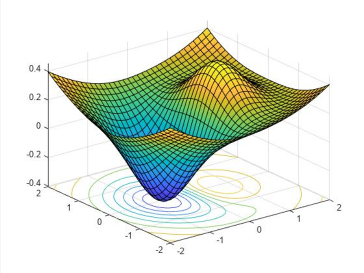
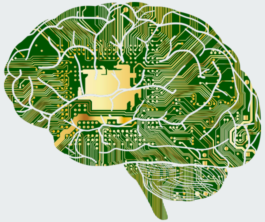
A Runtime Analysis for the Weighted Univariate Marginal Distribution Algorithm on LeadingOnes

by **Youssef Chaabouni**
Supervised by: **Benjamin Doerr** and **Martin S. Krejca**

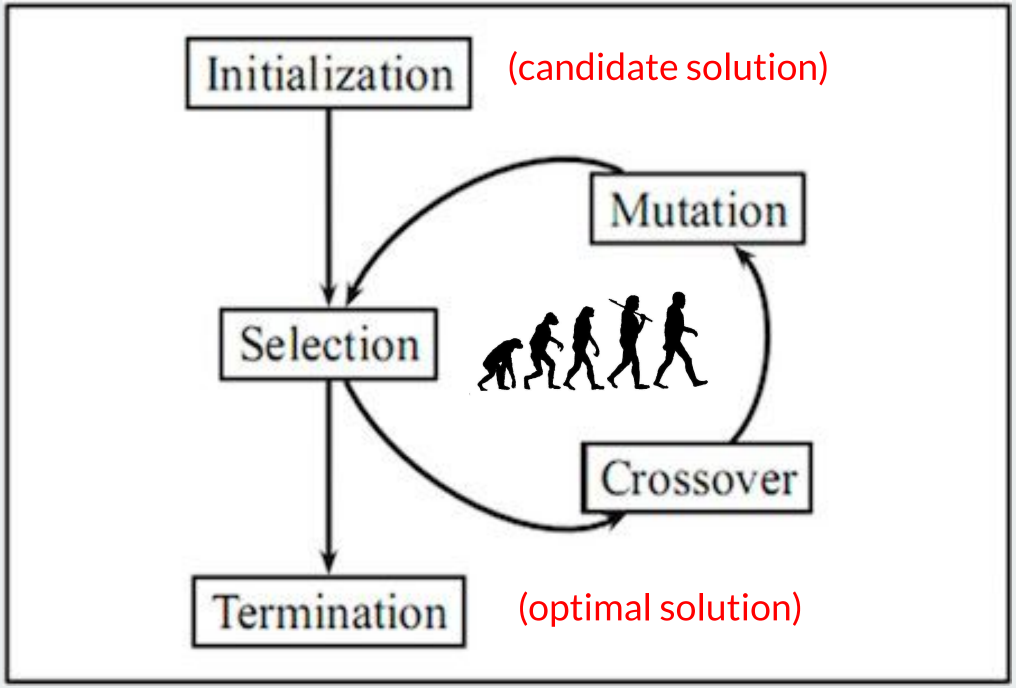
Evolutionary Algorithms: What Are They?



optimization problem

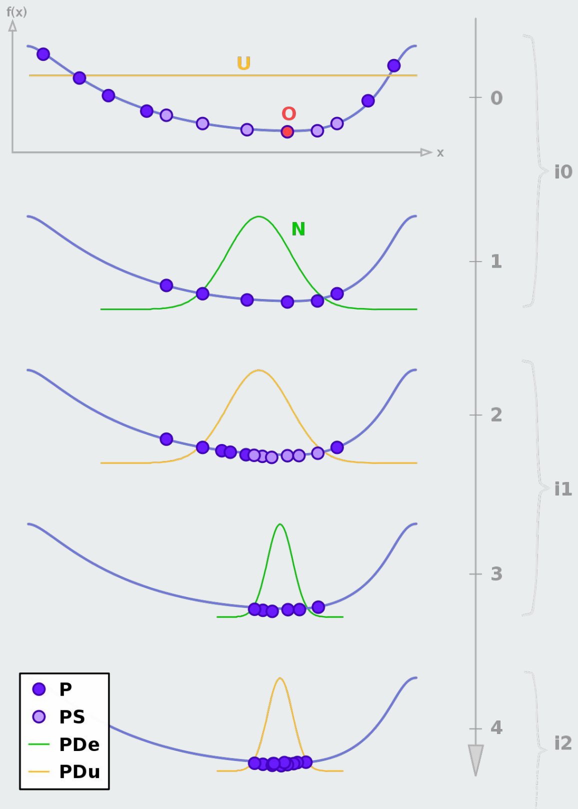


used in genetics & artificial intelligence



Estimation-of-Distribution Algorithms

probabilistic model ← uniform distribution
repeat
    sample a population
    select *best* individuals
    update probabilistic model based on selected individuals
until optimum is sampled;



