# Evolutionäre Algorithmen: Schemata und Formae

Karsten Weicker

HTWK Leipzig

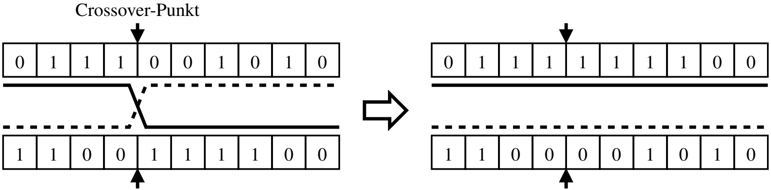
5. Januar 2013

# Überblick

# Überblick

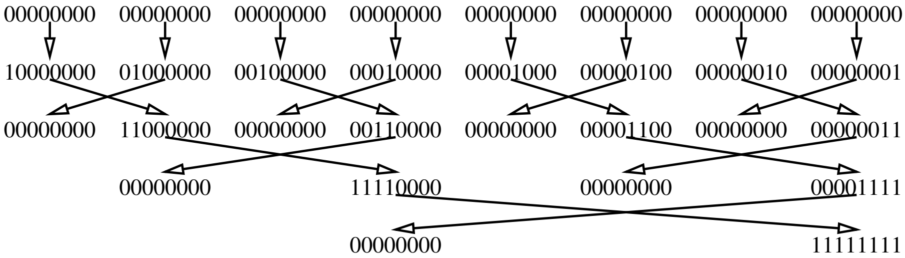
# Motivation

- zunächst: genetischer Algorithmus
- 1-Punkt-Crossover
- Frage: was für eine Suchdynamik ergibt sich durch die kombinierende Rekobmination?



# Motivation

Wie kann sich die Laufzeit durch Kombination ändern?



# Grundbegriffe

## Def.: Schema

- Sei $\mathcal{G} = \mathbb{B}^l$
- Dann: $H \in \{0, 1, *\}^l$ heißt Schema und beschreibt

$$\mathcal{I}(H) = \left\{ A.G_1 \cdots A.G_l \in \mathcal{G} \;\middle|\; \forall 1 \le i \le l : \; (H_i \ne *) \Rightarrow (A.G_i = H_i) \right\}.$$

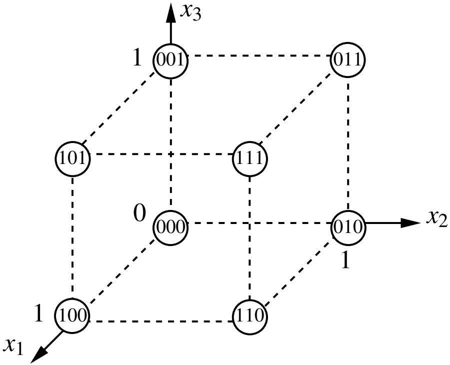
- Ordnung von $H$:

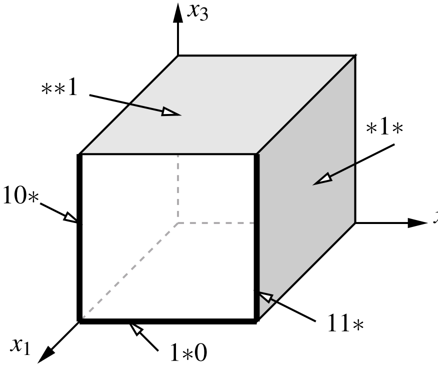
$$o(H) = \# \left\{ i \;\middle|\; (1 \le i \le l) \;\wedge\; (H_i \ne *) \right\}.$$

- definierende Länge von $H$:

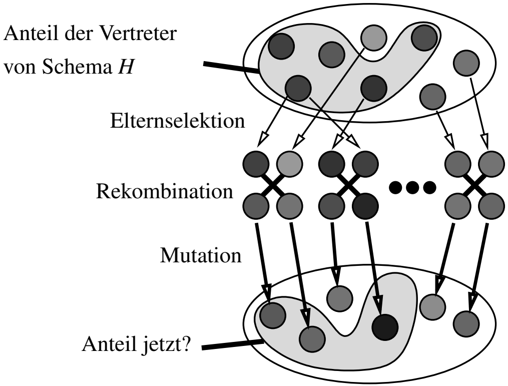
$$\delta(H) = \max \left\{ |i - j| \;\middle|\; (1 \le i, j \le l) \;\wedge\; (H_i \ne *) \;\wedge\; (H_j \ne *) \right\}$$

# Hyperebenen



# Schema-Theorem: Überblick




- welche Eigenschaften (Schemata) vermehren sich besonders stark?

