Motivations

- No explicit fitness function exists: to define fitness quantitatively
- Fitness evaluation is highly time-consuming: to reduce computational time
- Fitness is noisy: to cancel out noise
- Fitness is highly rugged: to smoothen the fitness landscape
- Search for robust solutions: to avoid additional expensive fitness evaluations

How To Estimate Fitness

- Ad hoc methods
  - Fitness inheritance (from parents)
  - Fitness imitation (from brothers and sisters)
- Problem approximation
  - To replace experiments with simulations
  - To replace full simulations with simplified simulations
- Data-driven functional approximation (meta-model)
  - Polynomials
  - Neural networks, e.g., multilayer perceptrons (MLPs), RBFN
  - Gaussian processes, Kriging models
  - Support vector machines

Fitness Inheritance and Imitation

- To estimate the fitness of an individual from that of its parents
  - weighted sum of parents’ fitness values (Smith et al, 1995, Salami et al, 2003)
- To estimate the fitness of an individual from that of other individuals of the same generation (Kim et al, 2001, Jin et al, 2004)
Issues in Choosing Meta-models

• Global model or local model
  ➢ For high-dimensional problems, local models are more practical
• Deterministic or stochastic models
  ➢ Stochastic models such as Gaussian models are able to provide an estimate and an error bound of the estimate
  ➢ Ensemble based error bound estimation
• Online learning capability
  ➢ Model is able to update recursively

Use Meta-models: Multiple Populations

• Injection island model (Elby et al, 1998)

• Hierarchical model (Sefrioui and Periaux, 2000)

Legends

- Model of lowest complexity
- Model of medium complexity
- Model of highest complexity
- Uni-directional migration
- Bi-directional migration

Meta-model in Initialization, Crossover and Mutations

• Generate a plurality of individuals randomly and choose the best ones according to the meta-model for the initial population
• Informed crossover (Rashed and Hirsch, 2000)
  ➢ Generate multiple offspring individuals (instead of 2) using crossover
  ➢ Keep the best two ones according to the meta-model
• Informed mutation (Rashed et al, 2000; Abboud et al, 2002)
  ➢ Generate multiple individuals randomly
  ➢ Choose the best according to the meta model

Meta-model in Fitness Evaluations (I)

• Use meta-models only: Risk of false convergence

Example of inconsistent error
Meta-model in Fitness Evaluations (II)

- Use meta-models together with the original fitness (if available)

<table>
<thead>
<tr>
<th>Individual-based control</th>
<th>Generation-based control</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st generation</td>
<td>1st generation</td>
</tr>
<tr>
<td>2nd generation</td>
<td>2nd generation</td>
</tr>
<tr>
<td>3rd generation</td>
<td>3rd generation</td>
</tr>
</tbody>
</table>

- Individuals evaluated with the original fitness
- Individuals evaluated with the meta-model

Individual-based Evolution Control

- Initialization
- Crossover/Mutation
- Evaluation with MM

- Evaluation with OF
- (Model update)
- Selection
- Evaluation with MM

Evolution Control Strategies

- Choose individuals randomly (Jin et al, 2000)
- Choose the best individuals according to the model (Jin et al, 2001; Jin et al, 2002a)
- Choose the potentially best individuals with the help of estimated error bound (Emmerich et al, 2002; Ulmer et al, 2003)
- Choose the most uncertain individuals (Branke, 2005)
- Choose the representative individuals (Kim et al 2001; Jin et al, 2004)

Random and Best Strategy

- 12-D Ackley function
- (3,12)-ES
- average over 10 runs

- The best strategy is more efficient than the random strategy
- In the best strategy, about half of the individual should be controlled to guarantee correct convergence
Clustering Method

- Group the population
- Choose the individual closest to the cluster center (Jin et al., 2004)
- Choose the best individual of each cluster (Gräning et al., 2005)

Meta-model Improvement

- Neural network ensemble (Jin et al., 2004)

Choose Uncertain Individuals

- Uncertainty measure (Branke et al., 2005)
  - Polynomials are used for meta-modeling
  - Define the following uncertainty measure:
    \[ u_i = \frac{1}{\sum_{j=1}^{K} d_{ij}} \]
    \( d_{ij} \): Euclidean distance between data point i and j;
    K: Number of data points in the neighborhood
  - Choose individuals for re-evaluation to minimize the following criterion
    \[ r = \sum u_i \]
  - A combination of quality and uncertainty:
    \[ r' = \sum p_i u_i, p_i \]: the probability of an individual to be selected

