Lecture 21 - Introductory Theoretical Analysis of Evolutionary Algorithms 2

Introduction to Evolutionary Computation

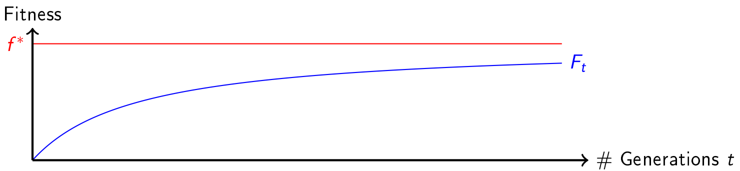
Per Kristian Lehre p.k.lehre@cs.bham.ac.uk

CERCIA, School of Computer Science

December 11, 2009

Outline

Convergence of Evolutionary Algorithms



Definition ([Rudolph, 1998])

An EA converges with probability 1 on fitness function $f : S \to \mathbb{R}$ if

$$\mathbf{Pr}\left[\lim_{t\to\infty} f^* - F_t = 0\right] = 1$$

where

$f^* := \max_{x \in S} f(x)$	Maximal value of fitness function f.
$F_t := \max_{x \in P_t} f(x)$	Maximal fitness among the individuals in population P_t at generation t.

Convergence Conditions

Two sufficient conditions[1] for convergence with probability 1

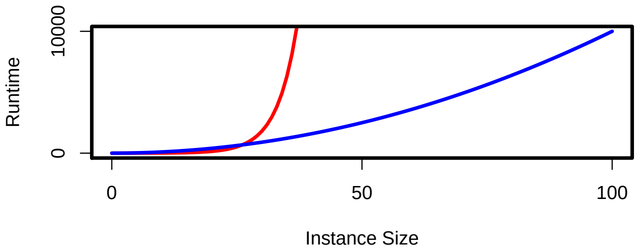
1. Elitism, i.e. $\forall t \geq 0 \;\; F_{t+1} \geq F_t$, and
2. Global mutation operator, i.e.

$$\forall x, y \in S \quad \mathbf{Pr}\left[x = \text{mut}(y)\right] \geq \delta_m > 0.$$

$\rightarrow$ Convergence easy to achieve.
How about time to reach optimum?

[1] See [Rudolph, 1998] for more convergence conditions.

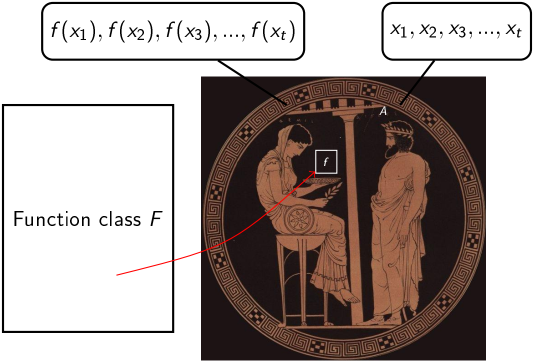
Computational Complexity of EAs



Prediction of resources needed for a given instance.
Usually *runtime* as function of *instance size*.

- Exponential runtime $\Longrightarrow$ Inefficient algorithm.
- Polynomial runtime $\Longrightarrow$ "Efficient" algorithm.

The Black Box Model

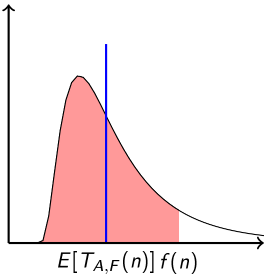


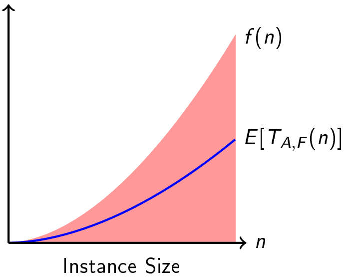
- **Worst case** runtime of algorithm A on function class F

$$T_{A,F} := \max_{f \in F} T_{A,f}$$

[Droste et al., 2006]

Expected Runtime and Success Probability





The **runtime** $T_{A,F}(n)$ is a **random variable** over the outcomes of the random number generator.

