Strategies
- Simulated Annealing
- Classifier Systems
- Neural Networks

Traditional Problem Solving Techniques Algorithms with analogies to Natural Processes

Introductory Course to Design Optimization

The Subclass we are interested in ...

Methods which are based on the principle of evolution (i.e. the survival of the fittest)

- Handling populations of candidate solutions
- Undergoing unary (mutation-type) operations
- Undergoing higher-order (crossover-type) operations
- Using a selection scheme biased towards fitter individuals



Introductory Course to Design Optimization

$$g = 0, \ S^g = S^{random}$$

 $\begin{bmatrix} \vec{y} = \vec{F}(\vec{x}), \ \forall \ s \in \ S^g \ \phi(\vec{y}), \ \forall \ s \in \ S^{g,\mu} \ S^{g+1} = T_3(T_2(T_1(S^g))) \ g = g + 1 \ ext{if}(ext{converge}(g)) ext{ end}$

Introductory Course to Design Optimization



Introductory Course to Design Optimization

Encoding the free variables – Binary Coding: $b_m, m=1,M$ Binary digits per variable

 x_m Gene

 \vec{x} Candidate solution

Chromosome: $\underbrace{0101}_{x_1}\underbrace{101}_{x_2}\dots\underbrace{10111}_{x_M}$ $x_m = L_m + \frac{U_m - L_m}{2^{b_m} - 1}\sum_{i=1}^{b_m} 2^{i-1}d_{m,i}$

 L_m , $\, U_m\,$ user defined bounds

Introductory Course to Design Optimization

Schemata in binary strings:

A <u>schema</u> is a similarity template describing a subset of strings with similarities at certain string positions (Holland, 1968)



Order of schema o(S)=3 (fixed digits)



m=length of string(=5)
r=number of *'s (=2)

2^m possible schemata

2^r strings matched by this schema

Introductory Course to Design Optimization

Order of schema o(S)=3 (fixed digits)

The <u>order</u> of a schema affects its survival probabilities during *mutation*

The <u>defining length</u> of a schema affects its survival probabilities during *crossover*

(Schema Theorem)

Introductory Course to Design Optimization

Representation:

$$(x_1, x_2, \ldots, x_m, \ldots, x_M)$$

A step ahead:

Representation including evolution parameters (the concept of ES):

$$(x_1, x_2, \dots, x_m, \dots, x_M, \sigma_1, \sigma_2, \dots, \sigma_m, \dots, \sigma_M)$$

Introductory Course to Design Optimization

The effect of selection:

(m_g) examples of a particular schema (S), generation (g) F(S)=average fitness of the strings matching schema (S)

 $m_{g+1} = m_g \frac{F(S)}{F_{mean,g}}$

Reproductive Schema Growth Equation

$$m_{g+1} = m_g(1+k)$$

m _{g+1} >m _g	if F(S)>F _{mean,g}	
m _{g+1} <m<sub>g</m<sub>	if F(S) <f<sub>mean,g</f<sub>	

$$m_{g+1} = m_0 (1+k)^{g+1}$$

Long term effect of selection (k=constant)

Introductory Course to Design Optimization

The effect of crossover:

Possibility of destructing the schema S during crossover:





Possibility of maintaining the schema S:

$$p_{main} = 1 - p_{Xover} \frac{d(S)}{m-1}$$

Schema Growth Equation (after selection & crossover):

$$m_{g+1}(S) \ge m_g(S) \frac{F(S)}{F_{mean,g}(S)} (1 - p_{Xover} \frac{d(S)}{m-1})$$

Introductory Course to Design Optimization

The effect of mutation:

Possibility of maintaining the schema S after mutation:

$$p_{surv} = (1 - p_{mut})^{o(S)} \approx 1 - p_{mut}o(S)$$

****11010*01*1*11*

Final Schema Growth Equation: $m_{g+1}(S) \ge m_g(S) \frac{F(S)}{F_{mean,g}(S)} (1 - p_{Xover} \frac{d(S)}{m-1} - o(S)p_{mut})$