# Schema-Theorem

- GENETISCHER-ALGORITHMUS auf $\mathcal{G} = \mathbb{B}^l$
- für beliebiges Schema $H$ und Population $P(t) = \langle A^{(t,i)} \rangle_{1 \le i \le \mu}$ gilt:

$$\mathrm{Erw}\left[p_H^{(t+1)}\right] \geq p_H^{(t)} \cdot \frac{\overline{F}_H^{(t)}}{\overline{F}^{(t)}} \cdot (1-p_m)^{o(H)} \cdot \left(1 - p_x \cdot \frac{\delta(H)}{l-1} \cdot \left(1 - p_H^{(t)} \cdot \frac{\overline{F}_H^{(t)}}{\overline{F}^{(t)}}\right)\right)$$

$$\text{mit } p_H^{(t)} \;=\; \frac{\#\{1 \le i \le \mu \mid A^{(t,i)}.G \in \mathcal{I}(H)\}}{\mu},$$

- mit durchschnittlicher Güte $\overline{F}^{(t)}$ in der Population und $\overline{F}_H^{(t)}$ der Schema-Vertreter

# Beispiel

| Individuum | Güte |
|---|---|
| 10101... | 3 |
| 01101... | 3 |
| 01100... | 2 |
| 11101... | 4 |
| 11000... | 2 |

| Individuum | Güte |
|---|---|
| 00001... | 1 |
| 10001... | 2 |
| 01001... | 2 |
| 11001... | 3 |
| 01110... | 3 |

# Korollar: Einfaches Schema-Theorem

## Einfaches Schema-Theorem

Durch einige einfache Abschätzungen erhält man die folgende bekanntere Form des Schema-Theorems:

$$\mathrm{Erw}\left[p_H^{(t+1)}\right] \geq p_H^{(t)} \cdot \frac{\overline{F}_H^{(t)}}{\overline{F}^{(t)}} \cdot \left(1 - o(H) \cdot p_m - p_x \cdot \frac{\delta(H)}{l-1}\right).$$

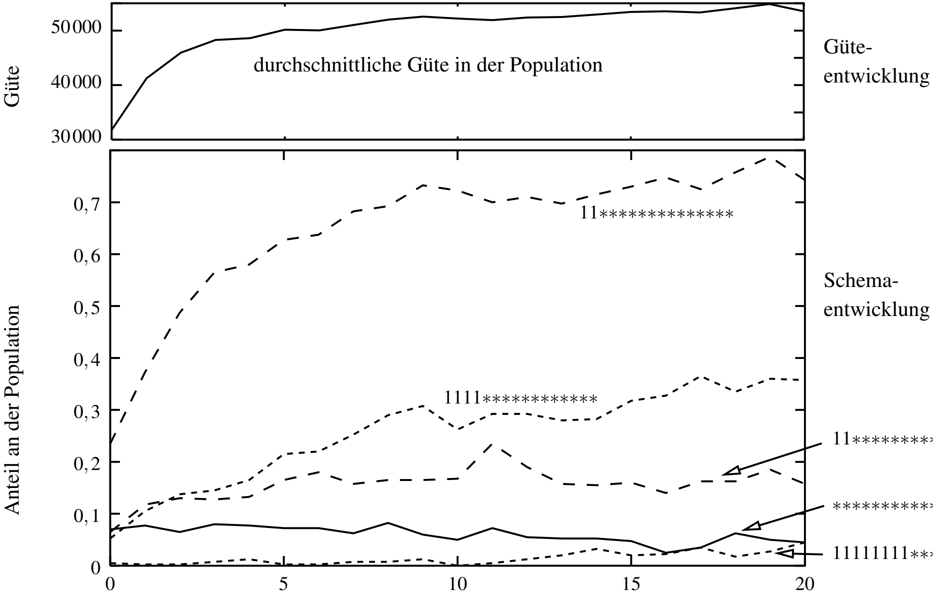
## Bausteine

Schemata mit

- überdurchschnittlicher Güte,
- kleiner definierender Länge und
- geringer Ordnung