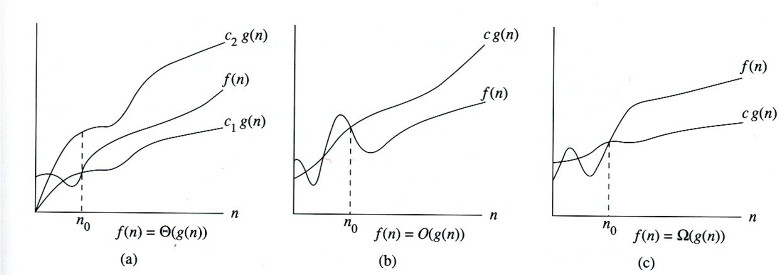
- **Expected Runtime**

$$\mathbf{E}\left[T_{A,F}(n)\right] = \sum_{t=1}^{\infty} t \cdot \mathbf{Pr}\left[T_{A,F}(n) = t\right].$$

- **Success Probability** within $f(n)$ steps

$$\mathbf{Pr}\left[T_{A,F}(n) \leq f(n)\right].$$

Asymptotic Notation



$f(n) \in O(g(n)) \iff$ there exist positive constants c and n_0 such that $0 \leq f(n) \leq c \cdot g(n)$ for all $n \geq n_0$.

$f(n) \in \Omega(g(n)) \iff g(n) \in O(f(n))$

$f(n) \in \Theta(g(n)) \iff f(n) \in O(g(n))$ and $f(n) \in \Omega(g(n))$

[Cormen et al., 1990]

(1+1) Evolutionary Algorithm

Algorithm 1 (1+1) EA

Sample x uniformly at random from $\{0,1\}^n$.
repeat
  $x' \leftarrow x$.
  Flip each bit of x' with probability $1/n$.
  if $f(x') \geq f(x)$ **then**
    $x \leftarrow x'$.
  end if
until termination condition met.

- Special case of the $(\mu+\lambda)$ EA.
- Starting point for rigorous runtime analysis of EAs, e.g.
  - [Mühlenbein, 1992]
  - [Garnier et al., 1999]
  - **[Droste et al., 2002]**

OneMax, a starting point for runtime analysis

Theorem

The expected runtime of (1+1) EA on OneMax *is* $\Theta(n \ln n)$,

$$\textsc{OneMax}(x) := x_1 + x_2 + \cdots + x_n.$$

(Upper bound proved in [Mühlenbein, 1992].)

Proof ingredients (upper bound).

- Linearity of Expectation $\mathbf{E}[X+Y] = \mathbf{E}[X] + \mathbf{E}[Y]$
- Geometric Distribution
- Artificial Fitness Levels
- n-th Harmonic Number

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = O(\ln n)$$

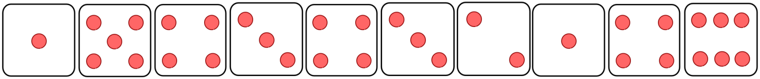
Geometric Distribution

Geometric Distribution

Number of indep. trials X until success, with success probability p.

$$\mathbf{Pr}[X = k] = (1-p)^{k-1} \cdot p \quad \text{for } k \in \{1, 2, 3, \ldots\}.$$

Eg. # dice throws until 6 pips is geom. distr. with $p = 1/6$.



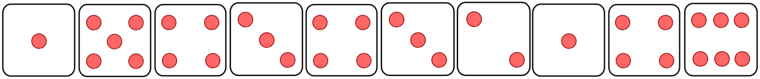
Geometric Distribution

Geometric Distribution

Number of indep. trials X until success, with success probability p.

$$\mathbf{Pr}[X = k] = (1-p)^{k-1} \cdot p \quad \text{for } k \in \{1, 2, 3, \ldots\}.$$

Eg. # dice throws until 6 pips is geom. distr. with $p = 1/6$.

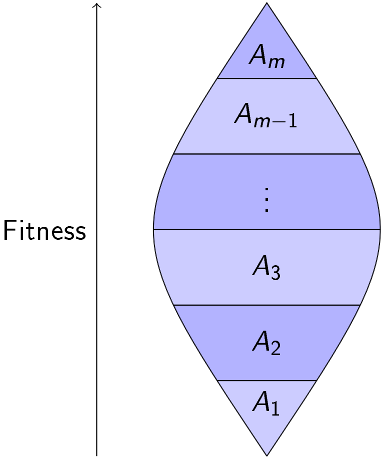


A geometrically distributed variable has expectation $\mathbf{E}[X] = 1/p$.

$$\begin{aligned} \mathbf{E}[X] &= p \cdot 1 + (1-p) \cdot (1 + \mathbf{E}[X]) \\ \mathbf{E}[X] \cdot (1 - (1-p)) &= p + (1-p) \\ \mathbf{E}[X] &= 1/p. \end{aligned}$$

Artificial Fitness Levels - Upper bounds

Search space partitioned into m subsets $A_1, A_2, ..., A_m$, with increasing fitness, *i.e.* $f(A_i) < f(A_j)$ for all $i < j$, and $f(A_m)$ optimal.



p_i : Probability of jumping from A_i to any A_j, $i < j$.

