INF FAKULTÄT FÜR INFORMATIK

Evolutionary Algorithms

Meta heuristics and related optimization techniques I/II

Prof. Dr. Rudolf Kruse **Pascal Held**

{kruse,cmoewes}@iws.cs.uni-magdeburg.de

Otto-von-Guericke-Universität Magdeburg

Fakultät für Informatik

Institut für Wissens- und Sprachverarbeitung

Outline

What is a meta heuristic?

- Algorithm for approximately solving of a combinatorial optimization problem
- defines abstract sequence of steps which can be applied on every kind of problem
- but: single steps has to be implemented in a problem-specific way

$\Rightarrow$ problem-specific heuristic

Application of meta heuristics

- on problems where no efficient solving algorithm is known
- e.g. combinatorial optimization problems
- finding an optimal solution is usually not guaranteed
- in comparison with optimal solution: good solutions can be arbitrarily bad
- success and runtime depends on:
  - problem definition and
  - implementation of particular steps

$\Rightarrow$ EAs are meta heuristics, too

Outline

Local Search Methods

according to [Weicker, 2007]

- given: optimization problem $(\Omega, f, \succ)$
- desired: find element $x \in \Omega$, which optimizes f (max. or min.)
- w.l.o.g.: find an element $x \in \Omega$, which maximizes f
  (should f be mimimized, then we consider $f' \equiv -f$)

$\Rightarrow$ **local search methods** to find local optima in Ω

- **assumption:** $f(x_1)$ and $f(x_2)$ differ slightly for similar $x_1, x_2 \in \Omega$
  (no huge jumps in f)
- applicable for arbitrary Ω to find local optima

Local Search Methods

- local search methods = special case of EA
- population: 1 solution candidate $\Rightarrow$ various consequences
  - recombination operator isn't reasonable as there is only one individual
  - changes: mutation resp. variation operator
  - selection: newly created individual instead of parental individual into next generation?

$\Rightarrow$ one fundamental algorithm for all local search methods

- variants by different acceptance criterion
- individuals contain usually no additional information $\mathcal{Z} = \emptyset$
- genotype $\mathcal{G}$ depends on problem (as always)

Algorithm 1 fundamental algorithm of local search

Input: objective function F
Output: solution candidate A

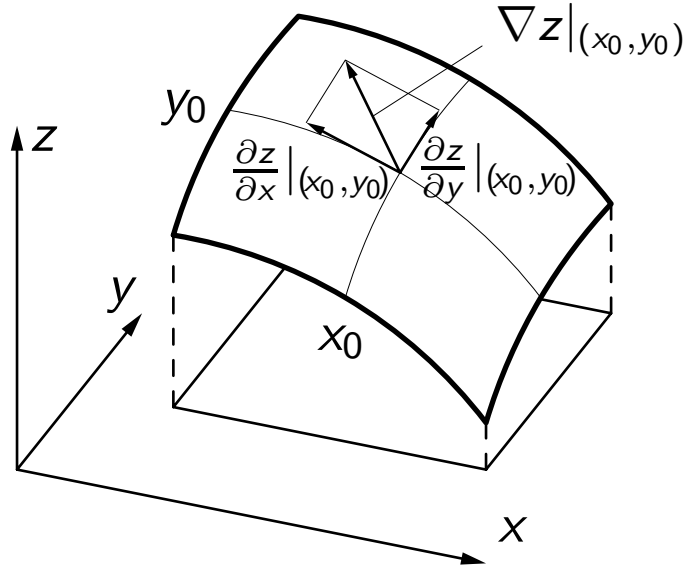
```
 1: t ← 0
 2: A(t) ← create solution candidate
 3: evaluate A(t) by F
 4: while termination criterion isn't fulfilled {
 5:     B = vary A(t)
 6:     evaluate B by F
 7:     t ← t + 1
 8:     if Acc(A(t − 1).F, B.F, t) {          /* acceptance criterion */
 9:         A(t) ← B
10:     } else {
11:         A(t) ← A(t − 1)
12:     }
13: }
14: return A(t)
```

- *Acc*: variably implemented on certain local search methods

Gradient Ascent or Descent

- **Assumption:** $\Omega \subseteq \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$ is differentiable
- **Gradient:** differential operation that creates a vector field

$\Rightarrow$ computes vector into the direction of the steepest ascent of the function in a point



- illustration of the gradient of $z = f(x, y)$ at point (x_0, y_0)

$$\nabla z|_{(x_0,y_0)} = \left(\frac{\partial z}{\partial x}|_{(x_0,y_0)}, \frac{\partial z}{\partial y}|_{(x_0,y_0)} \right)$$

Gradient Ascent or Descent

Idea: start at a randomly chosen point, then make small steps in the search space Ω in (or against) the direction of the steepest slope of the function until a (local) optimum is reached

1. Choose a (random) starting point $\boldsymbol{x}^{(0)} = \left(x_1^{(0)}, \ldots, x_n^{(0)}\right)$
2. Compute the gradient at the current point $\boldsymbol{x}^{(t)}$

$$\nabla_{\boldsymbol{x}} f\left(\boldsymbol{x}^{(t)}\right) = \left(\frac{\partial}{\partial x_1} f\left(\boldsymbol{x}^{(t)}\right), \ldots, \frac{\partial}{\partial x_n} f\left(\boldsymbol{x}^{(t)}\right)\right)$$

3. Make a small step in the direction of the gradient

$$\boldsymbol{x}^{(t+1)} = \boldsymbol{x}^{(t)} + \eta \, \nabla_{\boldsymbol{x}} f\left(\boldsymbol{x}^{(t)}\right)$$

   η: step width parameter („learning rate“ in ANN)
4. Repeat steps 2 and 3 until some termination criterion is fulfilled (e.g. user-specified number of steps has been executed, gradient is smaller than a user-specified threshold)

Problems

choice of the step width parameter

- too small value $\Rightarrow$ large runtime until optimum is reached
- too large value $\Rightarrow$ oscillations, jump back and forth in Ω
- Solution: momentum term, adaptive step width parameter (see also lecture „Neuronale Netze“)