Short, low-order, above-average schemata should receive an (exponentially) increasing number of strings in the next generations Introductory Course to Design Optimization

Lessons Learned:

- Short, low-order, above-average schemata sould receive an (exponentially) increasing number of strings in the next generations (Schema Theorem).
- GA explore the search space by short, low-order schemata.
- GAs seek near-optimal performance through the juxtaposition of short, low-order, high-performance schemata (the so-called <u>building blocks</u>, <u>Building Block</u> <u>Hypothesis</u>).

Introductory Course to Design Optimization

Exploration vs. Exploitation :

- Exploration: seeking the global optimum in new and unknown areas in the search space.
- Exploitation: making use the knowledge gained from the previously examined points to guide the search towards new better points in the search space.

Holland 1975: GAs

<u>but:</u>

- Infinite population size
- Fitness function value accurately reflects the utility of a solution
- Genes in a chromosome do not interact significantly

Introductory Course to Design Optimization

Binary Coding:

- Creates lengthy binary strings if high accuracy is required.
- Offers the maximum number of schemata per bit of information, compared to any other coding.
- Facilitates theoretical analysis and the development of new genetic operators.
- Smallest alphabet that allows a natural expression of the problem (Goldberg, 1989).

Real Coding:

- Problem-tailored genetic operators can readily be devised.
- High-cardinality alphabets contain more schemata (Antonisse, 1989).

Introductory Course to Design Optimization

Generalized Evolutionary Algorithm (EA)



Introductory Course to Design Optimization

EA – Schematic Presentation:

$$\begin{split} g &= 0, \ S^{g,e} = \varnothing, \ S^{g,\mu} = \varnothing, \ S^{g,\lambda} = S^{random} \\ \begin{bmatrix} \vec{y} &= \vec{F}(\vec{x}), \ \forall \ s \in S^{g,\lambda} \\ S^{g+1,e} &= T_e(S^{g,\lambda} \cup S^{g,e}) \\ \phi(\vec{y}), \ \forall \ s \in (S^{g,\mu} \cup S^{g,\lambda} \cup S^{g+1,e}) \\ S^{g,\lambda} &= T_{e_2}(S^{g,\lambda} \cup S^{g+1,e}) \\ S^{g+1,\mu} &= T_{\mu} \left(S^{g,\mu} \cup S^{g,\lambda}\right) \\ S^{g+1,\lambda} &= T_m \left(T_r(S^{g+1,\mu} \cup S^{g+1,e})\right) \\ g &= g+1 \\ \text{if}(\text{converge}(g)) \text{ end} \end{split}$$

Introductory Course to Design Optimization



Introductory Course to Design Optimization

- Proportional selection
- Linear ranking
- Roulette wheel
- Probabilistic Tournament selection
- $f \square$ Indirect Selection ($\mu{<}\lambda$)

$$S^{g+1,\mu} = \mathcal{T}_{\mu,1} \left(S^{g,\mu} \cup S^{g,\lambda} \right)$$
$$S^{g+1,\mu} = \mathcal{T}_{\mu,2} \left(S^{prov,\mu} \right)$$

Introductory Course to Design Optimization

Proportional selection:

Select ($\lambda \rho$) individuals out of (μ) preselected ones

$$\frac{\mu \phi^{(s)}}{\sum_{i=1}^{\mu} \phi^{(\iota)}} \ge 1$$

Premature convergence due to the presence of a "super-fit" individual.

$$number_of_copies^{(s)} = int \left| \frac{\lambda \rho \mu \phi^{(s)}}{\sum_{i=1}^{\mu} \phi^{(\iota)}} \right|$$

Introductory Course to Design Optimization

Roulette Wheel:

Select ($\lambda \rho$) individuals out of (μ) preselected ones

Premature convergence due to the presence of a "super-fit" individual.