# Beispiel

- Rekombinationswahrscheinlichkeit $p_x = 1.0$
- Mutationsrate $p_m = \frac{1}{16}$
- Genotyp: mit 16 Bit enkodierte Zahl (Maximum: 65536)
- Populationsgröße 400


Güte
50 000
40 000
30 000
durchschnittliche Güte in der Population
Güte-entwicklung
Anteil an der Population
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
5
10
15
20
11**************
1111************
Schema-entwicklung
11********
**********
11111111**

# Kritik am Schema-Theorem

- Relevanz der Aussage
  - Ergebnis lässt sich nicht iterativ anwenden
    $\Rightarrow$ Bausteinhypothese (exponentielles Wachstum von Schemata) gilt nicht
  - implizite Annahme: Kombination guter Schemata verbessert die Güte
    $\Rightarrow$ muss nicht gelten!
- Randbedingungen der evolutionären Algorithmen
  - Wahrscheinlichkeitsaussagen in sehr kleinen Mengen sind fragwürdig
  - beobachtete Güte eines Schemas muss nicht der tatsächlichen entsprechen
    Stichpunkt: Konvergenz einzelner Bits

# Überblick

# Beispiel zur Motivation der Formae

- $H_1 = *0*010$ beschreibt $\mathcal{I}(H_1) = \{000010, 001010, 100010, 101010\}$
- dieselben Werte an den Positionen $Pos = \{2,\ 4,\ 5,\ 6\}$
- Äquivalenzrelation

$$A.G \sim_{Pos} B.G :\Leftrightarrow \forall i \in Pos: \ A.G_i = B.G_i$$

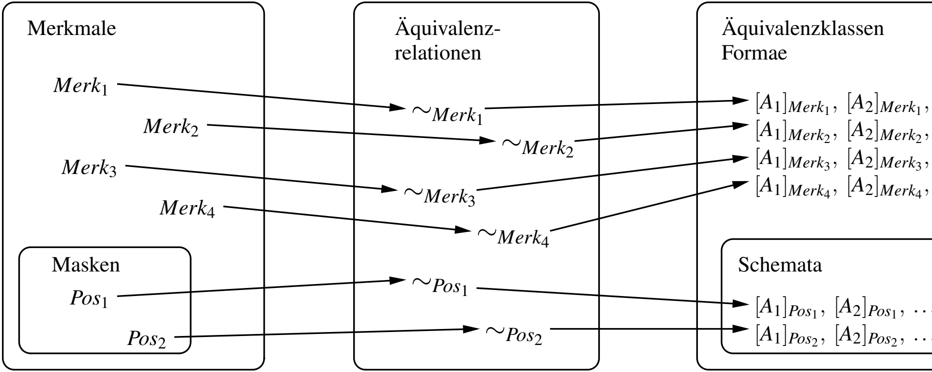
- ergibt 16 Äquivalenzklassen, u.a. gilt

$$\begin{aligned} \mathcal{I}(H_1) = [100010]_{\sim_{Pos}} &= \{000010,\ 001010,\ 100010,\ 101010\} \\ \text{bzw. } [101110]_{\sim_{Pos}} &= \{000110,\ 001110,\ 100110,\ 101110\}. \end{aligned}$$

# Masken und Merkmale

- für $\mathcal{G} = \mathbb{B}^l$ definiert $Pos \subseteq \{1, \ldots, l\}$ eine *Maske*
- Maske mit $\#Pos = k$ definierten Stellen erzeugt $2^{l-k}$ Äquivalenzklassen
- auch beliebige andere Merkmale möglich, z.B. Kante zwischen 4ter und 5ter Stadt