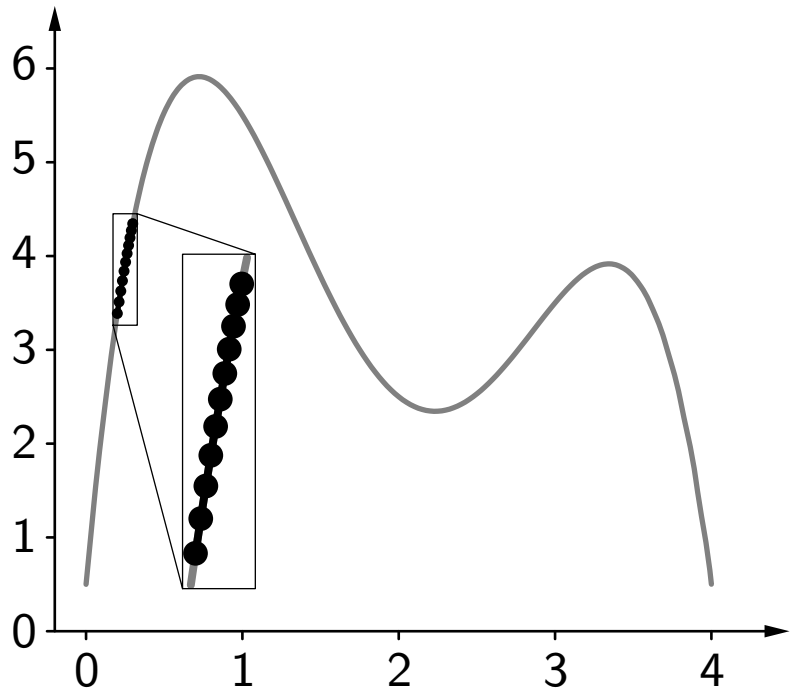
Getting stuck in local maxima

- due to local gradient information, maybe only local maxima is reachable
- problem *can't* be remedied in general
- chance improvement: multiple execution with different starting points

Examples

$$f(x) = -\frac{5}{6}x^4 + 7x^3 - \frac{115}{6}x^2 + 18x + \frac{1}{2}$$

t	x_t	$f(x_t)$	$f'(x_t)$	Δx_t
0	0.200	3.388	11.147	0.011
1	0.211	3.510	10.811	0.011
2	0.222	3.626	10.490	0.010
3	0.232	3.734	10.182	0.010
4	0.243	3.836	9.888	0.010
5	0.253	3.932	9.606	0.010
6	0.262	4.023	9.335	0.009
7	0.271	4.109	9.075	0.009
8	0.281	4.191	8.825	0.009
9	0.289	4.267	8.585	0.009
10	0.298	4.340		

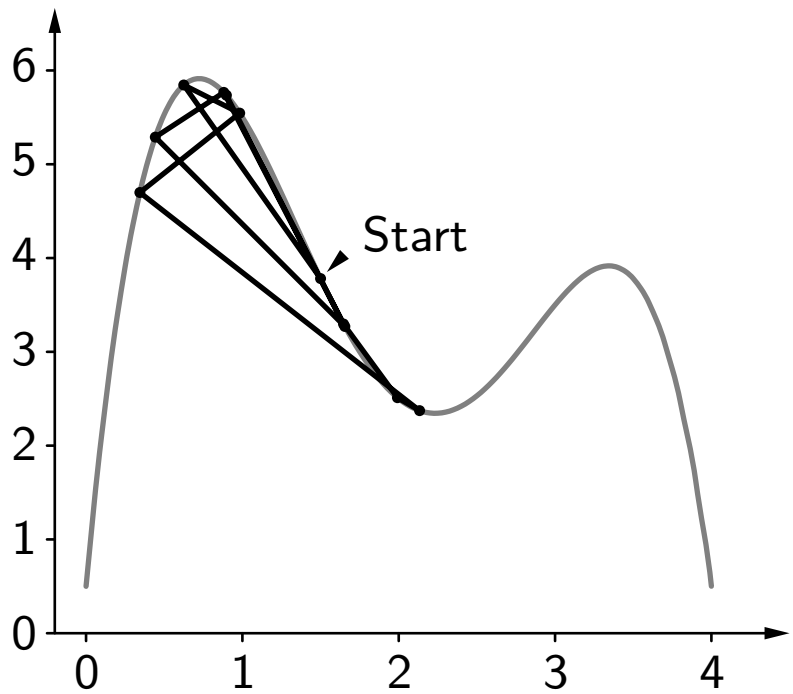


Gradient ascent with starting point 0.2 and $\eta = 0.001$

Examples

$$f(x) = -\frac{5}{6}x^4 + 7x^3 - \frac{115}{6}x^2 + 18x + \frac{1}{2}$$

t	x_t	$f(x_t)$	$f'(x_t)$	Δx_t
0	1.500	3.781	−3.500	−0.875
1	0.625	5.845	1.431	0.358
2	0.983	5.545	−2.554	−0.639
3	0.344	4.699	7.157	1.789
4	2.134	2.373	−0.567	−0.142
5	1.992	2.511	−1.380	−0.345
6	1.647	3.297	−3.063	−0.766
7	0.881	5.766	−1.753	−0.438
8	0.443	5.289	4.851	1.213
9	1.656	3.269	−3.029	−0.757
10	0.898	5.734		

