The Extended Breeder Genetic Algorithm
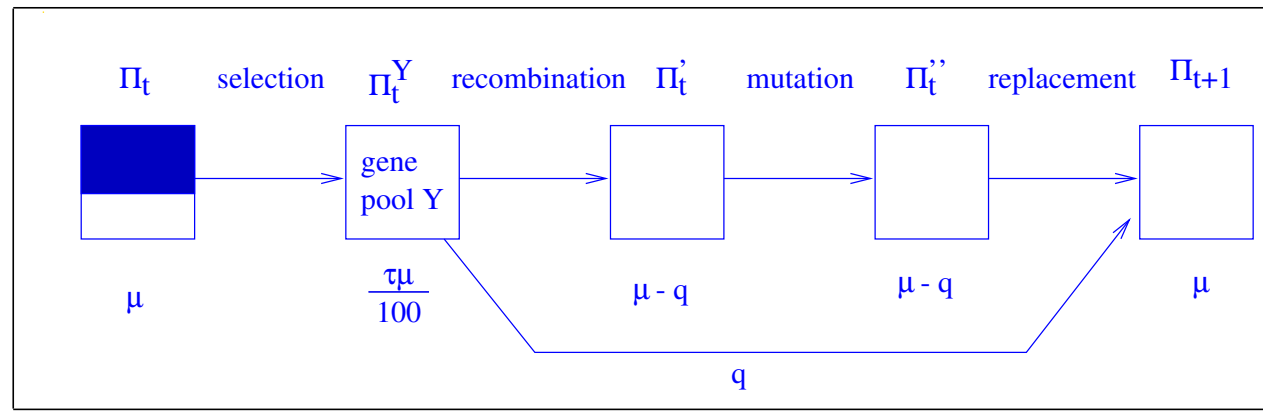
Lluís A. Belanche, René Alquézar

Soft Computing Research Group

LSI Department – UPC
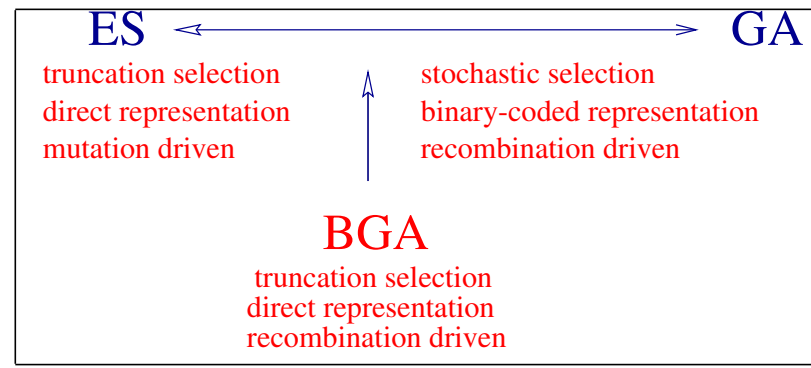
ATCI course, March 2014

The Breeder Genetic Algorithm

A scheme of the BGA procedure. Each box represents the population at different stages in the process to form a new generation. Notation on top of the boxes names the population at that point (see text) and the label from box to box (above the arrows) denotes operator sequencing (from left to right). The expressions at the bottom of the boxes indicate the population size at each step. Note how the final population size μ is formed by summing its two incoming values.

The Breeder Genetic Algorithm

ES ⟷ GA

truncation selection   stochastic selection
direct representation   binary-coded representation
mutation driven   recombination driven

BGA
truncation selection
direct representation
recombination driven

The BGA lies in a midway between ESs and GAs.

The Breeder Genetic Algorithm

Recombination (I)

In BGAs, recombination is applied unconditionally.
Let $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{y} = (y_1, \ldots, y_n)$ be two selected individuals.

1. Discrete Recombination (DR): $z_i \in \{x_i, y_i\}$ chosen with equal prob.

   Geometric effect: Let $H(\boldsymbol{x}, \boldsymbol{y})$ be the smallest hyperectangle containing both $\boldsymbol{x}$ and $\boldsymbol{y}$. Then $\boldsymbol{z}$ is one of the corners of $H(\boldsymbol{x}, \boldsymbol{y})$.