# Schemata in der Verallgemeinerung



# Definition: Formae

- Merkmal $Merk \in \mathcal{M}$ induziert eine Äquivalenzrelation $\sim_{Merk}$
- Damit ergibt sich zu einem Individuum $A.G$ seine *Forma*

$$[A.G]_{\sim_{Merk}} := \{B.G \in \mathcal{G} \mid A.G \sim_{Merk} B.G\}.$$

- *Genauigkeit* = Anzahl der Formae einer Äquivalenzrelation
- Formae $\Delta$ und $\Delta'$ sind *verträglich* wenn

$$\Delta \bowtie \Delta' :\Leftrightarrow \Delta \cap \Delta' \neq \emptyset$$

# Notation Schema/Forma

- sei $\mathcal{G} = M^l$
- Schema $H \in (M \cup \{*\})^l$ kann man beschreiben durch $Merk = \{i \mid (1 \leq i \leq l) \ \wedge \ (H_i \neq *)\}$ und $A.G \in \mathcal{I}(H)$
- Wir schreiben auch: $H = H_{Merk}(A.G)$
- Es gilt ebenso: $\mathcal{I}(H) = [A.G]_{\sim_{Merk}}$

# Allgemeines Schema-Theorem

- sei $P(t) = \langle A^{(t,i)} \rangle_{1 \leq i \leq \mu}$, $\mathcal{Z} = \{\perp\}$ und $\Delta$ ein Forma
- unabhängige Selektion *Sel* von $\mu$ Individuen
- $Rek^{\xi} : (\mathcal{G} \times \mathcal{Z})^2 \to (\mathcal{G} \times \mathcal{Z})^2$
- $Mut^{\xi} : \mathcal{G} \times \mathcal{Z} \to \mathcal{G} \times \mathcal{Z}$
- Dann gilt:

$$\mathrm{Erw}\left[p_H^{(t+1)}\right] \geq p_{sel}(\Delta, t) \cdot p_{\neg mut}(\Delta, t) \cdot p_{\neg rek}(\Delta, t),$$

$$p_{sel}(\Delta, t) = \sum_{A \in P(t) \text{ mit } A.G \in \Delta} \Pr_{\xi \in \Xi}[\widetilde{Sel}^{\xi}(P(t)) = \langle A \rangle]$$

$$p_{\neg mut}(\Delta, t) = \Pr_{\xi \in \Xi}[Mut^{\xi}(A).G \in \Delta \mid A \in P(t) \;\wedge\; A.G \in \Delta]$$

$$p_{\neg rek}(\Delta, t) = \Pr_{\xi \in \Xi, B.G \in \mathcal{G}}[Rek^{\xi}(A, B).G \in \Delta \mid A \in P(t) \;\wedge\; A.G \in \Delta].$$

# Bedingungen für Forma-Wachstum

- *minimale Redundanz* der Dekodierfunktion – bestenfalls: Bijektion
  sonst: aus $dec(A.G) = dec(B.G)$ folgt $[A.G]_{\sim_{Merk}} = [B.G]_{\sim_{Merk}}$
- *Ähnlichkeit in Formae* geringer Genauigkeit: Individuen haben ähnliche Güte
- *Abschluss gegen Schnitt*:

$$\forall \text{ Formae } \Delta, \Delta' \; \exists \text{ Forma } \Delta'' : \; \Delta \cap \Delta' = \Delta''$$

# Bedingungen für Forma-Wachstum

- *Verträglichkeit Forma/Rekombination:*

$$\forall\Delta\ \forall A, B \text{ mit } A.G, B.G \in \Delta\ \forall\xi \in \Xi: \ \ Rek^{\xi}(A, B).G \in \Delta$$

- *Übertragung der Gene* für Forma mit minimaler Genauigkeit

$$\forall A, B\ \forall\xi \in \Xi\ \forall \text{ minimales } \Delta: \\ Rek^{\xi}(A, B).G \in \Delta \ \Rightarrow\ (A.G \in \Delta\ \vee\ B.G \in \Delta)$$

- *Verschmelzungseigenschaft*

$$\forall\Delta, \Delta' \text{ mit } \Delta \bowtie \Delta'\ \forall A \text{ mit } A.G \in \Delta\ \forall B \text{ mit } B.G \in \Delta'\ \exists\xi \in \Xi \\ Rek^{\xi}(A, B).G \in \Delta \cap \Delta'$$