Pre-selection

- Initialization (µ, λ)-ES
- Crossover/Mutation
- Evaluation with MM
- Selection
- Evaluation with OF (Model update)
**Pre-selection Variants**

- Use RBF network and pre-selection based on fitness only (Won et al, 2003)
- Use Gaussian model, considering both fitness and uncertainty (Emmerich et al, 2002)
- Choose individuals having a large probability of improve (POI) (Ulmer et al, 2003)

\[ f(x) = f(x) - \alpha \sigma(x) \]

**Adaptation in Individual-based Evolution Control**

- Adaptation of \( \lambda' (\lambda' \geq \lambda) \) in pre-selection
- Adaptation of \( \lambda^* (1 \leq \lambda^* \leq \lambda) \) in best strategy or clustering methods
- Adaptation based on the quality of the meta-model
  - Increase \( \lambda'/\lambda \) if the quality of the meta-model becomes better or better than a certain criterion
  - Decrease \( \lambda'/\lambda \) if the quality of the model becomes worse, or worse than a certain criterion

**Performance Metrics**

- Approximation error: How good the meta-model approximates the fitness function
- Selection-based measure: For a (\( \mu, \lambda \))-ES, a meta-model is "perfect" if the same individuals are selected using meta-model or the original fitness function (Jin et al, 2002a; Hüsken et al, 2005)
- Calculate the correlation between the fitness value of the original fitness function and the meta-model (Jin et al, 2002b)
- Calculate the rank correlation (Jin et al, 2002b)

**Approximation Error**

\[
E^{\text{MDR}}(t-1) = E^{\text{NSR}}(t) = R(t-1) + R(t) \]

\[
\mu^{(\text{MDR})} \triangleq \frac{1}{\lambda} \sum \lambda \Delta \lambda \]

\[
\mu^{(\text{NSR})} \triangleq \frac{1}{\lambda} \sum \lambda \Delta \lambda \]

\[
\lambda' = \lambda + \mu^{(\text{MDR})} \Delta \lambda \]

\[
\lambda^* = \lambda - \mu^{(\text{NSR})} \Delta \lambda \]

\[
\lambda' = \lambda + \mu^{(\text{NSR})} \Delta \lambda \]

\[
\lambda^* = \lambda - \mu^{(\text{MDR})} \Delta \lambda \]
**Correct Selection**

- Assign a grade of $\lambda_i$ if the $i$-th best individual is correctly selected (Jin et al., 2002b; Hüsken et al., 2005)
  
  $\rho^{corr} = \sum (\lambda_i - 1)$

- The maximal grade that can be obtained is
  
  $\rho^{max} = \sum (\lambda_i - 1)$

- If $\mu$ individuals are selected randomly:
  
  $\rho^{rand} = \frac{2\mu - 1}{(\mu - 1)}$

- If the model is better than random guess ($\rho^{corr} > \rho^{rand}$)
  
  $\rho^{corr} = \sum \frac{(\lambda_i - \mu)}{\mu - 1}$

**Correlation**

- Fitness based correlation:
  
  $\rho^{corr} = 1 - \frac{\sum (y_i - \bar{y})^2}{n - 1}$

- Rank based correlation:
  
  $\rho^{corr} = \frac{1}{\sigma_y \sigma_x} \sum (r_{i,x} - \bar{r}_x) (r_{i,y} - \bar{r}_y)$

**Adaptive Pre-selection (I)**

- Impressive results have been obtained in adaptive pre-selection using selection based criterion (Ulmer et al., 2004)

- 10D Sphere function, no noise

- 10D Sphere function, changing noise level

**Adaptive Pre-selection (II)**

- Not that successful results (Gräning, 2005)

- Sphere function
Adaptive Pre-selection (III)

- Correct Selection
- Rank Correlation
- Fitness correlation

Generation-based Evolution Control

- Evolution control with a fixed frequency (Bull, 1999)
- Evolution control after convergence (Ratle, 1998)
- Evolution control with an adaptive control frequency (Jin et al, 2002a)
- Evolution control after convergence combined with uncertainty measure (Blühe et al, 2005)

Control Frequency Adaptation

- Adaptation of control frequency based on error estimation (Jin et al, 2001; Jin et al, 2002a)
- Estimation of model fidelity
- Adaptation of control frequency
- On-line model update

Meta-model Improvement

- Online structure optimization of neural networks (Hüsken et al, 2005)
Application Example

- 2-D aerodynamic design optimization
- B spline representation with 23 control points (46 parameters)
- Evolution strategy (2, 11) CMA-ES
- NN1: fully connected NN; NN2: structurally optimized NN