2. Extended Line Recombination (ELR): $z_i = x_i + \alpha(y_i - x_i)$ with $\alpha \in [-\delta, 1+\delta]$ chosen with uniform probability and $\delta \geq 0$ (typical $\delta = 0{,}25$)

   Geometric effect: Let $r(\boldsymbol{x}, \boldsymbol{y})$ be the line containing both $\boldsymbol{x}$ and $\boldsymbol{y}$. Then $r(\boldsymbol{x}, \boldsymbol{y})$ contains $\boldsymbol{z}$. That is, the resulting point lies in the line passing through $\boldsymbol{x}$ and $\boldsymbol{y}$. If $\delta = 0$ then the resulting point specifically lies in between $\boldsymbol{x}$ and $\boldsymbol{y}$.

3. Extended Intermediate Recombination (EIR) $z_i = x_i + \alpha_i(y_i - x_i)$ with $\alpha_i \in [-\delta, 1+\delta]$ chosen with uniform probability (same as ELR but a new α is chosen for each i).

   Geometric effect: Let $H(\boldsymbol{x}, \boldsymbol{y})$ be the smallest hyperectangle containing both $\boldsymbol{x}$ and $\boldsymbol{y}$. Then, for $\delta = 0$, $\boldsymbol{z} \in H(\boldsymbol{x}, \boldsymbol{y})$. For $\delta > 0$ the resulting point can also lie anywhere in the outside vicinity of $H(\boldsymbol{x}, \boldsymbol{y})$, further away the greater δ is.

The Breeder Genetic Algorithm

Recombination (II)

- Initially, EIR and ELR were not extended and yielded the parents's midpoint: $\alpha = \alpha_i = 0{,}5$

- The extension given by $\alpha_i \in [-\delta, 1 + \delta]$ such that $\delta \geq 0$ has been proven to be very useful, allowing more variety and increasing the variance of the operator.

The Breeder Genetic Algorithm

Fuzzy Recombination (FR)(I)

- This operator replaces the uniform *pdf* by a bimodal one, where the two modes are located at x_i and y_i, the two parents.

- It thus favours offspring values close to them, and not in any intermediate point with equal probability, as with previous operators.

- The label "fuzzy" comes from the fact that both parts of the pdf resemble fuzzy numbers (triangular in the original formulation)

- This idea need not be limited to the original triangular numbers.

- The geometric effect is in this case clear from the figures.

The Breeder Genetic Algorithm

Fuzzy Recombination (FR)(II)
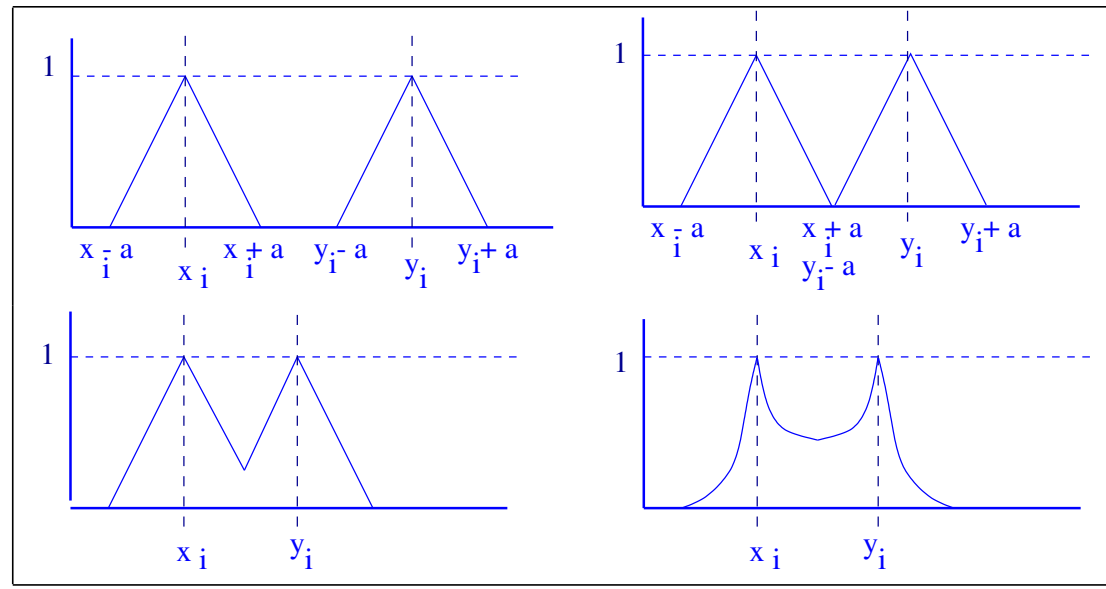
Bimodal probability distribution for the FR operator, where $a = e|y_i - x_i|$.

