



OTTO VON GUERICKE  
UNIVERSITÄT  
MAGDEBURG

INF

FAKULTÄT FÜR  
INFORMATIK

# Evolutionäre Algorithmen

## Kapitel 6: Das Schematheorem, No Free Lunch Theorem

Sanaz Mostaghim

Intelligente Systeme

SS 2016

# Outline

## Motivation

Schemata

Einfluss der Selektion

Einfluss der Mutation

Einfluss des Crossover

## Herleitung des Schematheorems

## Baustein-Hypothese

## Zusammenfassung

## No-Free-Lunch-Theorem

# Warum funktionieren EAs?

**Ansatz** von [Holland, 1975]:

- ▶ betrachte Chromosomenschemata (d.s. nur teilweise festgelegte Chromosomen)
- ▶ untersuche, wie sich Zahl der Chromosomen, die zu Schema passen, über Generationen hinweg entwickelt

**Ziel:** zumindest grobe stochastische Aussage darüber, wie  $\Omega$  von EAs durchforstet wird

zur **Vereinfachung** der Darstellung: Beschränkung auf

- ▶ Bitfolgen (Chromosomen aus Nullen und Einsen) mit fester Länge  $L$
- ▶ fitnessproportionale Selektion (Glücksradauswahl)
- ▶ Binär-Mutation (Mutieren jedes Bits mit  $W$ 'keit)
- ▶ 1-Punkt-Crossover (Durchschneiden an einer Stelle und Vertauschen)

# Algorithm 1 Genetischer Algorithmus

**Input:** Zielfunktion  $F$

```
1:  $t \leftarrow 0$ 
2:  $P(t) \leftarrow$  erzeuge Population mit  $\mu$  Individuen          /*  $\mu$  muss gerade sein */
3: bewerte  $P(t)$  durch  $F$ 
4: while Terminierungsbedingung nicht erfüllt {
5:    $P'(t) \leftarrow$  selektiere  $\mu$  Individuen  $A^{(1)}, \dots, A^{(\mu)}$  aus  $P(t)$  mittels Glücksradauswahl
6:    $P'' \leftarrow \emptyset$ 
7:   for  $i \leftarrow 1, \dots, \frac{\mu}{2}$  {
8:      $u \leftarrow$  wähle Zufallszahl gemäß  $U([0, 1])$ 
9:     if  $u \leq p_x$  {                                           /* Rekombinationsw'keit  $p_x$  */
10:       $B, C \leftarrow$  1-Punkt-Crossover( $A^{(2i-1)}, A^{(2i)}$ )
11:    } else {
12:       $B \leftarrow A^{(2i-1)}$ 
13:       $C \leftarrow A^{(2i)}$ 
14:    }
15:     $B \leftarrow$  Binär-Mutation( $B$ )
16:     $C \leftarrow$  Binär-Mutation( $C$ )
17:     $P'' \leftarrow P'' \cup \{B, C\}$ 
18:  }
19:  bewerte  $P''$  durch  $F$ 
20:   $t \leftarrow t + 1$ 
21:   $P(t) \leftarrow P''$ 
22: }
23: return bestes Individuum aus  $P(t)$ 
```

# Schemata

## Definition (Schema)

Ein **Schema**  $h$  ist eine Zeichenkette der Länge  $L$  über dem Alphabet  $\{0, 1, *\}$ , d.h.  $h \in \{0, 1, *\}^L$ .

Das Zeichen  $*$  heißt **Jokerzeichen** oder **Don't-Care-Symbol**.

## Definition (Passung)

Ein Chromosom  $c \in \{0, 1\}^L$  **passt zu einem Schema**  $h \in \{0, 1, *\}^L$ , in Zeichen:  $c \triangleleft h$ , wenn es mit  $h$  an allen Stellen übereinstimmt, an denen  $h$  eine 0 oder eine 1 enthält. (Stellen, an denen ein  $*$  steht, bleiben unberücksichtigt.)