Surrogate Approach

- Generate a surrogate with initial data (Büche et al, 2005)
- Search on the meta-model until converges, restricting the search within
  the neighborhood of the current best solution:
  \[ x^\ast - d/2 \leq x \leq x^\ast + d/2 \]
  \[ d_i = \max(x_i) - \min(x_i), x_i \in NC \] closest neighbors
- Train the model using the NC closest data and NR most recently
  evaluated data to prevent the model from getting stuck

Local Search Using Meta-model

- Meta-model is used and updated only in local search (Ong et al, 2003)
- The trust-region framework for managing meta-models is applied to
  each individual during local search
  - The “trust-region” is a range in which the meta-model is trustful,
    first initialized as the range of the data used to construct the meta-
    model
  - The radius of the trust-region is reduced if the accuracy of the
    meta-model is low; enlarged if the accuracy is high, otherwise not
    changed (Dennis and Torczon, 1997)
  - The result of local search is coded back to the chromosome using
    the Lamarckian mechanism

Meta-model in Search for Robust Solutions

- Averaging based approach to robust solutions need a large number
  of additional fitness evaluations
- Meta-model can be used reduce computational cost (Paenke et al,
  2005)
Informed Operators

- Idea: replace randomness with decisions informed by the surrogate model
- Examples:
  - Informed initialization
  - Informed Crossover (parents, type)
  - Informed mutation (type, amplitude)

Informed mutation

- Crossover is done, followed by several random mutations
- Random mutations are evaluated using meta model, best becomes newborn

OEGADO (Objective Exchange GADO) for two objectives (Chafekar et al., 2005)

- 1. Both the GAs are run concurrently with each GA optimizing one of the two objectives while also forming a meta model of it.
- 2. At intervals equal to twice the population size, each GA exchanges its meta model with the other GA.
- 3. Informed operators are used. The IOs generate multiple children and use the meta model to compute the approximate fitness of these children. The best child is selected to be the newborn.
- 4. The true fitness function is then called to evaluate the actual fitness of the newborn corresponding to the current objective.
- 5. The individual is then added to the population using the replacement strategy.
- 6. Steps 2 through 5 are repeated till the maximum number of evaluations is exhausted.

OEGADO for three or more objectives

- 1. Each GA optimizes its objective and forms its own surrogate model.
- 2. After a given interval of evaluations each GA offers its meta model to one of the other GAs and obtains its meta model to use by its informed operators.
- 3. After the second interval each GA exchanges its meta model with one of the other remaining GAs.
- 4. This process continues and the GAs continue to exchange their meta models in a round-robin fashion.
Results (3-objective problem Remy)

- OEGADO did better with a good distribution over the covered area
- NSGA II & εMOEA did not achieve a good distribution

Results (more difficult 3-objective DTLZ8)

- OEGADO did well with a good distribution over the covered area
- NSGA II & εMOEA did not achieve a good distribution

Summary

- Meta-modeling and other fitness approximation techniques have found a wide range of applications
- Proper control of meta-models plays a critical role in the success of using meta-models
- Proper choice of a meta-model: with/without error estimation, local/global
- Application of meta-models to multi-objective optimization, dynamic optimizations, search for robust solutions poses many challenging problems
- Stringent theoretical analysis of EA dynamics using meta-models is very desirable

Further Information

- On-line bibliography: [http://www.soft-computing.de/amre.html](http://www.soft-computing.de/amre.html)
- Special issues:
  - Special issue on “Approximation and Learning in evolutionary computation”, Soft Computing, 9(1), 2005
  - Special issue on “Knowledge incorporation and extraction in evolutionary computation”, IEEE Transactions on Systems, Man, and Cybernetics, Part C – Applications and Reviews, 2005
  - Special issue on “Evolutionary optimization in the presence of uncertainties”, IEEE Transactions on Evolutionary Computation (to appear)
  - Special issue on “Evolutionary computation in dynamic and uncertain environments”, Genetic Programming and Evolvable Machines (submission deadline: [details])
Acknowledgements

Yaochu Jin would like to thank his colleagues in the ELTEC group at Honda Research Institute Europe,
Bernhard Sendhoff
Markus Olhoefer
Martina Haensjöger
Stefan Menzel
former collaborator,
and former students
Michael Hülsken
Ingo Paeske
Lars Gräning
Khaled Rasheed would like to thank his students in the Computer Science Dept. at the University of Georgia,
Deepti Chadékar
Liang Shi