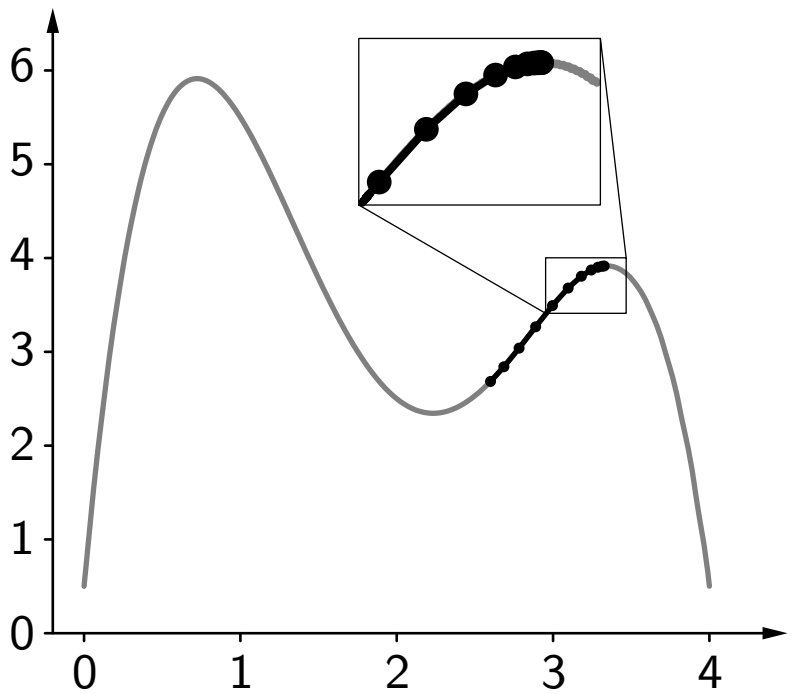


Gradient ascent with starting point 1.5 and $\eta = 0.25$

Examples

$$f(x) = -\frac{5}{6}x^4 + 7x^3 - \frac{115}{6}x^2 + 18x + \frac{1}{2}$$

t	x_t	$f(x_t)$	$f'(x_t)$	Δx_t
0	2.600	2.684	1.707	0.085
1	2.685	2.840	1.947	0.097
2	2.783	3.039	2.116	0.106
3	2.888	3.267	2.153	0.108
4	2.996	3.492	2.009	0.100
5	3.097	3.680	1.688	0.084
6	3.181	3.805	1.263	0.063
7	3.244	3.872	0.845	0.042
8	3.286	3.901	0.515	0.026
9	3.312	3.911	0.293	0.015
10	3.327	3.915		



Gradient ascent with starting point 2.6 and $\eta = 0.05$

Hill Climbing

Idea: If f is not differentiable, determine direction in which f increases by evaluating random points in the vicinity of the current point

1. Choose a (random) starting point $x_0 \in \Omega$
2. Choose a point $x' \in \Omega$ „in the vicinity“ of x_t (e.g. by a small random variation of x_t)
3. Set
$$x_{t+1} = \begin{cases} x' & \text{if } f(x') > f(x_t), \\ x_t & \text{otherwise} \end{cases}$$
4. Repeat steps 2 and 3 until a termination criterion is fulfilled

Hill Climbing

Pseudocode of **acceptance criterion** in **fundamental algorithm**:

Algorithm 2 Acceptance criterion of Hill Climbing

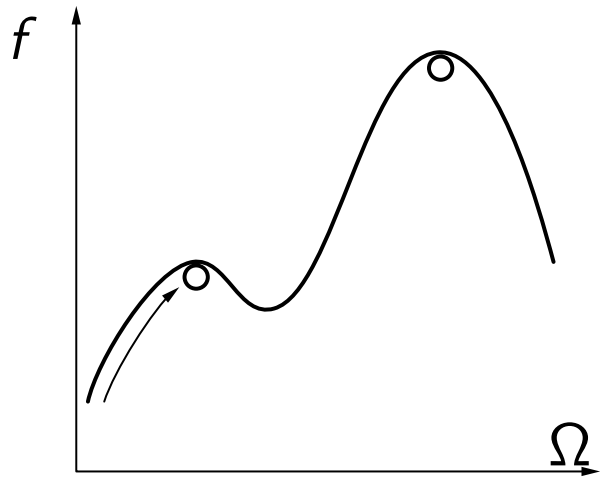
Input: fitness of parents $A.F$, fitness of offspring $B.F$, generation t
Output: **true** or **false**

1: **return** $B.F \succ A.F$

- **Problem:** Getting stuck in local maxima
- all following methods try to remedy this problem

Simulated Annealing

- Extension of hill climbing and gradient ascent which avoids getting stuck
- **Idea:** passage from lower into higher (local) maxima should be more probable than reversed



Principe:

- random variants of current solution are created
- better solutions are **always** accepted
- worse solutions are accepted with certain probability which depends on
  - quality difference between current and new solution and
  - temperature parameter (decreases over the time)

Motivation (Minimizing instead of Maximizing)

- physical minimizing of (lattice) energy when heated metal is cooling down slowly
- process is called **Annealing**
- Purpose: getting softer metal by removing stresses and strains as well as instabilities $\Rightarrow$ easier metal processing

Alternative motivation: (minimizing as well)

- ball rolls around on irregularly curved surface
- function to minimize: potential energy of ball
- at the beginning: certain kinetic energy overcome slopes
- friction: energy is decreasing $\Rightarrow$ stopped moving in a valley finally

Attention: no guarantee to find global optimum

Simulated Annealing

1. Choose a (random) starting point $x_0 \in \Omega$
2. Choose a point $x' \in \Omega$ „in the vicinity“ of current point x_t (for example, by a small random variation of x_t)
3. Set

$$x_{t+1} = \begin{cases} x' & \text{if } f(x') \geq f(x_t), \\ x' \text{ with probability } p = e^{-\frac{\Delta f}{kT}} & \\ x_t \text{ with probability } 1 - p & \text{otherwise} \end{cases}$$

$\Delta f = f(x_t) - f(x')$	quality reduction of the solution
$k = \Delta f_{\max}$	estimate of the range of quality values
T	temperature parameter (decreased over time)

4. Repeat steps 2 and 3 until a termination criterion is fulfilled

for small T method is almost identical to hill climbing

Simulated Annealing

Algorithm 3 Acceptance criterion of Simulated Annealing

Input: parental fitness $A.F$, fitness of offspring $B.F$, generation t
Output: **true** oder **false**

```
1: if B.F ≻ A.F {
2:     return true
3: } else {
4:     u ← choose randomly from U([0, 1])        /* random number
       between 0 and 1 */
5:     if u ≤ exp(−(A.F − B.F)/(k T_{t−1})) {
6:         return true
7:     } else {
8:         return false
9:     }
10: }
```

Threshold Accepting

Idea: very similar to simulated annealing, worse solutions are sometimes accepted again, however, with an upper bound for the quality degradation

1. Choose a (random) starting point $x_0 \in \Omega$
2. Choose a point $x' \in \Omega$ „in the vicinity“ of the current point x_t (for example, by a small random variation of x_t)
3. set

$$x_{t+1} = \begin{cases} x' & \text{if } f(x') \geq f(x_t) - \theta, \\ x_t & \text{otherwise.} \end{cases}$$

   θ threshold for accepting worse solution candidates
   (is (slowly) decreased over time)
   ($\theta = 0$ is equivalent to standard hill climbing)
4. Repeat steps 2 and 3 until a termination criterion is fulfilled

Threshold Accepting

Algorithm 4 Acceptance criterion of Threshold Accepting

Input: parental fitness $A.F$, fitness of offspring $B.F$, generation t

Output: **true** oder **false**

```
1: if B.F ≻ A.F oder A.F − B.F ≤ θ {
2:     return true
3: } else {
4:     return false
5: }
```

