Landscapes and Hardness for Genetic Algorithms

Timo Latvala

Outline

- Genetic Algorithms
- Preliminaries
- Epistatic measures
- SWO and RSAO
- Other measures

Genetic Algorithms

function GenAlg ($popSize$, $maxIter$) $=$

Select an initial population P of with $popSize$ feasible solutions
$best$:= best solution of P;
for i := 1 to $maxIter$
  Construct a pairing of the elements in P; $Q = P$;
  for each pair w, x in the pairing
    $(y, z) := mate(w, x)$;
    $y := mutate(y); z := mutate(z)$;
    $Q := Q \cup \{y, z\}$;
  $P := select(Q, popSize)$;
  if$(f(max(P) > f(best))$
    $best := max(P)$;
return $best$

Genetic Algorithms

- Recombination is usually done using uniform crossover.
- Some times fitter parents can be biased to produce more children.
- Mutation is usually implemented as changing the value of a site with a small probability.
- Linear ranking selection vs proportional selection.

Idea

- Analyse and evaluate different measures of hardness
- The measures are mostly applied on easy problems.
- Some of the methods also relate to the fitness landscape the algorithm perceives.
- The performance is measured by comparing *convergence* in several ways.

Notation and Definitions

- Σ is a finite alphabet, $S = \Sigma^l$, the universe.

- A *schema* is a string in $\Sigma^{l'}$, where $\Sigma = \Sigma \cup \{\#\}$. $\#$ functions as a wild card symbol.

- $a^n = a \ldots a$, i.e. n a:s.

- $d(s, s'), s, s' \in S$ Hamming distance.

- A *first order* function (1ORD):

$$s \mapsto \sum_i g_i(s_i)$$

Notation and Definitions

- A function is *linear* if

$$s \mapsto C_1 - C_2 d(s, s^*),$$

where C_1 and C_2 are constant and s^* is the optimum configuration

- A function f is *monotone* (MON) if

$$d(s^1, s^*) < d(s^2, s^*) \Rightarrow f(s^1) > f(s^2)$$

- A fitness function f is unimodal (UNI) it has a unique optimum, which is the global optimum.

Epistatic Measures

- epistasis = "interaction between the sites in the expression of the fitness function"

- Any fitness function can be written in the form

$$
\begin{aligned}
f(s) &= c + \sum_{i=0}^{l-1} g_i(s_i) + \\
&+ \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} g_{ij}(s_i, s_j) + \cdots \\
&+ \text{random error}
\end{aligned}
$$

- $\bar{f}_P = \frac{1}{|P|} \sum_{t \in P} f(t)$, $\bar{f}_{P_i}(s) = \frac{1}{|\{t \in P; t_i = s_i|} \sum_{t \in P, t_i = s_i} f(t)$

Epistatic Measures

- $$\xi_P(s) = \sum_{i=1}^{l} (\bar{f}_{P_i}(s) - \bar{f}_P) + \bar{f}_P$$

- Epistasis variance:

$$epiv_P(f)^2 = \frac{\sum_{s \in P}(f(s) - \xi(s))^2}{\sum_{t \in P} f(t)^2}$$

- Epistasis correlation:

$$epic_P(f) = \frac{\sum_{s \in P}(f(s) - \bar{f}_P)(\xi(s) - \bar{\xi})}{\sqrt{\sum_{t \in P}(f(t) - \bar{f})^2}\sqrt{\sum_{t \in P}(\xi(t) - \bar{\xi})^2}}$$

Epistatic Measures: Discussion

- $epiv(f) = 0$ iff f is a first order function
- Epistasis correlation is translation invariant i.e. $epic(af + b) = epic(f)$.
- $epic(f) = 1$ for a subset of the first order functions.
- Epistasis variance returns 0 for both constant and other first order functions.
- Epistasis variance measures more the absensce of epistasis than its presence.

Epistatic Measures: Discussion

- Both epistatic measures are sensitive to non-linear scaling, due to their reference classes.
- GA are rarely sensitive to non-linear scaling.
- Epistatic measures does not distinguish between the signs of the interactions
- Epistatic measure can be generalised to measure higher order effects
- As a difficulty measure they fare poorly.