# Überblick

# „Fehlendes“ Schema-Theorem

- laut Altenberg: Es fehlt ein Schema-Theorem mit einer Aussage zum Suchfortschritt
- Ausgangssituation:
  - Genetischer-Algorithmus mit fitnessprop. Selektion und ohne Mutation
  - allgemeine Rekombination $Merk \subseteq \{0, \ldots, l\}$
  - mögliche Eltern von Individuum $A$: $H_{Merk}(A.G)$ und $H_{\widetilde{Merk}}(A.G)$ mit $\widetilde{Merk} = \{1, \ldots, l\} \setminus Merk$
  - kommen mit Wahrscheinlichkeiten $p_{Merk}$ vor

# Verknüpfung Schema/Suchfortschritt

Auf zwei Ebenen:

- am Ende im Ergebnis: Wie ändert sich die durchschnittliche Güte?
- durch Verknüpfen von Eltern und Kindgüte

$$\mathrm{Cov}\left[A.F, \frac{\overline{F}^{(t)}_{H_{Merk}(A.G)} \cdot \overline{F}^{(t)}_{H_{\widetilde{Merk}}(A.G)}}{\left(\overline{F}^{(t)}\right)^2}\right]$$

# Lemma

Es gilt:

$$\mathrm{Cov}\left[A.F, \frac{\overline{F}^{(t)}_{H_{Merk}(A.G)} \cdot \overline{F}^{(t)}_{H_{\widetilde{Merk}}(A.G)}}{(\overline{F}^{(t)})^2}\right]$$

$$= \sum_{A.G \in \mathcal{G}} \left(A.F - \overline{F}^{(t)}\right) \cdot \frac{\overline{F}^{(t)}_{H_{Merk}(A.G)} \cdot \overline{F}^{(t)}_{H_{\widetilde{Merk}}(A.G)}}{(\overline{F}^{(t)})^2} \cdot p_{A.G} \;,$$

mit Häufigkeit $p_{A.G}$ von Individuum $A$ in der Population

# Beispiel:

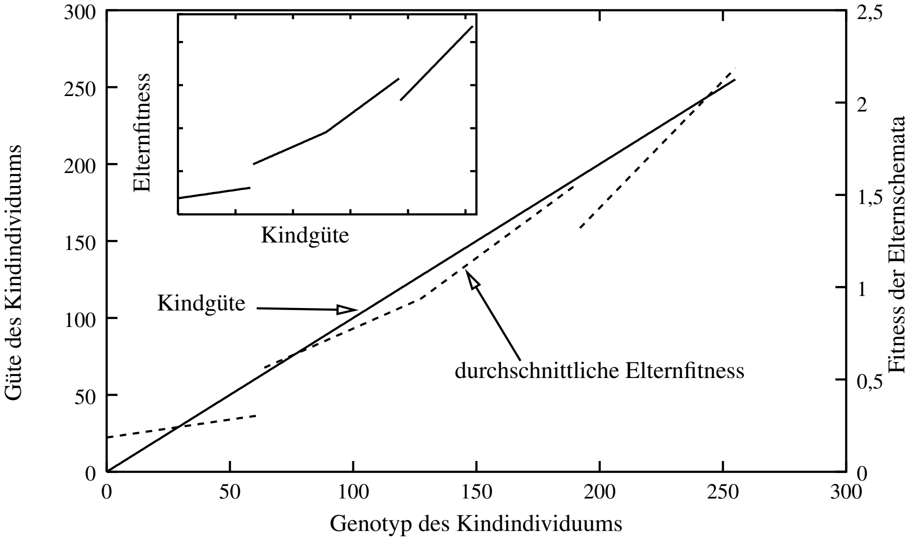
- $\mathcal{G} = \mathbb{B}^8$ enkodiert standardbinär $\{0, \ldots, 255\}$