Great Deluge Algorithm

Idea: very similar to simulated annealing, worse solutions are sometimes accepted again, absolute lower bound is used

1. Choose a (random) starting point $x_0 \in \Omega$
2. Choose a point $x' \in \Omega$ „in the vicinity“ of the current point x_t (e.g. by a small random variation of x_t)
3. Set

$$x_{t+1} = \begin{cases} x' & \text{if } f(x') \geq \theta_0 + t \cdot \eta, \\ x_t & \text{otherwise} \end{cases}$$

   θ_0 lower bound for the quality of the candidate solutions at $t = 0$ (initial „water level“)

   η step width parameter („speed of the rain“)
4. Repeat steps 2 and 3 until a termination criterion is fulfilled

Great Deluge Algorithm

Algorithm 5 Acceptance criterion of Great Deluge Algorithm

Input: parental fitness $A.F$, fitness of offspring $B.F$, generation t
Output: **true** oder **false**

```
1: if B.F ≻ θ_0 + η · t {
2:     return true
3: } else {
4:     return false
5: }
```

Record-to-Record Travel

Idea: similar to great deluge algorithm, rising water level is used, linked to the fitness of the best found individual

- possbile degradation: always seen in relation to the best found individual
- only if there is an improvement: current individual is important
- similar to threshold accepting: a monotonously increasing sequence of real numbers controls the selection of poor individuals

Record-to-Record Travel

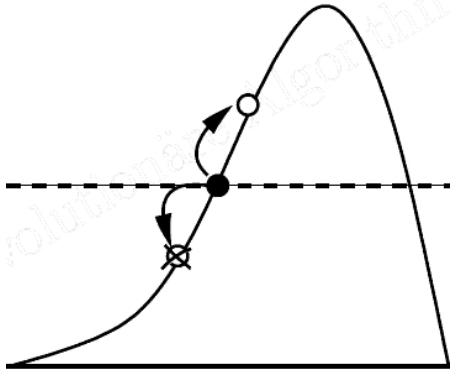
Algorithm 6 Acceptance criterion of Record-to-Record Travel

Input: parental fitness $A.F$, fitness of offspring $B.F$, t, best found quality F_{best}

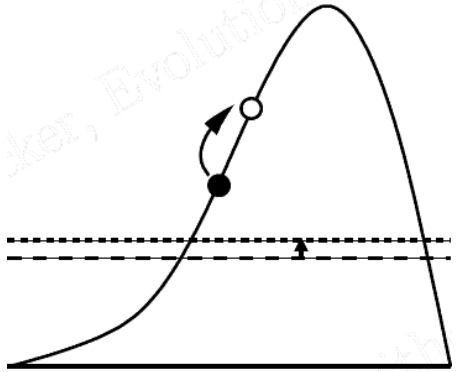
Output: **true** oder **false**, F_{best}

1: **if** $B.F \succ F_{\text{best}}$ {
2: $F_{\text{best}} \leftarrow B.F$
3: **return** **true**, F_{best}
4: } **else** {
5: **if** $B.F - F_{\text{best}} < T_t$ {
6: **return** **true**, F_{best}
7: }
8: }
9: **return** **false**, F_{best}

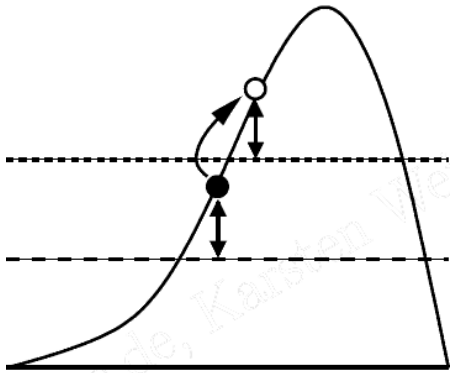
Comparison of local search algorithms



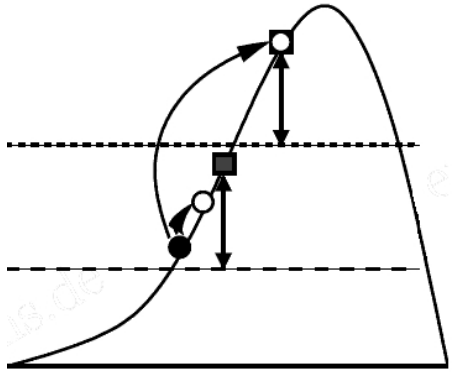
Hill Climbing



Great Deluge Algorithm



Threshold Accepting



Record-to-Record Travel

Outline

Example: The Traveling Salesman Problem

dt. Problem des Handlungsreisenden (TSP)

- **given:** set of n cities (idealized as points on a plane) and distances/costs of the routes between the cities
- **desired:** round trip with minimum distance through all n cities so that each city is visited exactly once
- **mathematically:** Find a so-called Hamiltonian cycle (contains each node once) with mimimal total weight in a weighted graph
- **known:** TSP is NP-complete $\Rightarrow$ no algorithm is known that solves this problem in polynomial time
- **therefore:** for large n only approximate solutions can be computed in reasonable time (best solution could be found in some cases, but not every time)
- **here:** using hill climbing and simulated annealing to solve the problem

Example: The Traveling Salesman Problem

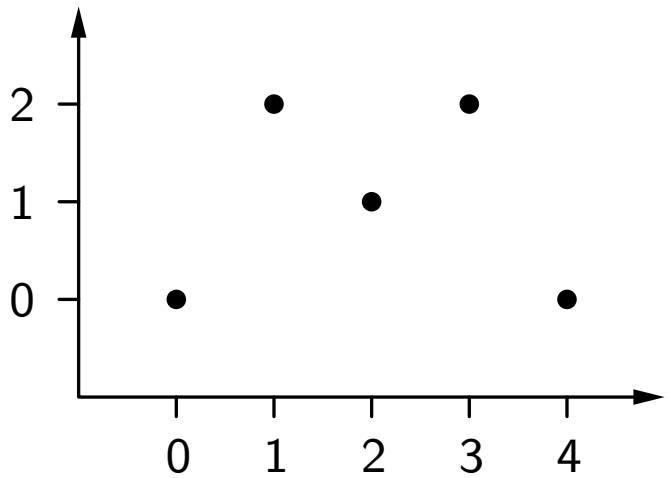
1. Order the cities randomly (that is, create a random round trip)
2. Randomly choose two pairs of cities such that each pair consists of cities that are neighbors in the current round trip and such that all four cities are distinct, split the round trip between the cities of each pair and reverse the interjacent part
3. If this new round trip is better (that is, shorter or cheaper) than the old, then replace the old round trip with the new one.
   Otherwise replace the old round trip with probabilityt $p = e^{-\frac{\Delta Q}{kT}}$

