



Wintersemester 2005/06

**Fundamente der Computational Intelligence**  
**(Vorlesung)**

Prof. Dr. Günter Rudolph

Fachbereich Informatik

Lehrstuhl für Algorithm Engineering





## Inhalt

- Grundlagen Optimierung
- Nachbarschaftssuche
- (1+1)-EA
  - Erwartete Laufzeit
    - Counting Ones
    - Leading Ones
  - Konvergenz
- Populationsbasierte EA



## Konvergenzaussagen

### **Satz:**

Sei  $D_k = |f(x_k) - f^*|$  für  $k \geq 0$  durch einen (1+1)-EA generiert,  
 $S^* = \{x^* \in S : f(x^*) = f^*\}$  die Menge der optimalen Lösungen und  
 $P_m(x, S^*)$  die W'keit, von  $x \in S$  durch eine Mutation nach  $S^*$  zu gelangen.

Wenn für jedes  $x \in S \setminus S^*$  gilt  $P_m(x, S^*) \geq \delta > 0$ , dann  $D_k \rightarrow 0$  vollständig.

### **Beweis:**

Für den (1+1)-EA gilt:  $P(x, S^*) = 1$  für  $x \in S^*$  wg. Selektion des Besseren.

Es reicht also zu zeigen, dass der EA durch Mutation  $S^*$  sicher erreicht:

Erfolg in 1. Iteration:  $P_m(x, S^*) \geq \delta$ .

Kein Erfolg in 1. Iteration  $\leq 1 - \delta$ .

Kein Erfolg in k. Iteration  $\leq (1 - \delta)^k$ .

$\Rightarrow$  Erfolg in k. Iteration  $\geq 1 - (1 - \delta)^k \rightarrow 1$  für  $k \rightarrow \infty$ .

Wg.  $P\{D_k > \varepsilon\} \leq (1 - \delta)^k \rightarrow 0$  folgt Konvergenz in W'keit und da  
folgt sogar vollständige Konvergenz.

$$\sum_{k=0}^{\infty} (1 - \delta)^k < \infty \quad \blacksquare$$



Die Bedingung:  $\forall x \in S \setminus S^*$  gilt  $P_m(x, S^*) \geq \delta > 0$

- ist hinreichend, aber nicht notwendig und
- etwa für die globale Mutation mit Mutationsw'keit  $p \in (0,1)$  erfüllt, da

$$P_m(x, y) = p^{H(x,y)} (1-p)^{n-H(x,y)} > 0 \text{ für alle } x, y \text{ in } S,$$

wobei  $n$  = Dimension und  $H(x,y)$  der Hamming-Abstand zwischen  $x$  und  $y$ .

### Achtung:

Die gleiche Konvergenzaussage erhält man für reine Zufallssuche:

```
Wähle  $x_0 \in S$ , setze  $k = 0$ .
repeat
  Wähle  $Y_k$  zufällig gleichverteilt aus  $S$ .
  Falls  $f(Y_k) < f(X_k)$  dann  $X_{k+1} = Y_k$ 
                        sonst  $X_{k+1} = X_k$ 
until Terminierung
```



### Folgerung:

Solche Aussage sind von geringer Bedeutung, weil  $\delta$  astronomisch klein!

### Beispiel:

alle Bits müssen invertiert werden  $\Rightarrow p^n = n^{-n} = \delta$

Übergang zum Optimum erfolgt mit W'keit  $\delta$  (geometrische Verteilung)

$\Rightarrow$  mittlere Zeit bis zum Eintreffen des Ereignisses  $1/\delta = n^n$

Sei  $n = 20$  und wir haben einen Tera-Hertz-Rechner ( $10^{12}$  Iterationen/Sekunde).

Dann warten wir im Mittel  $20^{20} / 10^{12}$  Sekunden =  $2^{20} \times 10^8$  Sekunden auf Lösung,

also über 3,325 Millionen Jahre!

$\Rightarrow$  Das ist mathematisch gesehen zwar endlich, praktisch jedoch **unendlich!**



### **Also:**

Stochastische Konvergenzaussagen (auch für Simulated Annealing etc.) sind mit Vorsicht zu genießen!

### **Aber:**

Negative Aussage (keine Konvergenz bzw. kein sicheres Besuchen des Optimums) ist schon von Bedeutung!

### **Beispiel:**

Sicheres Auffinden des Optimums in Zeit  $T$  mit  $W$ 'keit  $\gamma > 0$ .

⇒ Multistart!  $W$ 'keit des Mißerfolgs sinkt exponentiell schnell!