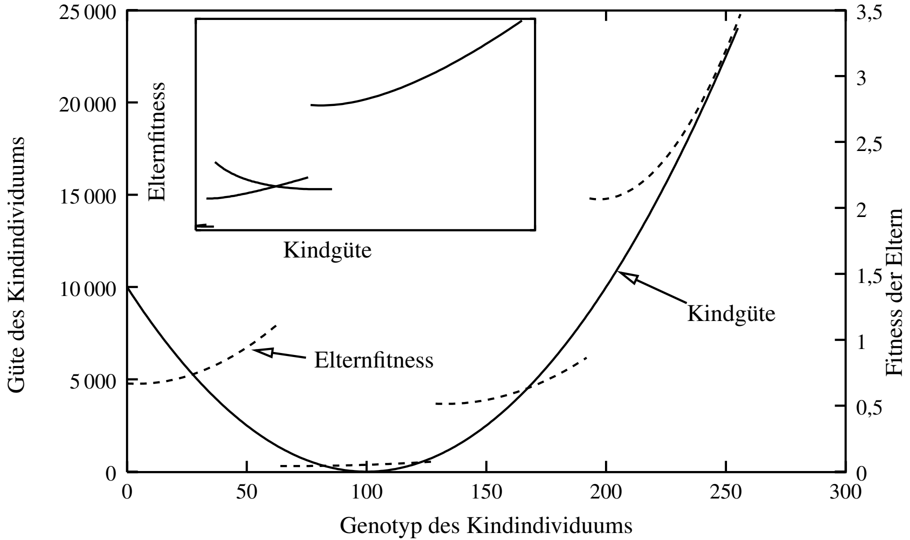
Crossover-Punkt

$$\begin{array}{|c|c||c|c|c|c|c|c|}\hline 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ \hline \end{array} \quad \text{Kindindividuum}$$

- Rekombination:

$$\left.\begin{array}{cccccccc} 1 & 0 & * & * & * & * & * & * \\ * & * & 0 & 1 & 0 & 1 & 1 & 0 \end{array}\right\} \begin{array}{l}\text{mögliche} \\ \text{Eltern}\end{array}$$


Elternfitness
Kindgüte
Kindgüte
durchschnittliche Elternfitness
Güte des Kindindividuums
Fitness der Elternschemata
Genotyp des Kindindividuums
300
250
200
150
100
50
0
2,5
2
1,5
1
0,5
0
0
50
100
150
200
250
300


25 000
20 000
15 000
10 000
5 000
0
Güte des Kindindividuums
0
50
100
150
200
250
300
Genotyp des Kindindividuums
3,5
3
2,5
2
1,5
1
0,5
0
Fitness der Eltern
Elternfitness
Kindgüte
Elternfitness
Kindgüte

# Satz: „Fehlendes“ Schema-Theorem

Für einen genetischen Algorithmus nur mit Rekombination gilt:

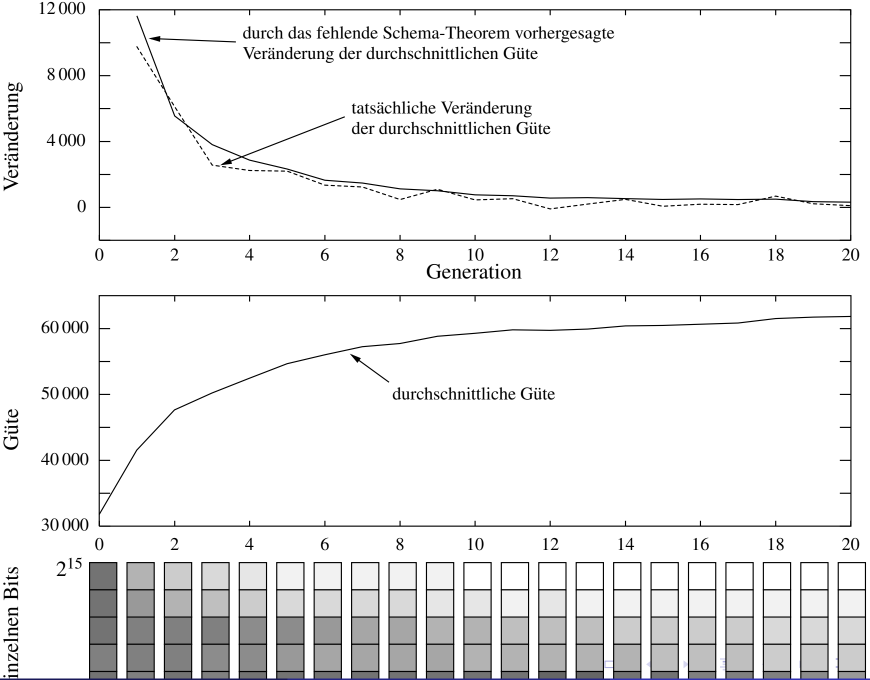
$$\mathrm{Erw}\left[\overline{F}^{(t+1)}\right]-\overline{F}^{(t)} = \sum_{Merk\subseteq\{1,\ldots,l\}} p_{Merk}\cdot\left(\mathrm{Cov}\left[A.F, \frac{\overline{F}_{H_{Merk}(A.G)}\cdot\overline{F}_{H_{\widetilde{Merk}}(A.G)}}{(\overline{F}^{(t)})^2}\right]\right.$$

