Genetic Algorithms

Parameter Control

- Motivation
- Parameter setting
  - Tuning
  - Control
- Examples
- Where to apply parameter control
- How to apply parameter control

Motivation

An EA has many parameters that affect the search, e.g.
- mutation operator and mutation rate
- crossover operator and crossover rate
- selection mechanism and selective pressure (e.g. tournament size)
- population size

Good parameter values facilitate good performance

Q1 How to find good parameter values?

Motivation

EA parameters are rigid (constant during a run)

BUT

an EA is a dynamic, adaptive process

THUS

optimal parameter values may vary during a run

Q2: How to vary parameter values?
Parameter Setting

**Parameter Settings:**

**Tuning**
- The traditional way of testing and comparing different values before the "real" run

**Problems:**
- Users’ mistakes in settings can be sources of errors or sub-optimal performance
- Parameters interact: exhaustive search is not practicable
- Costs much time even with "smart" tuning
- Good values may become bad during the run

---

**Parameter Settings:**

**Control**
- Setting values on-line, during the actual run, e.g.
  - Predetermined time-varying schedule $p = p(t)$
    - Finding optimal $p$ is hard, finding optimal $p(t)$ is harder
  - Using feedback from the search process
    - Still user-defined feedback mechanism, how to "optimize"?
  - Encoding parameters in chromosomes and rely on selection
    - Will natural selection work for strategy parameters?
    - How to implement effectively?

---

**Examples:**

**Varying mutation step size**
- Problem to solve:
  - $\min f(x_1, \ldots, x_n)$
  - $L_i \leq x_i \leq U_i$ for $i = 1, \ldots, n$ bounds
  - $g_j(x) \leq 0$ for $j = 1, \ldots, q$ inequality constraints
  - $h_k(x) = 0$ for $k = q+1, \ldots, m$ equality constraints

- Algorithm:
  - EA with real-valued representation $x = (x_1, \ldots, x_n)$
  - Arithmetic averaging crossover
  - Gaussian mutation: $x_i' = x_i + N(0, \sigma)$
  - Standard deviation $\sigma$ is called mutation step size
Examples: Varying mutation step size, option 1

Replace the constant \( \sigma \) by a function \( \sigma(t) \)

\[
\sigma(t) = 1 - 0.9 \times \frac{t}{T}
\]

\( 0 \leq t \leq T \) is the current generation number

- Characteristics:
  - changes in \( \sigma \) are independent from the search progress
  - strong user control of \( \sigma \) by the above formula
  - \( \sigma \) is fully predictable
  - a given \( \sigma \) acts on all individuals of the population

\[ \sigma(i-n)/c \quad \text{if} \quad p > 0.2 \]
\[ \sigma(i-n)/c \quad \text{if} \quad p < 0.2 \quad 0 < c < 1 \]
\[ \sigma(i-n) \quad \text{otherwise} \]

---

Examples: Varying mutation step size, option 2

Replace the constant \( \sigma \) by a function \( \sigma(t) \) updated after every \( n \) steps by the 1/5 success rule:

1/5 success rule (Rechenberg 1973):
1/5 of mutations should be successful – mutant more fit than parent

Examples: Varying mutation step size, option 3

- Assign a personal \( \sigma \) to each individual
- Incorporate this \( \sigma \) into the chromosome: \((x_1, \ldots, x_n, \sigma)\)
- Apply variation operators to \( x \)'s and \( \sigma \)

\[ \sigma' = \sigma \times e^{N(0,\sigma)} \]
\[ x_i' = x_i + N(0,\sigma') \]

- Characteristics:
  - changes in \( \sigma \) are results of natural selection
  - (almost) no user control of \( \sigma \)
  - \( \sigma \) is not predictable
  - a given \( \sigma \) acts on one individual

---

Examples: Varying mutation step size, option 2

- Assign a personal \( \sigma \) to each individual
- Incorporate this \( \sigma \) into the chromosome: \((x_1, \ldots, x_n, \sigma)\)
- Apply variation operators to \( x \)'s and \( \sigma \)

\[ \sigma' = \sigma \times e^{N(0,\sigma)} \]
\[ x_i' = x_i + N(0,\sigma') \]

- Characteristics:
  - changes in \( \sigma \) are results of natural selection
  - (almost) no user control of \( \sigma \)
  - \( \sigma \) is not predictable
  - a given \( \sigma \) acts on one individual
Examples:
Varying mutation step size, option 4

Assign a personal \( \sigma \) to each variable in each individual
Incorporate \( \sigma \)'s into the chromosomes: \((x_1, \ldots, x_n, \sigma_1, \ldots, \sigma_n)\)
Apply variation operators to \( x_i \)'s and \( \sigma_i \)'s

\[
\sigma_i' = \sigma_i e^{N(0, \tau)}
\]
\[
x_i' = x_i + N(0, \sigma_i')
\]