(a) Not all the intermediate values are possible ($e < 0{,}5$).

(b) Original formulation with $e = 0{,}5$.

(c) Overlapped distribution $0{,}5 < e \leq 1{,}0$.

(d) A general "fuzzy" symmetric probability distribution.

The Breeder Genetic Algorithm

Mutation (I)

Mutation is applied to each gene with some probability $\Pr(\Psi) = 1/n$ so that, on average, one gene is mutated for each individual.

Discrete Mutation (DM): $z_i = x_i + sign \cdot range_i \cdot \delta$

with $sign \in \{-1, +1\}$ chosen with equal prob., $range_i = \rho(r_i^+ - r_i^-)$, $\rho \in [0{,}1, 0{,}5]$ and

$$\delta = \sum_{i=0}^{k-1} \varphi_i 2^{-i}$$

where $\varphi_i \in \{0, 1\}$ from a Bernouilli pdf such that $\Pr(\varphi_i = 1) = 1/k$.

- $k \in \mathbb{N}^+$ is a parameter originally related to the *precision* with which the optimum was to be located, a machine-dependent constant.

- In practice, k is related to the *expected* value of mutation steps: the higher k is, the more fine-grained is the resultant mutation operator.

- The expected value of δ for a given k is that of a Binomial $(k, 1/k)$:

$$E\{\delta\}_{(k)} = \frac{1}{k} \sum_{i=0}^{k-1} 2^{-i} = \frac{1}{k}(2 - 2^{-k+1}) < \frac{2}{k}$$

- It can be proven that $E\{\delta\}_{(k')} < E\{\delta\}_{(k)}$ for $k' > k$ and thus expected mutation steps are lower.

- The factor ρ is the *range ratio*, related to the *maximum* step that mutation is allowed to produce as a ratio of variable range.

The Breeder Genetic Algorithm

Mutation (II)

Continuous Mutation (CM) Same as DM but with

$$\delta = 2^{-k\beta}$$

where $\beta \in [0, 1]$ with uniform prob. and $E\{\delta\}_{(k)} = \sqrt{2^{-k}}$

Again, $E\{\delta\}_{(k')} < E\{\delta\}_{(k)}$ for $k' > k$ and expected mutation is lower.

Extension to heterogeneous problems

- The algorithm manipulates the variables as a unique entity at all levels.

- Obviously, **real-valued** variables are directly treated as such, initialized at random within a pre-declared range, and recombined and mutated with the operators already described.

- Extension:

  - Ordinal variables

  - Nominal variables

  - Fuzzy variables

  - Missing values

Extension to heterogeneous problems

Ordinal variables

- Ordinal m-valued variables are represented as positive natural numbers in the interval $[1, m]$ an initialized at random within the interval.

- For recombination, there are three possibilities, which mimic the real-valued operators: DR (generally valid but ignores the order), LR (respects the order), and EIR (idem, needs an δ parameter).

- Some preliminary investigations lead to the choice of LR ($\alpha = 0,5$), that is, the *median* of the parents.