$$angle_slot = \frac{\phi^{(s)}}{\sum_{i=1}^{\mu} \phi^{(\iota)}} 360^{o}$$

Introductory Course to Design Optimization

Parent Selection Operator T_{μ} (5/5)

Fitness Ranking:

- Individuals are sorted in the order of fitness values
- Reproductive trials are assigned according to rank
- Linear Ranking
- Exponential Ranking, etc

- Overcomes the problem of the presence of an extreme individual
- **Fitness ranking performs better than fitness scaling.**

Introductory Course to Design Optimization



- Two-point
- One- or two-point per variable
- Discrete (variables' interchanging)
- Uniform (with parent-depending probability)

Introductory Course to Design Optimization

Recombination Operation T_r – Real Coding

- $\begin{array}{c|c} \bullet & \bullet = 2 & x_1 x_2 x_3 x_4 x_5 x_6 \\ & x_1 x_2 x_3 x_4 x_5 x_6 \end{array} \right\} \quad x_1 x_2 x_3 x_4 x_5 x_6 \end{array}$
 - $x_4 = x_4 + r(x_4 x_4), r \in [0, 1]$

- Two-point
- □ M-point $x_m = x_m + r_m (x_m x_m), m = 1, M, r_m \in [0, 1]$ □ Discrete, $\rho = 2$
- Discrete Panmictic ho=M (50% selection probability from each parent $\{1,2\}, \{1,3\}, ..., \{1,M\}$)
- Generalized Intermediate Panmictic

$$x_{\rm m} = x_{\rm m} + r_m (x_{\rm m,\rho} - x_{\rm m}), \ {\rm m} = 1, {\rm M}, \ r_m \in [0,1]$$

D Blend Xover (BLX-a), c=(1+2a)r-a, a=0.5

Introductory Course to Design Optimization

$$P_m \sim \frac{1...5}{\sum_{m=1}^M b_m}$$

0110001100 → 0010101100

<u>Dynamic Adjustment</u> of ${m P}_m$, depending on

- the number of generations without any improvement
- the number of evaluations without any improvement

Introductory Course to Design Optimization

Dynamic mutation probability per variable ${m P}_m$

$$x_{m} = \begin{cases} x_{m} + D(g, U_{m} - x_{m}), & r_{1} < 0.5\\ x_{m} - D(g, x_{m} - L_{m}), & r_{1} \ge 0.5 \end{cases}$$
$$D(g, a) = a \cdot r_{2} \cdot \left(1 - \frac{g}{g_{\max}}\right)^{p}$$
$$\dot{\eta} \quad D(g, a) = a \cdot \left(1 - r_{2}^{(1 - \frac{g}{g_{\max}})^{p}}\right)$$

p~0.2

... or, using number of evaluations, instead of number of generations

Introductory Course to Design Optimization

Mutation Operator T_m – Real Coding

$$\sigma_m = \sigma_m \cdot \exp(\tau \cdot N(0,1) + \tau \cdot N_m(0,1))$$
$$x_m = x_m + \sigma_m \cdot N_m(0,1)$$

$$\tau = \left(\sqrt{2\sqrt{M}}\right)^{-1}$$
$$\tau' = \left(\sqrt{2M}\right)^{-1}$$

Introductory Course to Design Optimization

Genetic Algorithms or Evolution Strategies or ...