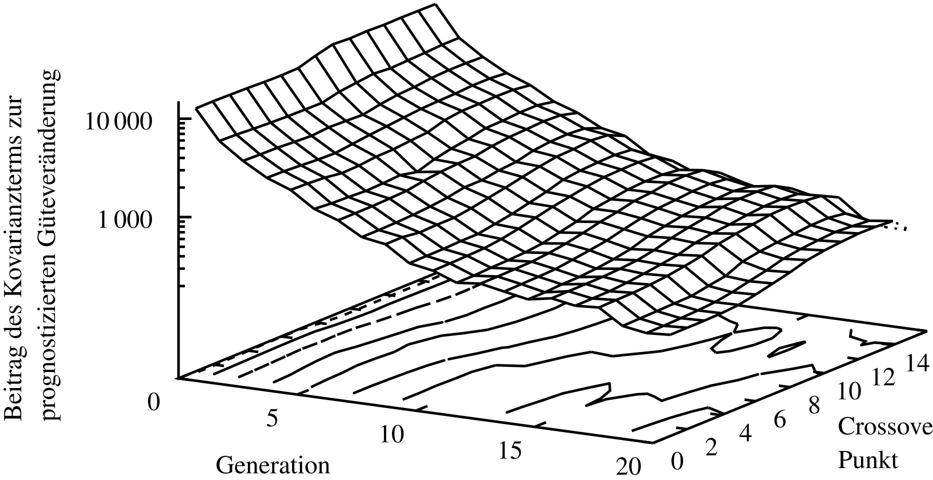
$$\left.-\sum_{A.G\in\mathcal{G}}\left(p_{A.G}-p_{H_{Merk}(A.G)}\cdot p_{H_{\widetilde{Merk}}(A.G)}\right)\cdot\left(A.F-\overline{F}^{(t)}\right)\cdot\frac{\overline{F}_{H_{Merk}(A.G)}\cdot\overline{F}_{H_{\widetilde{Merk}}(A.G)}}{(\overline{F}^{(t)})^2}\right)$$

mit Häufigkeiten $p_{H_{Merk}(A.G)}$ und $p_{H_{\widetilde{Merk}}(A.G)}$ der erzeugenden Schemata von $A.G$ und durchschnittlichen Schematagütewerten $\overline{F}_{H_{Merk}(A.G)}$ und $\overline{F}_{H_{\widetilde{Merk}}(A.G)}$ .

# Beispiel

- 1-Punkt-Crossover
- Rekombinationswahrscheinlichkeit $p_x = 1.0$
- keine Mutation
- Genotyp: mit 16 Bit enkodierte Zahl (Maximum: 65536)
- Populationsgröße 400


12 000
8 000
4 000
0
Veränderung
durch das fehlende Schema-Theorem vorhergesagte
Veränderung der durchschnittlichen Güte
tatsächliche Veränderung
der durchschnittlichen Güte
0 2 4 6 8 10 12 14 16 18 20
Generation
60 000
50 000
40 000
30 000
Güte
durchschnittliche Güte
0 2 4 6 8 10 12 14 16 18 20
$2^{15}$
inzelnen Bits


Beitrag des Kovarianzterms zur
prognostizierten Güteveränderung
10 000
1 000
0
5
10
15
20
Generation
0
2
4
6
8
10
12
14
Crossove
Punkt