Fitness Distance Correlation

- Computes the correlation between a fitness function f and the distance to the global optimum s^*.

- fitness distance correlation

$$fdc_P(f) = \frac{\sum_{t \in P}(f(t) - \bar{f}_P(t))(d(t) - \bar{d})}{\sqrt{\sum_{t \in P}(f(t) - \bar{f})^2}\sqrt{\sum_{t \in P}(d(t) - \bar{d})^2}}$$

- $fdc(f) = -1$ iff f is a linear function

- Sensitive to non-linear scaling

- Detects constant functions.

Sitewise Optimisation Measure

- We generalise the strict monotone linearity condition for the reference class of the fdc.

- If one comes closer to the optimum the fitness has to increase.

- The measure is defined algorithmically.

- $(f \in SWO)$ if the output of the SWO algorithm is optimal regardless of the input string.

Sitewise Optimisation Measure

function SWO (function f, string s) $=$

$\quad$ **for each** $i \in L$

$\quad\quad$ $A_i = \{\alpha \in \Sigma |\ f(s[i|\alpha]) > f(s[i|\beta])$ for all $b \in \Sigma\}$

$\quad\quad$ **for each** $\beta \in \Sigma$

$\quad\quad\quad$ **if** $(s_i \in A_i)$

$\quad\quad\quad\quad$ $m_i = s_i$;

$\quad\quad\quad$ **else** $m_i :=$ arbitrary element of A_i;

$\quad$ **return** m;

- SWO-functions have a unique global optimum.

- They do not possess local constantness.

Sitewise Optimisation Measure

- Denote by $SWO(f,P)$ the set formed by applying the SWO-algorithm individually on the members of P

- $swo_P(f) = \frac{w(SWO(f,P))}{w(P)}$

- $w(P) = \frac{1}{P^2} \sum_{p,q \in P} d(p,q)$

- $swo(f) = 0$ iff f has a unique global optimum and is SWO.

- The measure is invariant to non-linear scaling.

Beyond Monotone Fitness Functions

- Steepest ascent optimisable (SAO)
- $SWO \subset SAO \subset UNI$.
- restricted SAO: a site can only be modified once
- $SWO \subset RSAO \subset UNI$.

Experimental Design Perspective

- Instead of just computing the effects of the first order interactions also the higher order terms are considered.

- Interaction does not necessarily affect convergence

- Already problems with second order effects can be NP-complete

- Total knowledge of the interactions is equivalent to knowing all the Walsh coefficients.

Metropolis Sampling

- Sample only the relevant part of the space.
- Give states a Boltzmann probability.
- Can be used to compute the *density of states*
- Hypothesis: problems with a fast decay in density should be hard to solve
- It is an invariant measure w.r.t. the landscape of the problem.
- No derivative work available.

On-Line Sampling

- Sample only the relevant space
- Use the states encountered during a given run
- The landscape directly affects which states are chosen.
- Run the same GA with a hamming fitness function $f(s) = l - d(s, s^*)$.
- Compute the ratio of the two measures.

Final Remarks

- Major flaws of the measures: sensitivity to non-linear fitness scaling, constantness and averaging.
- None of the measures can identify all easy cases for GAs.
- Monotonicity and unimodality are too small reference classes.
- All experiments have been performed with a specific GA - generality?

References

- L. Kallel, B. Naudts, C.R. Reeves, "Properties of fitness functions and search landscapes." Theoretical Aspects of Evolutionary Computing (L. Kallel, B. Naudts, A. Rogers, eds.) Springer-Verlag, Berlin Heidelberg 2001, pp. 175-206.

- B. Naudts, L. Kallel, "A comparison of predictive measures of problem difficulty in evolutionary algorithms." IEEE Transactions on Evolutionary Computation 4 (2000), pp 1-15.

- H.Rosé, W.Ebeling and T.Asselmeyer, "The density of states - A measure of difficulty of optimization problems". In Proc. 4th Conf. Parallel Problem Solving from Nature. Berlin, Germany, Springer-Verlag, 1996, pp 263–290.