ΔQ	quality difference between the old and the new round trip
k	estimate of the range of round trip qualities, e.g. $k_t = \frac{t+1}{t} \max_{i=1}^{t} \Delta Q_i$, where ΔQ_i is the quality difference in the i-th step and t is the current step)
T	temperature parameter (reduced over time, e.g. $T = \frac{1}{t}$)

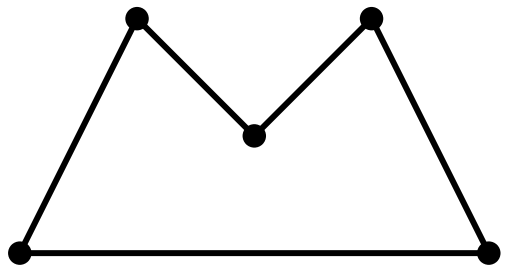
4. Repeat steps 2 and 3 until a termination criterion is met

Example: The Traveling Salesman Problem

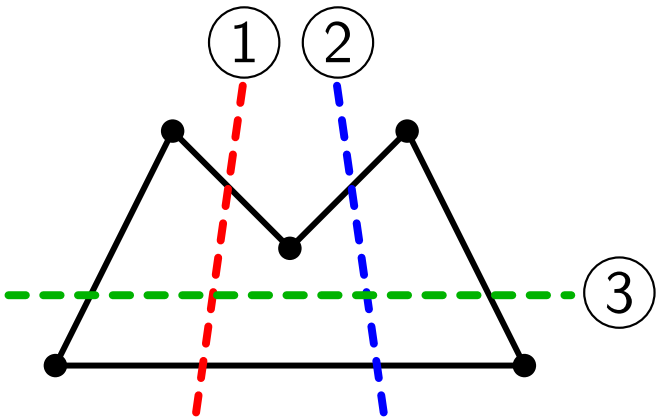
Pure hill climbing can get stuck in local minimum:



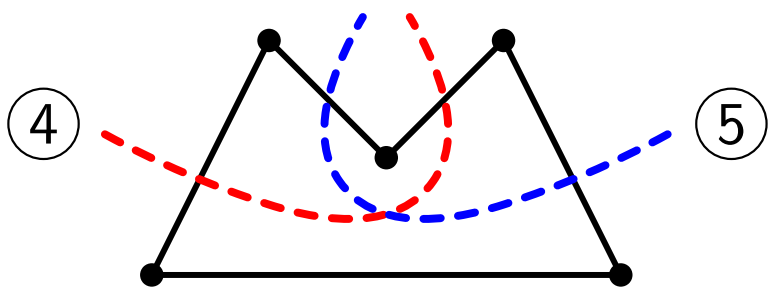
initial round trip



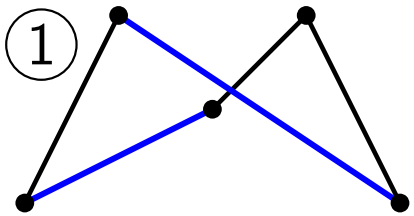
Length: $2\sqrt{2} + 2\sqrt{5} + 4 \approx 11.30$

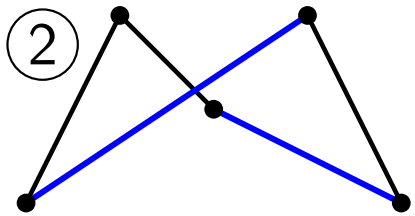


possible seperations of the round trip

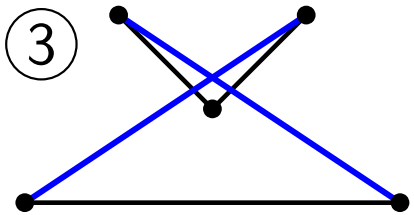


Example: The Traveling Salesman Problem

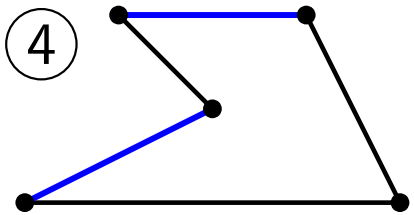


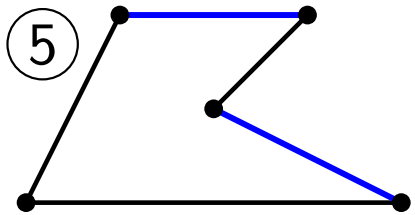


$\sqrt{2} + 3\sqrt{5} + \sqrt{13} \approx 11.73$



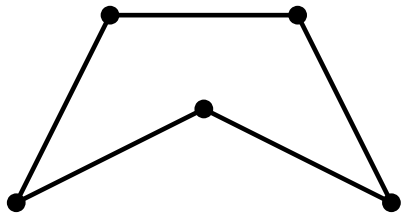
$\sqrt{2} + 2\sqrt{13} + 4 \approx 14.04$





$\sqrt{2} + 2\sqrt{5} + 2 + 4 \approx 11.89$

global optimum:



$4\sqrt{5} + 2 \approx 10.94$

Example: The Traveling Salesman Problem

- all modifications of the initial round trip results in worse round trips $\Rightarrow$ global optimum (of this round trip) is not reachable by pure hill climbing
- in contrast, simulated annealing accepts sometimes worse solutions $\Rightarrow$ way to find global optimum (*but*: no guarantee that this will happen!)
- **be careful:** it depends on permitted operations if search can get stuck in local optimum:
  - when using another operation (change of the position of a city in round trip resp. delete of current position and insert on another one), getting stuck can be avoided in considered example
  - for this set of operations: construction of an example that gets stuck in local minimum

Outline

Tabu search

- local search method that considers history when creating new individual
- *tabu lists* avoid recurrence on previously considered solution candidates

- tabu-list = FIFO-queue with fixed length
  - entries are complete solution candidate or aspects
  - mutations are not allowed to create tabu entries
  - often: FIFO list with beneficial properties can break tabu
  - even possbile: new best entire quality breaks tabu