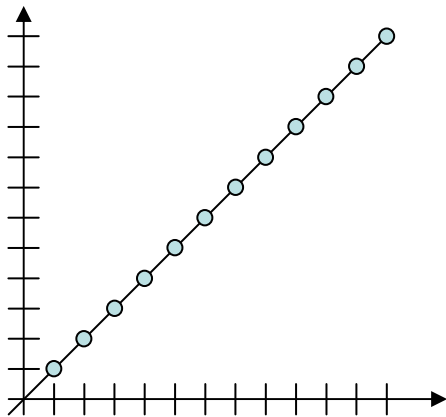
Allerdings spielt Größe von  $\gamma$  eine Rolle!



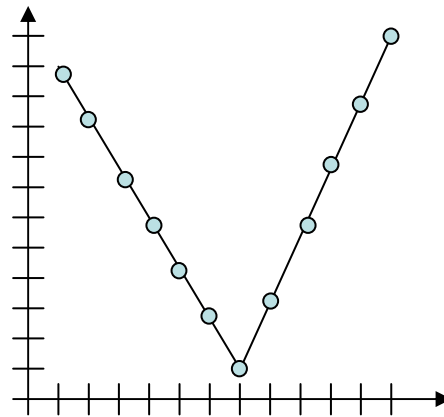
“Functions of Unication“ (Goldberg/Deb 1993):

Es existiert  $g: \{0, 1, \dots, n\} \rightarrow \mathbb{R}$  mit  $f(x) = g(\|x\|)$ , wobei  $\|\cdot\|$  Hamming-Norm.

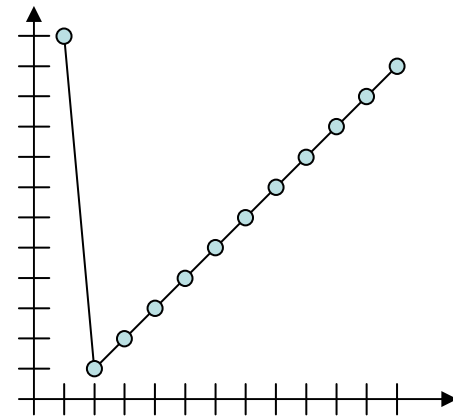
**Beispiel:** Counting Ones  $f(x) = \|x\|$



$\gamma = 1$



$\gamma = 0.5$



$\gamma$  klein

NIAH  
(needle-in-a-haystack)



Populationen:  $(\mu + \lambda)$ -EA

```
Wähle  $\mu$  Individuen  $\in S$ 
```

```
repeat
```

```
  do  $\lambda$  times:
```

```
    wähle zufällig ein Elter aus und
```

```
    mutiere Individuum
```

```
  end ( $\rightarrow \lambda$  Nachkommen)
```

```
  wähle  $\mu$  beste aus  $\mu + \lambda$  Individuen aus
```

```
until Terminierung
```

Reproduktion  
durch Mutation

Selektion

$\mu = \lambda = 1$  ist (1+1)-EA





Bei Populationen ist Variation nicht auf Mutation begrenzt:

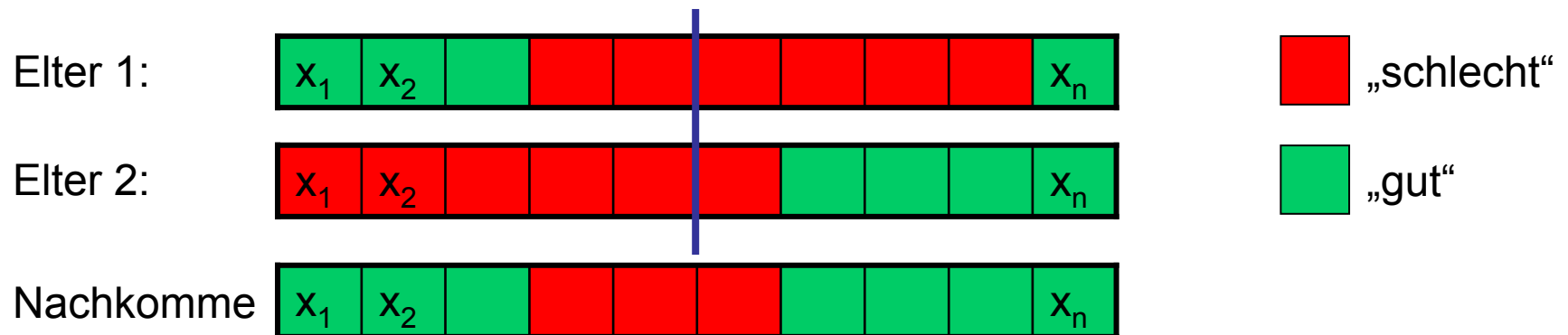
Sexuelle Reproduktion! → **Rekombination / Crossover**

Warum manchmal sinnvoll?

**Annahme:** Separierbares Problem  $f(x) = \sum_{i=1}^n f_i(x_i)$

d.h. Lösung des Gesamtproblems durch Lösung kleiner Teilprobleme

⇒ Rekombination von Teillösungen





- **1-Punkt-Crossover:**

Wähle 1 Position  $k \in \{ 1, 2, \dots, n-1 \}$  zufällig aus.