T_i : Time to jump from A_i to any A_j, $i < j$.

Expected runtime

$$
\begin{aligned}
\mathbf{E}[T] &\leq \mathbf{E}[T_1 + T_2 + \cdots + T_m] \\
&= \mathbf{E}[T_1] + \mathbf{E}[T_2] + \cdots + \mathbf{E}[T_m] \\
&= 1/p_1 + 1/p_2 + \cdots 1/p_m
\end{aligned}
$$

Artificial Fitness Levels - Upper bound on OneMax

Partitioning of search space in fitness levels

$$\textsc{OneMax}(x) := x_1 + x_2 + \cdots + x_n$$

A_i : all bitstrings with i 1-bits.

p_i : probability of increasing number of 1-bits, from within A_i
(at least prob. of flipping one 0-bit, and no other bits)

$$p_{n-i} \geq i \cdot \frac{1}{n} \cdot \underbrace{\left(1 - \frac{1}{n}\right)^{n-1}}_{\geq 1/e} \geq \frac{i}{en}$$

Artificial Fitness Levels - Upper bound on OneMax

Partitioning of search space in fitness levels

$$\textsc{OneMax}(x) := x_1 + x_2 + \cdots + x_n$$

A_i : all bitstrings with i 1-bits.

p_i : probability of increasing number of 1-bits, from within A_i (at least prob. of flipping one 0-bit, and no other bits)

$$p_{n-i} \geq i \cdot \frac{1}{n} \cdot \underbrace{\left(1 - \frac{1}{n}\right)^{n-1}}_{\geq 1/e} \geq \frac{i}{en}$$

Expected runtime

$$\mathbf{E}\left[T_{\textsc{OneMax}}\right] \leq \sum_{i=1}^{n} \frac{1}{p_{n-i}} \leq \sum_{i=1}^{n} \frac{en}{i} = O(n \ln n)$$

Artificial Fitness Levels - Exercise

Exercise

Estimate an upper bound on the expected runtime of (1+1) EA on

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

$$x = \overbrace{1111111111111111}^{\text{Leading 1-bits.}} \; \underbrace{0}_{\text{First 0-bit.}} \; \overbrace{****************}^{\text{Random bitstring.}} .$$

(1+1) EA on Linear Functions

Definition

A function $f : \{0,1\}^n \to \mathbb{R}$ is *linear* iff there exist $w_i, c \in \mathbb{R}$ s.t.

$$f(x) = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + c$$

Theorem ([Droste et al., 2002, He and Yao, 2003])

The expected runtime of (1+1) EA on the class of linear functions with all non-zero weights is $\Theta(n \log n)$.

→ All linear functions are easy for the (1+1) EA.

Enforcing Expected Runtimes

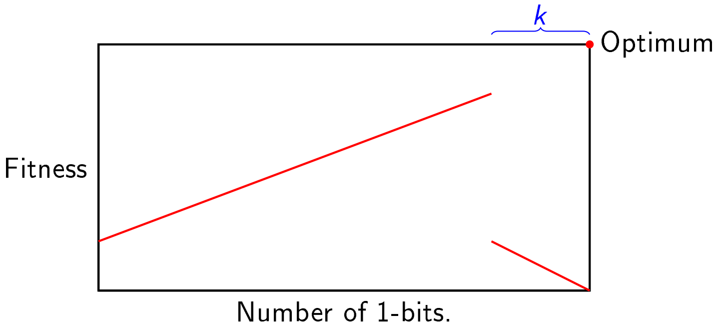


Figure: Illustration of $\textsc{Jump}_k$.

Theorem ([Droste et al., 2002])

The expected runtime of (1+1) EA on $\textsc{Jump}_k$ *is* $\Theta(n^k + n \log n)$ *for* $k \in \{1, 2, ..., n\}$.

$\rightarrow$ Difficulty of problem determined by size of the gap k.

When Offspring Populations are Helpful

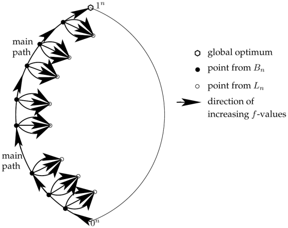


Figure: Illustration of SufSamp.