The Weighted Univariate Marginal Distribution Algorithm

by Dufay and Doerr, 2022 [DD22]

Individuals:

bit-strings of length n

$$\{0,1\}^n$$

Problem:

Find the maximum of a pseudo-Boolean function

$$f : \{0,1\}^n \to \mathbb{R}$$

Algorithm 1 The weighted univariate marginal distribution algorithm (*weighted UMDA*) with parameters λ, μ and weights $\gamma_1 \geq \cdots \geq \gamma_\mu \geq \gamma_{\mu+1} = \cdots = \gamma_\lambda = 0$ such that $\sum_{i=1}^{\lambda} \gamma_i = 1$; optimizing a pseudo-Boolean function f.

$t \leftarrow 0$
$p^{(t)} \leftarrow (\frac{1}{2})_{i \in [n]}$
repeat $\triangleright$ *iteration t*
  for $i \in [\lambda]$ **do** $x^{(i)} \leftarrow$ individual sampled via $p^{(t)}$;
  let $y^{(1)}, \ldots, y^{(\mu)}$ denote the μ individuals out of $x^{(1)}, \ldots, x^{(\lambda)}$ with the best fitness (breaking ties uniformly at random)
  for $i \in [n]$ **do** $p_i^{(t+1)} \leftarrow \sum_{i=1}^{\mu} \gamma_i y_j^{(i)}$;
  restrict $p^{(t+1)}$ to the interval $[\frac{1}{n}, 1 - \frac{1}{n}]$
until termination criterion met;

The LeadingOnes Benchmark by Rudolph, 1997 [Rud97]

Objective function:

LeadingOnes(x) = length of the longest prefix of 1s of the bit-string x

$$\textsc{LeadingOnes}(x) = \sum_{i=1}^{n} \prod_{j=1}^{i} x_i$$

Theoretical Results

We provide a runtime upper bound for the **weighted UMDA** optimizing **LeadingOnes**.

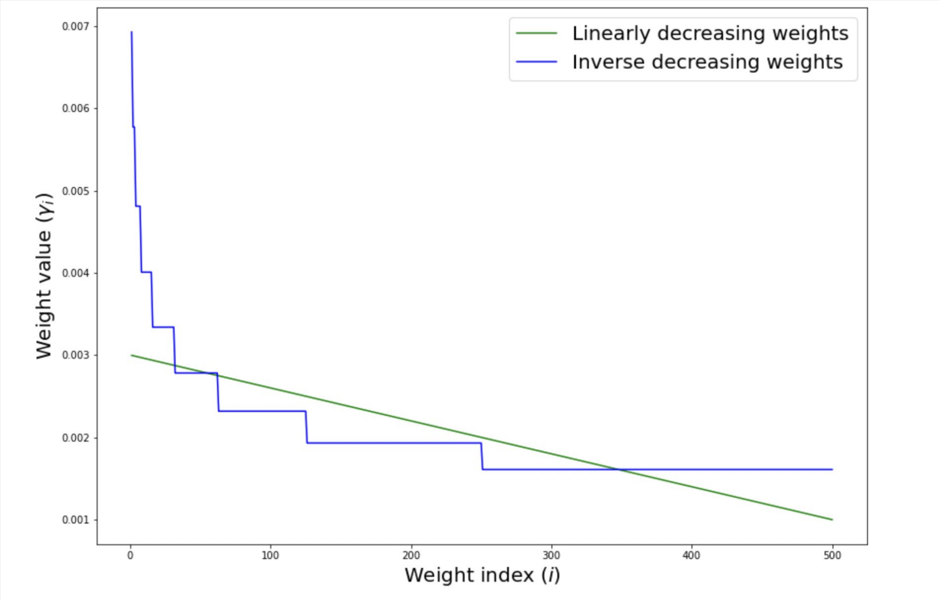
This is a generalization of the results of Doerr and Krejca about the **UMDA** (special case when all the weights are equal) optimizing **LeadingOnes** [DK21a].