Binary	Coding
--------	--------

- \square $\mu = \lambda$, parents=offspring [Holland, 1970]
- **ρ=2, two-parent recombination**[Goldberg, 1989]
- κ=0, zero life-span
- **P**_r<1, recombination probability^[Fogel]
- □ *K*=1, one target
- Real Coding, including the evolution parameters
 Without parent Selection Operator (μ<λ)
- □ *μ*<<*λ*
- $\square \quad \kappa=0 \ (\mu,0,\lambda) = (\mu,\lambda) \text{ or } \kappa=\infty \ (\mu,\infty,\lambda) = (\mu+\lambda)$
- *ρ*=2
- *P_r*=1 [Schwefel, Rechenberg, 1965]
 K=1
 - [Bäck, 1996]

Introductory Course to Design Optimization

[Michalewicz, 1994]

K.C. Giannakoglou, Associate Professor NTUA, Greece

ES

GA

Multi-objective Optimization



(K = number of objectives)

Introductory Course to Design Optimization

Pareto Front



Dominant Solution:

 $\begin{aligned} \vec{x}^{(p)} &< \vec{x}^{(q)} \Leftrightarrow s^{(p)} < s^{(q)} \Leftrightarrow \\ \forall \ k \in \{1, \dots, K\} : F_k^{(p)} \leq F_k^{(q)} \land \\ \exists \ k \ \in \{1, \dots, K\} : F_k^{(p)} < F_k^{(q)} \end{aligned}$

Pareto Optimal Solution:

 $\vec{x} \Leftrightarrow \not\exists \vec{x'} \in \mathbb{R}^M : \vec{x'} > \vec{x}$

Introductory Course to Design Optimization

Multi-Objective Optimization – Computation of φ

$$\phi = \sum_{k=1}^{K} w_k \cdot F_k$$

$$\phi = F_k, \ k \in [1, K]$$

For parent selection, VEGA [Schaffer, 1984]

$\phi = \text{Pareto}_{\text{Rank}}(F_k), \ k = 1, K$

[Goldberg, 1989]

Introductory Course to Design Optimization

Vector Evaluated Genetic Algorithm (VEGA)



Recombination, Mutation

Introductory Course to Design Optimization

VEGA: Guess the final solutions...



φ computation using the Pareto front:

- **Front Ranking [Goldberg, 1989]**
- Niched Pareto GA (NPGA) [Horn, Nafpliotis, 1993]
- Nondominated Sorting GA (NSGA) [Srinivas, Deb, 1994]
- **Strength Pareto EA (SPEA)** [Zitzler, Thiele, 1998]
- Pareto Envelope-based Selection Algorithm (PESA) [Corne, Knowles, Oates, 2000]
- **NSGA-II** [Deb, Agrawal, Pratap, Meyarivan, 2000]
- Strength Pareto EA II (SPEA II) [Zitzler, Laumanns, Thiele, 2001]

Introductory Course to Design Optimization

φ computation – Sorting (NSGA)



Introductory Course to Design Optimization

φ computation – Niching (NSGA)



Introductory Course to Design Optimization

φ computation – SPEA



Ways to reduce the number of evaluations:



- Improved evolution operators
- Hybridization with other optimization methods
- Distributed EAs (island model)
- Hierarchical EAs (faster solvers)
- Use of <u>surrogate models</u> (<u>metamodels</u>, fast approximate model)

Introductory Course to Design Optimization

Hybridization with other Optimization Methods



Introductory Course to Design Optimization

Use of Surrogate Evaluation Models

- Polynomial Interpolation
- Artificial Neural Networks
 - Multilayer Perceptron
 - Radial Basis Function Networks
- Kriging



Ways of using the surrogate evaluation model:

- Decoupled from the exact evaluation tool (+Design of Experiments, DoE)
- In combination with the exact evaluation tool
 - Regular Training (depending on the number of new entries in the DB)
 - Dynamical Training (separately, for each new individual)

Introductory Course to Design Optimization

Use of Surrogate Models (with Off-Line Training)



Introductory Course to Design Optimization



Introductory Course to Design Optimization

Inexact Pre-Evaluation (IPE) – The Concept:

Exact Evaluation of the most promising solutions



Introductory Course to Design Optimization

Local Surrogate Models - Training Set:



Introductory Course to Design Optimization

Surrogate Models – What else do they tell us???