Example

Graph coloring with k colors so that two nodes connected with an edge are colored differently

- $\Omega = \{1, \ldots, k\}^n$
- minimizing of $f(x) = \sum_{(v_i, v_j) \in E} \begin{cases} 1 & \text{if } x_i = x_j \\ 0 & \text{otherwise} \end{cases}$
- mutating colors v_i from c to $d \Rightarrow$ tabu entry (i, c)

Algorithm 7 Tabu search

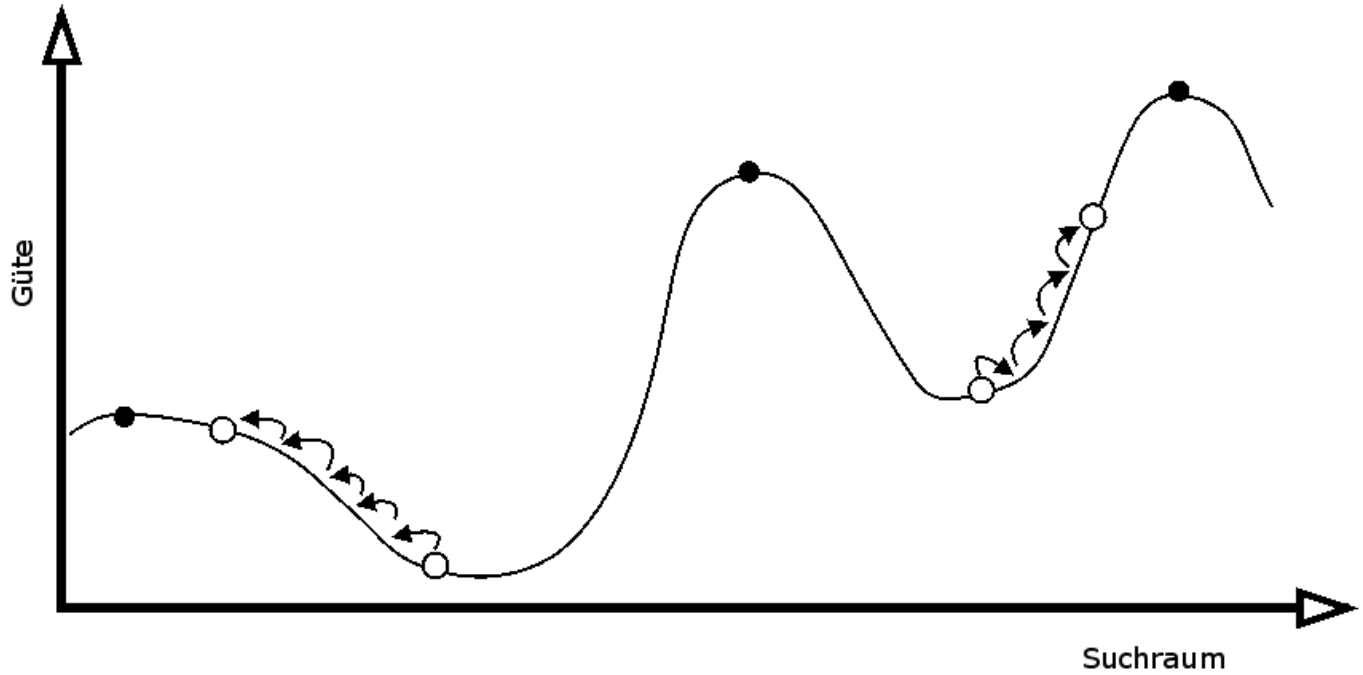
Input: Target function F
Output: best individual A_{best}

1: $t \leftarrow 0$
2: $A(t) \leftarrow$ create random solution candidate
3: evaluate $A(t)$ by F
4: $A_{\text{best}} \leftarrow A(t)$
5: initalize Tabu-list
6: **while** termination criterion isn't fulfilled {
7: $P \leftarrow \emptyset$
8: **while** $|P| < \lambda$ {
9: $B \leftarrow$ mutate $A(t)$
10: evaluate B by F
11: **if** $(A(t), B) \notin$ Tabu-list or $B.F \succ A_{\text{best}}.F$ {
12: $P \leftarrow P \cup \{B\}$
13: }
14: }
15: $t \leftarrow t + 1$
16: $A(t) \leftarrow$ best individual of P
17: **if** $A(t).F \succ A_{\text{best}}.F$ {
18: $A_{\text{best}} \leftarrow A(t)$
19: }
20: Tabu-list $\leftarrow$ update by $(A(t-1), A(t))$
21: }
22: **return** A_{best}

Memetic Algorithms

- on the one hand: population-based algorithms
  - Advantage: widespread trawling of the (search) space
  - Disadvantage: slow
- otherwise: local search
  - Advantage: fast optimization
  - Disadvantage: susceptible to local optima
- **Memetic Algorithms:** Combination of both techniques
- origin of the name (Richard Dawkins): „memes“ are elements of the behaviour which can be acquired individually (in contrast to genes)
- Procedure: each new individual will be optimized immediately
  - only few steps or
  - till local optimum

Example: SAGA



„Simulated Annealing Genetic Algorithm“

Memetic Algorithms

Algorithm 8 Memetic Algorithm

Input: Evaluation function F
Output: best individual

1: $t \leftarrow 0$
2: $P(t) \leftarrow$ initialize population of size μ
3: $P(t) \leftarrow \text{LOCAL SEARCH}(F)$ on each individual in $P(t)$
4: evaluate $P(t)$ by F
5: **while** termination criterion isn't fulfilled $\{$
6: $\quad E \leftarrow$ select parents for λ offspring in $P(t)$
7: $\quad P' \leftarrow$ create offspring by recombination on E
8: $\quad P'' \leftarrow$ mutate individual in P'
9: $\quad P''' \leftarrow \text{LOCAL SEARCH}(F)$ on each individual in P''
10: $\quad$ evaluate P''' by F
11: $\quad t \leftarrow t + 1$
12: $\quad P(t) \leftarrow$ environmental selection on P'''
13: $\}$
14: **return** best individual of $P(t)$