- Mutation involves an *increase* (to the immediately following value w.r.t. the linear order) or a *decrease* (idem, but in the opposite sense), and the decision is taken with equal probability.

# Extension to heterogeneous problems

Nominal variables

- Nominal m-valued variables are also represented as an interval $[1, m]$, but no order relation is assumed.

- The clear choice for recombination is DR, being the only one explicitly assuming no underlying order.

- Mutation is realized by *switching* to a new value in the interval, with equal probability.

Extension to heterogeneous problems

Fuzzy variables (I)

- The extension to handle *fuzzy numbers* is given by a tuple of reals (three in the general case, two if the chosen representation is symmetric).

- Recombination is taken as the corresponding extension of the operators for real-valued quantities. In particular, for EIR the mode is obtained following its formula (involving the selection of δ), and the spread is computed using the same formula with the *same α*.
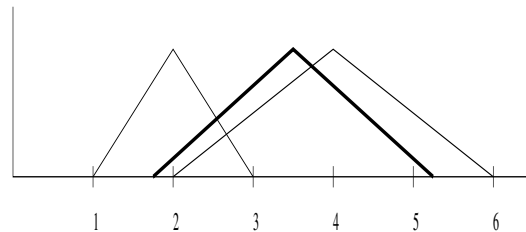
EIR recombination for fuzzy numbers with $\delta = 0{,}25$, and $\alpha = 0{,}75$ uniformly chosen in $[-0{,}25, 1{,}25]$. Mode and spread for the two parents are 2.0, 1.0 and 4.0, 2.0. The thicker number is the result of recombination. As for real numbers, the value of α makes offspring resemble its bigger parent more (a factor of $\frac{3}{4}$) than its smaller one. The mode is 3.5 and the spread 1.75.

Extension to heterogeneous problems

Fuzzy variables (II)

- Mutation of fuzzy numbers is also developed as an extension of the real-valued operators, by taking into account that mode and spread are collectively expressing a single (fuzzy) number.

- Both continuous and discrete operators can be used, as follows:

  1. The change on the mode is determined using the respective formulas.

  2. The change on the spread uses the *same* sign and δ (which are the terms depending on probabilities) as used for the mode.

Extension to heterogeneous problems

Missing values (I)

- Missing values are dealt with in a specially meaningful way.

- They are initially generated according to the estimated probability of a missing value in the variable. This makes sense since for variables containing high numbers of missing values, the probability of placing one in the corresponding gene increases.

- If this probability is zero a missing value could still be introduced by mutation (signaling the temporal loss of a gene or trait).

- A mutation operator sets a missing value in the allele with a certain probability (usually very low). If this change leads to improved performance, it will be retained.

- A missing value cannot be mutated back to a non-missing one. A definite value can only be recovered by recombination to the (non-missing) gene of another individual.

Extension to heterogeneous problems

Missing values (II)

- Recombination is treated as *discrete* (DR) whenever at least one of the parents have a missing trait. This is coherent with the philosophy of EA: recombination stands for the *transmission* of the parents's genetic material to their offspring.

- If a parent is lacking a gene, this characteristic has to be given the chance to be passed on. Besides, if the trait or gene is lacking for both parents, it will be so for the offspring, since nothing can be "invented from scratch" (this is the role of mutation).

Given Ω a recombination operator (possibly heterogeneous), it is extended to a $\Omega_{\mathcal{X}}$ (where $\mathcal{X}$ denotes the missing value) as:

$$\Omega_{\mathcal{X}}(x_i, y_i) = \begin{cases} \Omega(x_i, y_i) & \text{if } x_i \neq \mathcal{X} \ \& \ y_i \neq \mathcal{X} \\ DR(x_i, y_i) & \text{if } x_i = \mathcal{X} \ XOR \ y_i = \mathcal{X} \\ \mathcal{X} & \text{otherwise} \end{cases}$$

This manipulation differs from treating a missing value as any other value, for its generation and propagation would be carried out blindly. The proposed treatment has the added appeal of being simple, and natural from the point of view of an EA in the sense that it is taken as a missing *gene* and is independent of the data type.