- Characteristics:
  - changes in \( \sigma_i \) are results of natural selection
  - (almost) no user control of \( \sigma_i \)
  - \( \sigma_i \) is not predictable
  - a given \( \sigma_i \) acts on one gene of one individual

Examples:
Varying penalties

Constraints
- \( g_j(x) \leq 0 \) for \( j = 1, \ldots, q \) inequality constraints
- \( h_k(x) = 0 \) for \( k = q+1, \ldots, m \) equality constraints
are handled by penalties:

\[
eval(x) = f(x) + W \times \text{penalty}(x)
\]

where

\[
\text{penalty}(x) = \sum_{j=1}^{m} \begin{cases} 1 & \text{for violated constraint} \\ 0 & \text{for satisfied constraint} \end{cases}
\]

Examples:
Varying penalties, option 1

Replace the constant \( W \) by a function \( W(t) \)

\[
W(t) = (C \times t)^b
\]

\( 0 \leq t \leq T \) is the current generation number

- Characteristics:
  - changes in \( W \) independent from the search progress
  - strong user control of \( W \) by the above formula
  - \( W \) is fully predictable
  - a given \( W \) acts on all individuals of the population

Examples:
Varying penalties, option 2

Replace the constant \( W \) by \( W(t) \) updated in each generation

\[
W(t+1) = \begin{cases} 
\beta \times W(t) & \text{if last } k \text{ champions all feasible} \\
\gamma \times W(t) & \text{if last } k \text{ champions all infeasible} \\
W(t) & \text{otherwise}
\end{cases}
\]

\( \beta < 1, \gamma > 1, \beta \times \gamma \neq 1 \) champion: best of its generation

- Characteristics:
  - changes in \( W \) are based on feedback from the search progress
  - some user control of \( W \) by the above formula
  - \( W \) is not predictable
  - a given \( W \) acts on all individuals of the population
Examples: Varying penalties, option 3

Assign a personal $W$ to each individual in population
Incorporate this $W$ into the chromosome: $(x_1, \ldots, x_n, W)$
Apply variation operators to $W$ and each $x_i$

Alert:
\[
\text{eval}((x, W)) = f(x) + W \times \text{penalty}(x)
\]

while for mutation step sizes we had
\[
\text{eval}((x, s)) = f(x)
\]
this option is thus “cheating” \Rightarrow algorithm can improve the
evaluation by evolving smaller weights $W$ rather than
improving $f(x)$

Examples: Lessons learned

Various forms of parameter control can be distinguished by:

- primary features:
  - what component of the EA is changed
  - how the change is made

- secondary features:
  - evidence/data backing up changes
  - level/scope of change

Where to apply parameter control

Practically any EA component can be parameterized and
thus controlled on-the-fly:
- representation
- evaluation function
- variation operators
- selection operator (parent or mating selection)
- replacement operator (survival or environmental
  selection)
- population (size, topology)
How to apply parameter control

Three major types of parameter control:

- **deterministic**: some rule modifies strategy parameter without feedback from the search (based on some counter)
- **adaptive**: feedback rule based on some measure monitoring search progress
- **self-adaptive**: parameter values evolve along with solutions; encoded onto chromosomes they undergo variation and selection

Evidence: Informing the change

The parameter changes may be based on:

- **time or nr. of evaluations** (deterministic control)
- **population statistics** (adaptive control):
  - progress made
  - population diversity
  - gene distribution, etc.
- **relative fitness** of individuals created with given values (adaptive or self-adaptive control)

Evidence: Informing the change

- **Absolute evidence**: predefined event triggers change, e.g. increase $p_m$ by 10% if population diversity falls under threshold $x$
  - Direction and magnitude of change is fixed
- **Relative evidence**: compare values through solutions created with them, e.g. increase $p_m$ if top quality offspring came by high mutation rates
  - Direction and magnitude of change is not fixed
Evidence: Refined taxonomy

- Combinations of types and evidences
  - Possible: +
  - Impossible: -

<table>
<thead>
<tr>
<th></th>
<th>Deterministic</th>
<th>Adaptive</th>
<th>Self-adaptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Relative</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Scope/level

The parameter may take effect on different levels:
- environment (fitness function)
- population
- individual
- sub-individual

Note: given component (parameter) determines possibilities
Thus: scope/level is a derived or secondary feature in the classification scheme

Evaluation/Summary

- Parameter control offers the possibility to use appropriate values in various stages of the search
- Adaptive and self-adaptive parameter control
  - offer users “liberation” from parameter tuning
  - delegate parameter setting task to the evolutionary process
  - the latter implies a double task for an EA: problem solving + self-calibrating (overhead)