Properties

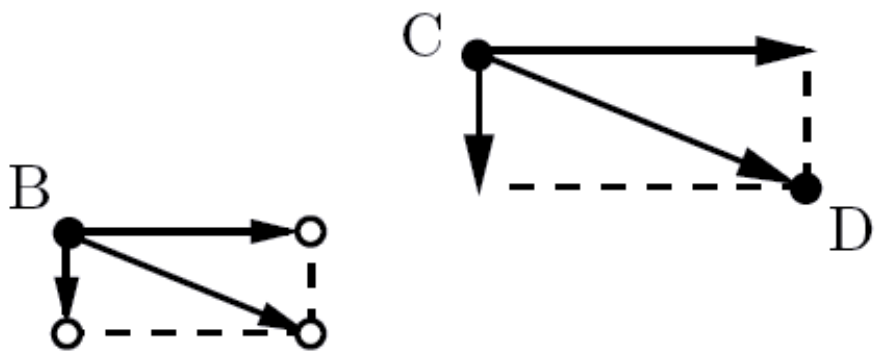
- often strongly accelerated optimization
- but: search dynamic can be limited crucially
  - mutation gets stuck frequently in (wide) local optima
  - recombination has bounded starting situation
  - parts of the space are perhaps unreachable

- method correlates with lamarckian evolution

Differential evolution

Idea

- no adaptation of the step width in $A.S$
- but: relations of individuals in the population can be used as basis for the step width



Differentialevolution

Algorithm 9 DE-Operator

Input: individuals A, B, C, D

Output: optimized individual A'

1: index $\leftarrow$ choose random number according to $U(\{1, \ldots, l\})$
2: **for each** $i \in \{1, \ldots, l\}$ $\{$
3: $\quad u \leftarrow$ choose random number according to $U([0, 1))$
4: $\quad$ **if** $u \leq \tau$ **or** $i =$ index $\{$
5: $\quad\quad A'.G_i \leftarrow B'.G_i + (C.G_i - D.G_i) \cdot \alpha$
6: $\quad$ $\}$ **else** $\{$
7: $\quad\quad A'.G_i \leftarrow A.G_i$
8: $\quad$ $\}$
9: $\}$
10: **return** A'

Details

- DE-operator: combinations of recombination and mutation
- selection: a new individual replaces parental individual if and only if it has a better fitness

Parameter	Range
Population size μ	10–100, $10 \cdot n$
Weighting of recombination τ	0.7–0.9
Scaling factor α	0.5–1.0

Algorithm 10 Differentialevolution

Input: evaluation function F
Output: best individual of $P(t)$

1: $t \leftarrow 0$
2: $P(t) \leftarrow$ create population of size μ
3: evaluate $P(t)$ by F
4: **while** termination criterion isn't fulfilled $\{$
5: $\quad P(t+1) \leftarrow \emptyset$
6: $\quad$ **for** $i \leftarrow 1, \ldots, \mu$ $\{$
7: $\quad\quad$ **do** $\{$
8: $\quad\quad\quad A, B, C, D \leftarrow$ select parents uniformly random from $P(t)$
9: $\quad\quad$ $\}$ **while** A, B, C, D pairwise distinct
10: $\quad\quad A' \leftarrow \text{DE-OPERATOR}(A, B, C, D)$
11: $\quad\quad$ evaluate A' by F
12: $\quad\quad$ **if** $F(A') \succ F(A)$ $\{$
13: $\quad\quad\quad P(t+1) \leftarrow P(t+1) \cup \{A'\}$
14: $\quad\quad$ $\}$ **else** $\{$
15: $\quad\quad\quad P(t+1) \leftarrow P(t+1) \cup \{A\}$
16: $\quad\quad$ $\}$
17: $\quad$ $\}$
18: $\}$
19: $t \leftarrow t+1$
20: **return** best individual of $P(t)$

Scatter Search

Ingredients

- population with solution candidates
- Variation operators
- selections pressure
- further: local search

But...

- a deterministic method!
- spacious exploration of the space:
  - wide initializing
  - systematical generation of new individuals

Procedure

iterative process with two phases

1. generating of new individuals and selection of those which guarantee the most possible variety
2. recombination of all „pairings“ of the chosen individuals
3. selection of the best and iteration until nothing changes anymore

Example

- real-valued problem spaces with $\mathcal{G} = \Omega = \mathbb{R}^n$

Steo 1

1. Diversity generator creates μ individuals in P
   on the example:
   - per dimension of the space a value range containing four partitions
   - notice of the frequency of the generated individuals per partition
   - inverse proportionally choosing of the partition

2. separate: population of the best α individuals is expanded by those β individuals of P which maximize $\min_{B \in P_{\text{best}}} d(A.G, B.G)$ $(A \in P)$

Step 2

1. Subset generator chooses individuals of the set with the best ones
   Example: (here it's obvious) all possible pairs
2. Applying combination operator
   (Example: arithmetic Crossover with $U\left(\left[-\frac{1}{2}, \frac{3}{2}\right]\right)$)
3. Local optimization
4. If all ind. are created $\Rightarrow$ selection of the $\alpha + \beta$ best ones
5. Iterate until set is not changing anymore
6. α best individuals for next step 1

Algorithm 11 Scatter Search

Input: Evaluation function F

1: $P_{\text{best}} = \emptyset$; $P = \emptyset$
2: **for** $t \leftarrow 1, \ldots,$ maxIter {
3: **while** $|P| < \mu$ {
4: $A \leftarrow$ create individual with the diversity generator
5: $A \leftarrow$ LOCAL SEARCH(F) applied on A; evaluate A by F
6: **if** $A \notin P \cup P_{\text{best}}$ {
7: $P \leftarrow P \cup \{A\}$
8: }
9: }
10: **if** $t = 1$ {
11: $P_{\text{best}} \leftarrow$ select α individuals of P with BEST-SELECTION
12: $P \leftarrow$ discard individuals of P_{best} in P
13: }
14: **for** $k \leftarrow 1, \ldots, \beta$ {
15: $A \leftarrow$ the one individual of P that maximizes $\min_{B \in P_{\text{best}}} d(A.G, B.G)$
16: $P \leftarrow$ discard individual A in P
17: $P_{\text{best}} \leftarrow P_{\text{best}} \cup \{A\}$
18: }
19: **do** {
20: $P' \leftarrow \emptyset$; Sets $\leftarrow$ create subsets of P_{best} by using subset operator
21: **for each** $M \in$ Sets {
22: $A \leftarrow$ apply combination operator on M
23: $A \leftarrow$ LOCAL SEARCH(F) applied on A; evaluate A by F
24: **if** $A \notin P_{\text{best}} \cup P'$ {
25: $P' \leftarrow P' \cup \{A\}$
26: }
27: }
28: $P_{\text{best}} \leftarrow$ select $\alpha + \beta$ Ind. of $P_{\text{best}} \cup P'$ with BEST-SELECTION
29: } **while** P_{best} has not changed
30: $P_{\text{best}} \leftarrow$ select α individuals of P with BEST-SELECTION
31: }
32: **return** best individual in P_{best}