Nachkomme besteht aus ersten  $k$  Komponenten des 1. Elter und  $n-k$  letzten Komponenten des 2. Elter.

- **Multi-Point-Crossover:**

Wähle mehrere Positionen zufällig ohne Zurücklegen aus.

Positionen sortieren.

Dann immer abwechselnd Komponenten übernehmen bis zur jeweiligen Position vom 1. oder 2. Elter.

- **Uniform Crossover:**

Für jede Komponente Elter zufällig auswählen.

Normalerweise gleichverteilt.

Es gibt: Verallgemeinerungen für mehr als 2 Eltern.

Zudem: Das funktioniert bei allen Produkträumen! Auch  $\mathbb{R}^n$ ,  $\mathbb{Z}^n$ , ...



Selektion: seien  $v_1, v_2, \dots, v_\lambda$  Fitnesswerte der Population

- **Proportionale Selektion** (Holland 1975)  
Annahme:  $v_i > 0$

W'keit, das  $i$ -te Individuum zu selektieren:

$$s_i = \frac{v_i}{\sum_{j=1}^{\lambda} v_j}$$

Bei  $\lambda$  Selektionen:

Individuum  $i$  wird im Mittel  $\lambda s_i$  mal gezogen

Wenn  $v_i \geq \lambda^{-1} \sum_j v_j$  („above average fitness“),  
dann wird Individuum  $i$  im Mittel mindestens 1 mal gezogen

**Allerdings:** keine Garantie, dass das beste Individuum „überlebt“



Selektion: seien  $v_1, v_2, \dots, v_\lambda$  Fitnesswerte der Population

- **Stochastic Universal Sampling** (Baker 1987)

Annahme:  $v_i > 0$

Man will erzwingen, dass etwa  $\lambda$  si Individuen gezogen werden:

Sei  $c_i$  die Summe der Selektionsw'keiten bis Index  $i$ :  $c_i = s_1 + s_2 + \dots + s_i$

ziehe  $U \in (0,1)$  gleichverteilt, setze  $a = 0$

```
for i = 1 to  $\lambda$  do
  while  $\lambda c_i > a + U$  do
    selektiere Individuum  $i$ 
    a++
  endwhile
endfor
```

Das beste Individuum  
überlebt sicher!



Selektion: seien  $v_1, v_2, \dots, v_\lambda$  Fitnesswerte der Population

- **Tunierselektion** (Goldberg & Deb 1991)

Man zieht  $q$  mal gleichverteilt mit W'keit  $1 / \lambda$   
Das beste der  $q$  Individuen wird selektiert.

Annahme:  $v_i$  aufsteigend sortiert (Rang 1 ist der beste)

$Q_i = \{ i, i+1, \dots, \lambda \}^q$  mit  $|Q_i| = (\lambda - i + 1)^q$  enthält Ind. mit Rang  $\geq i$ ,

$Q_i$  enthält auch Tupel ohne Index  $i$ , diese müssen wir abziehen:

$Q_i \setminus Q_{i+1}$  mit  $|Q_i \setminus Q_{i+1}| = |Q_i| - |Q_{i+1}| = (\lambda - i + 1)^q - (\lambda - i)^q$

also:  $s_i = [(\lambda - i + 1)^q - (\lambda - i)^q] / \lambda^q$  und  $s_i < 1$



Selektion: seien  $v_1, v_2, \dots, v_\lambda$  Fitnesswerte der Population

- $\mu$  **beste** (Schwefel 1975)

- $(\mu + \lambda)$ -**Selektion**

- wähle  $\mu$  beste Individuen aus  $\mu$  Eltern und  $\lambda$  Nachkommen  
 $\Rightarrow$  der beste Elter überlebt mit W'keit 1

- $(\mu, \lambda)$ -**Selektion**

- wähle  $\mu$  beste Individuen nur von den  $\lambda$  Nachkommen  
 $\Rightarrow$  der beste Elter kann nicht überleben

- begrenzte Lebensdauer

manchmal auch „truncation selection“ genannt



Selektion: seien  $v_1, v_2, \dots, v_\lambda$  Fitnesswerte der Population

- **EP-Turnierselektion** (Fogel)

selektieren aus  $\mu$  Eltern und  $\lambda$  Nachkommen

Für jedes Individuum  $i$ :

ziehe  $q$  Individuen gleichverteilt

jedesmal wenn  $i$  nicht schlechter, dann erhält es Bonuspunkt

Selektiere  $\mu$  Individuen mit meisten Bonuspunkten

Man kann Fälle konstruieren, dass der beste Elter nicht überlebt!



### **Restriktionen / Nebenbedingungen:**

- Lethalmutationen
- Straffunktionen
- Reparatur