- Fitness Function Approximation
- Confidence Intervals
- Hessian Matrix Approximation
- Sensitivity Derivatives (Importance Factors)



Introductory Course to Design Optimization



□ Master: EA Module - waiting list for evaluations

□ Slave: discrete (remote) evaluation process

Introductory Course to Design Optimization

Distributed EA

- Why ?
 - L (<< λ) < CPUs</p>
 - Persistant diversity in populations
 - Straightforward parallelization
- Additional Parameters:
 - Number of islands
 - Communication topology
 - Communication frequency
 - Migration algorithm
 - EA parameters per island



Introductory Course to Design Optimization

Distributed EA on a Multi-Processor System



Introductory Course to Design Optimization

Applications



Developed by the National Technical University of Athens, (NTUA), Greece

Introductory Course to Design Optimization

Problem: Rastrigin's Function



Introductory Course to Design Optimization

Results: Rastrigin's Function



Introductory Course to Design Optimization

Problem: Three-Element Airfoil, Lift Maximization



Introductory Course to Design Optimization

Results: Three-Element Airfoil, Lift Maximization

	Initial	Optimal
Rotation Angle	28.1º	28.02°
Δx	-0.020	-0.03078
Δу	0.0269	0.01982
Rotation Angle	-37º	-36.96°
Δx	0.020	0.02016
Δγ	0.0249	0.02469





Results: Three-Element Airfoil, Lift Maximization



Introductory Course to Design Optimization



Problem: Design of a Compressor Cascade



- Minimum Total Pressure Losses
- Constraint on the minimum (maximum thickness)
- Desirable Flow turning

$$F_1 = \omega \cdot P_1 \cdot P_2$$

$$P_{1} = \exp(\frac{|t_{\max} - t_{thres}|}{t_{thres}}), \ t_{thres} = 0.9t_{\max,ref}, \ t_{\max} < t_{thres}$$
$$P_{2} = \exp(-\max(1, \frac{\Delta\alpha_{ref} - \Delta\alpha}{\Delta\alpha_{ref}})), \ \Delta\alpha = \alpha_{1} - \alpha_{2} > 0$$

Introductory Course to Design Optimization

Results: Design of a Compressor Cascade



Introductory Course to Design Optimization

Problem: Compressor Multi-Point Design



Compressor Multi-Point Design (1 OP – 1 target) (20, 2, 100)1 23 54 0.0182 0.0155 0.0275 ω L=5 0.0244 ω REF 0.0234 0.0208 0.0189 0.0237 0.03740.019 REF 0.5 REF L=100 0.0185 L=5 L=50.4 0.018 പ്റ0.3 0.2 0.0175 щ 0.1 0.017 0 0.0165 -0.10.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 'n. 0.016 x/c 0.0155 500 1000 1500 2000 2500 3000 **Evaluations** Introductory Course to Design Optimization



Introductory Course to Design Optimization

Compressor Multi-Point Design (5 OP – 3 targets)



Introductory Course to Design Optimization

Compressor Multi-Point Design (5 OP – 5 targets)



Introductory Course to Design Optimization

Problem: High-Lift, Low-Drag Optimization



Results: High-Lift, Low-Drag Optimization





Optimization of Combined Cycle GT Power Plants



Design variables

- > HP steam pressure
- > LP steam pressure
- superheated HP steam temperature
- feedwater temperature at the inlet to the HP evaporator
- feedwater temperature at the outlet from the first HP economizer
- feedwater temperature at the inlet to the LP evaporator
- superheated LP steam temperature
- steam pressure fed to the water tank
- exhaust gas mass flow ratio (percentage of mass flowrate traversing the LP economizer)
- exhaust gas temperature at the HRSG outlet
- steam extraction pressure from the LP steam turbine
- exhaust gas temperature at the inlet to the condensate preheater

Natural gas fired, dual-pressure CCGTPP configuration GT: 260 MWe, 38% efficiency, exhaust gas mass flow 615 kg/sec at 600C.

Introductory Course to Design Optimization

Optimization of Combined Cycle GT Power Plants