It states that in the regime where *the inverse of the squared weights is at least quasilinear*, the runtime of **weighted UMDA** in objective function evaluations is *at most linear in the problem size divided by the logarithm of the selection rate.*

Theorem 3. *Let $\delta \in (0,1)$ be a constant, and let $\zeta = \frac{1-\delta}{4e}$. Consider the weighted UMDA optimizing* LeadingOnes *with $\sum_{i=1}^{\lambda} \gamma_i^2 \leq \frac{1}{128 n \ln n}$ and $\lambda \geq \frac{\mu}{\zeta}$. Further, let $d = \lfloor \log_4(\zeta \frac{\lambda}{\mu}) \rfloor$. Then the weighted UMDA samples the optimum after at most $\lambda(\lceil \frac{n}{d+1} \rceil + \lceil \frac{n}{n-1} e \ln n \rceil)$ fitness function evaluations with a probability of at least $1 - 5n^{-1}$.*

Experimental Settings

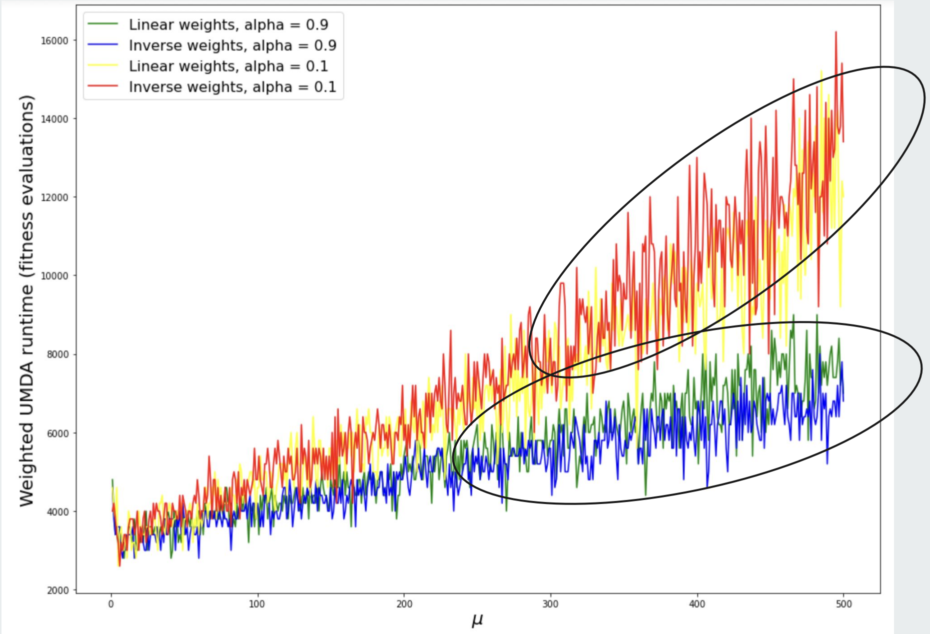
We design two weight distribution models (see Figure) and normalize them such that <u>the sum of the squared weights is the same</u>.



Experimental Results

We run experiments on the **weighted UMDA** optimizing LeadingOnes.

We observe that weight distributions with the same sum of squared weights behave similarly.



We conjecture that the sum of the squared weights characterizes the behavior of the **weighted UMDA** by determining the force that pulls the frequencies of the probabilistic model to the extremities 0 and 1 (called *genetic drift*).

Conclusion

We generalize the upper bound for **UMDA** to the more general version of the algorithm, the **weighted UMDA**.

The fact that the condition for which the upper bound holds is only affected by the weight distribution through the sum of the squared weights motivates an experimental analysis.

We conjecture that the sum of the squared weights characterizes the behavior of the **weighted UMDA**.

Thank you!

References

[DD22] Benjamin Doerr and Mark Dufay. General univariate estimation- of-distribution algorithms. In Günter Rudolph, Anna V. Kononova, Hernán E. Aguirre, Pascal Kerschke, Gabriela Ochoa, and Tea Tusar, editors, *Parallel Problem Solving From Nature, PPSN 2022*, pages 470–484. Springer, 2022.

[DK21a] Benjamin Doerr and Martin S. Krejca. A simplified run time analysis of the univariate marginal distribution algorithm on LeadingOnes. *Theoretical Computer Science*, 851:121–128, 2021.

[Rud97] Günter Rudolph. *Convergence Properties of Evolutionary Algorithms*. Verlag Dr. Kovac, 1997.