Scatter Search

Recommended parameters

parameter	value range
Population size μ	50–150
Total number of best individuals α	5–20
Expanding of best individuals β	5–20

Cultural Algorithm

Motivation:

- further information memory in addition to the genetic layer
- culture determines behaviour based on certain values
- layer of culture is introduced in EAs

collectively cultural knowledge:

- Memory: belief space (dt. *Überzeugungsraum*)
- is modified by the best individuals of a generation
- situational and normative knowledge

Algorithm 12 Cultural Algorithm

Input: Evaluation function F

Output: best individual in $P(t)$

1: $t \leftarrow 0$
2: $P(t) \leftarrow$ initialize population
3: $\mathcal{BS}(t) \leftarrow$ initialize belief space
4: evaluate $P(t)$ by F
5: **while** termination criterion is not fulfilled $\{$
6: $\quad P' \leftarrow$ determine important individuals of $P(t)$
7: $\quad \mathcal{BS}(t+1) \leftarrow \mathcal{BS}(t)$ is adapted by P'
8: $\quad P'' \leftarrow$ create offspring of $P(t)$ on the basis of $\mathcal{BS}(t+1)$
9: $\quad$ evaluate $P(t)$ by F
10: $\quad t \leftarrow t+1$
11: $\quad P(t) \leftarrow$ selection from $P' \cup P(t-1)$
12: $\}$
13: **return** best individual of $P(t)$

Cultural Algorithm

Situational knowledge

- for $\Omega = \mathbb{R}^n$: two best individuals of the prior generation
- mutation should be orientated in those direction if necessary

Normative knowledge

- for $\Omega = \mathbb{R}^n$: upper and lower bound per dimension
- detect biggest/smallest value of the 20% best individuals of the prior generation
- always: accept increase
- only on „bound indivdual“ with better quality: accept decrease

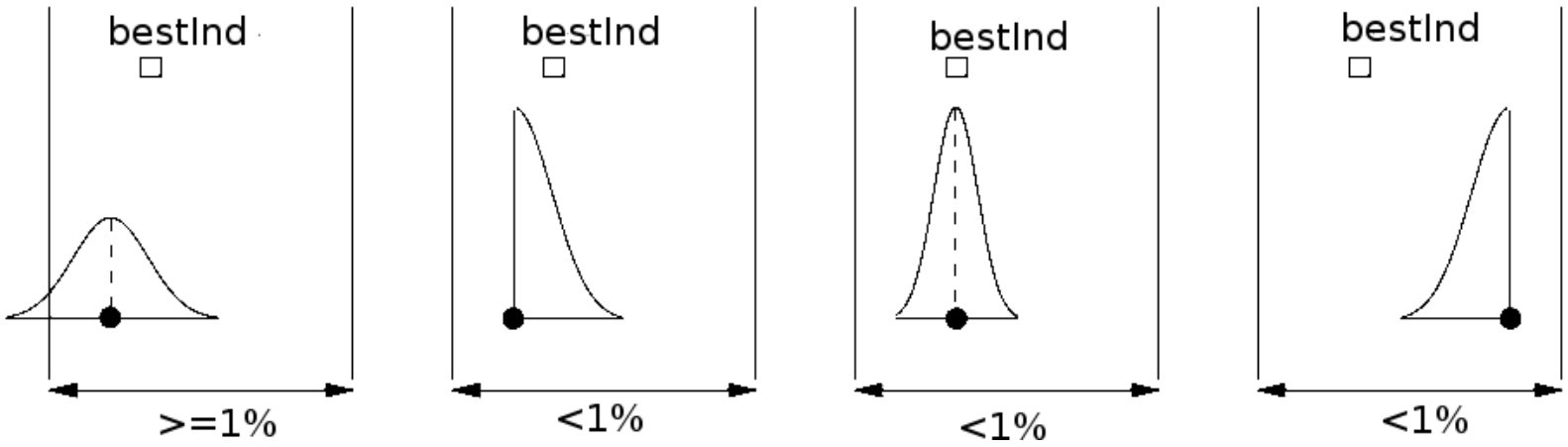
Cultural Algorithm

Stable space dimensions

- if range is limited on $< 1\%$ and
- contain last best individuals

Mutation

- if stable: orientate on best individual
- otherwise: self-adapted step width



Cultural Algorithm

Algorithm 13 KA-Mutation

Input: Individual A
Output: Individual B

1: $u' \leftarrow$ choose randomly according to $\mathcal{N}(0,1)$
2: **for each** $i \in \{1,\ldots,l\}$ {
3: $\quad u_i'' \leftarrow$ choose randomly according to $\mathcal{N}(0,1)$
4: $\quad B.S_i \leftarrow A.S_i \cdot \exp\left(\frac{1}{\sqrt{2l}} \cdot u' + \frac{1}{\sqrt{2\sqrt{l}}} \cdot u_i''\right)$
5: $\quad u \leftarrow$ choose randomly according to $\mathcal{N}(0, B.S_i)$
6: $\quad$ **if** $\mathcal{BS}$ is stable on dimension i {
7: $\quad\quad$ **switch**
8: $\quad\quad$ **case** $A.G_i < \text{BestInd.G}_i : B.G_i \leftarrow A.G_i + |u|$
9: $\quad\quad$ **case** $A.G_i > \text{BestInd.G}_i : B.G_i \leftarrow A.G_i - |u|$
10: $\quad\quad$ **case** $A.G_i = \text{BestInd.G}_i : B.G_i \leftarrow A.G_i + \frac{u}{2}$
11: $\quad$ } **else** {
12: $\quad\quad B.G_i \leftarrow A.G_i + u$
13: $\quad\quad B.G_i \leftarrow \max\{\text{ug}_i, \min\{\text{og}_i, B.G_i\}\}$
14: $\quad$ }
15: }
16: **return** B

Literatur zur Lehrveranstaltung

Weicker, K. (2007).
Evolutionäre Algorithmen.
Teubner Verlag, Stuttgart, Germany, 2nd edition.