Theorem ([Jansen et al., 2005])

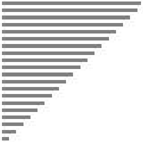
The success probabilities of (1+λ) EA on SufSamp are

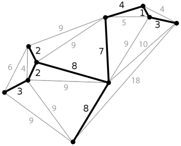
For $\lambda = 1$: Within $n^{O(1)}$ with prob. $2^{-\Omega(\sqrt{n})}$.

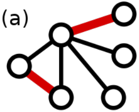
For $\lambda \geq cn \log n$: Within $O(\lambda n)$ with prob. $\geq 1 - \varepsilon$, for any $\varepsilon > 0$.

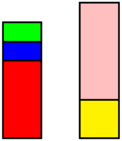
→ Increasing the offspring population size
can decrease the runtime from exponential to polynomial.

(1+1) EA in Combinatorial Optimisation









Sorting	$\Theta(n^2 \log n)$	[Scharnow et al., 2002]
MST	$\Theta(m^2 \log(nw_{max}))$	[Neumann and Wegener, 2007]
Max. Matching	$e^{\Omega(n)}$	[Giel and Wegener, 2003]
	$(1+\varepsilon)$-approx in $O(m^{\lceil 2/\varepsilon \rceil})$	
Partition	$4/3$-approx. in $O(n^2)$	[Witt, 2005]

More results in survey [Oliveto et al., 2007].

Summary

Convergence of EAs

- Optimum found in finite time?
- Two simple conditions sufficient.

Computational Complexity of EAs

- Expected Runtime and Success Probability
- Artificial Fitness Levels
- (1+1) EA optimises linear functions in $\Theta(n \log n)$.
- (1+1) EA on classical combinatorial optimisation problems.

Compulsory Reading

1. Section 1 and Section 2 in [Rudolph, 1998].
2. [Oliveto et al., 2007].

References I

Cormen, T. H., Leiserson, C. E., and Rivest, R. L. (1990).
Introduction to Algorithms.
MIT Press.

Droste, S., Jansen, T., and Wegener, I. (2002).
On the analysis of the (1+1) Evolutionary Algorithm.
Theoretical Computer Science, 276:51–81.

Droste, S., Jansen, T., and Wegener, I. (2006).
Upper and lower bounds for randomized search heuristics in black-box optimization.
Theory of Computing Systems, 39(4):525–544.

Garnier, J., Kallel, L., and Schoenauer, M. (1999).
Rigorous hitting times for binary mutations.
Evolutionary Computation, 7(2):173–203.

Giel, O. and Wegener, I. (2003).
Evolutionary algorithms and the maximum matching problem.
In *Proceedings of the 20th Annual Symposium on Theoretical Aspects of Computer Science (STACS 2003)*, pages 415–426.

He, J. and Yao, X. (2003).
Towards an analytic framework for analysing the computation time of evolutionary algorithms.
Artificial Intelligence, 145(1-2):59–97.

References II

Jansen, T., Jong, K. A. D., and Wegener, I. (2005).
On the choice of the offspring population size in evolutionary algorithms.
Evolutionary Computation, 13(4):413–440.

Mühlenbein, H. (1992).
How genetic algorithms really work I. Mutation and Hillclimbing.
In *Proceedings of the Parallel Problem Solving from Nature 2, (PPSN-II)*, pages 15–26. Elsevier.

Neumann, F. and Wegener, I. (2007).
Randomized local search, evolutionary algorithms, and the minimum spanning tree problem.
Theoretical Computer Science, 378(1):32–40.

Oliveto, P. S., He, J., and Yao, X. (2007).
Time complexity of evolutionary algorithms for combinatorial optimization: A decade of results.
International Journal of Automation and Computing, 4(1):100–106.

Rudolph, G. (1998).
Finite markov chain results in evolutionary computation: A tour d'horizon.
Fundamenta Informaticae, 35(1):67–89.

References III

Scharnow, J., Tinnefeld, K., and Wegener, I. (2002).
Fitness landscapes based on sorting and shortest paths problems.
In *Proceedings of 7th Conf. on Parallel Problem Solving from Nature (PPSN–VII)*, number 2439 in LNCS, pages 54–63.

Witt, C. (2005).
Worst-case and average-case approximations by simple randomized search heuristics.
In *In Proceedings of the 22nd Annual Symposium on Theoretical Aspects of Computer Science (STACS'05)*, number 3404 in LNCS, pages 44–56.