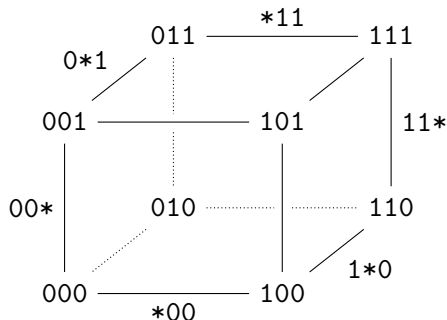
# Schemata: Ein Beispiel

$h = \quad **0*11*10*$       Schema der Länge 10  
 $c_1 = \quad 1100111100$       passt zu  $h$ ,      also  $c_1 \triangleleft h$   
 $c_2 = \quad 1111111111$       passt nicht zu  $h$ , also  $c_2 \not\triangleleft h$

- ▶ es gibt  $2^L$  Chromosomen und  $3^L$  Schemata
- ▶ jedes Chromosom passt zu  $\sum_{i=0}^L \binom{L}{i} = 2^L$  Schemata
- ▶ Population der Größe  $\mu$  kann bis zu  $\mu 2^L$  Schemata haben (normalerweise aber viel weniger aufgrund ähnlicher Chromosomen)
- ▶ Betrachtung eines Chromosoms  $\hat{=}$  Betrachtung vieler Schemata
- ▶ **dadurch: impliziter Parallelismus**

# Schemata: Hyperebenen

Jedes Schema beschreibt Hyperebene im Hypereinheitswürfel (aber nur Ebenen, die parallel oder senkrecht zu Achsen stehen)



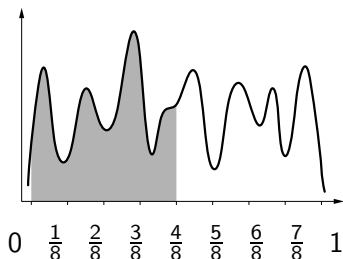
- Beispiele:**  $*00 \hat{=}$  Kante von 000 nach 100 (vorne unten)  
 $0^{**} \hat{=}$  linke Würfelfläche  
 $*** \hat{=}$  gesamter Würfel

# Schemata: Wertebereiche von Funktionen

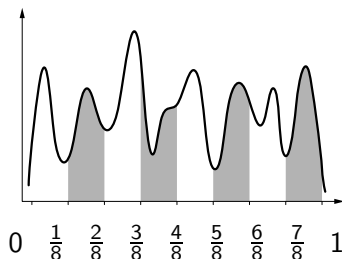
geg.: reelle Funktion  $f : x \in [0, 1] \rightarrow \mathbb{R}$

Annahme: Binärkodierung von  $x$  (kein Gray-Kode)

Schema  $\hat{=}$  „Streifenmuster“ (periodischer Fkt.) in  $\text{dom}(f) = [0, 1]$



Schema  $0^{**}\dots^*$



Schema  $^{**}1^*\dots^*$

Schemata mit Gray-Kodierung: siehe Übungsaufgabe



# Einfluss der Selektion

- ▶ Auswirkungen von **Selektion** und **genetischen Operatoren** (Mutation und Crossover) untersuchen
- ▶ Verfolgung der Vermehrung von Chromosomen, die zu Schema passen
- ▶ für Selektion: welche Fitness haben Chromosomen, die zu Schema  $h$  passen? Ansatz: Mittelung über alle Chromosomen

## Definition (Mittlere Fitness)

Die **mittlere relative Fitness** der Chromosomen, die in der Generation  $P(t)$  zum Schema  $h$  passen, ist

$$f_{\text{rel}}(h) = \frac{\sum_{A \in P(t), A.G \triangleleft h} A.F_{\text{rel}}}{|\{A \in P(t) \mid A.G \triangleleft h\}|}$$

# Einfluss der Selektion

Die durchschnittliche Anzahl Nachkommen eines zu Schema  $h$  passenden Chromosoms ist

$$f_{\text{rel}}(h) \cdot |P|$$

Die zu erwartende Zahl Chromosomen, die nach Auswahl zu Schema  $h$  passen, ist

$$(\text{Zahl vorher passender Chromosomen}) \cdot f_{\text{rel}}(h) \cdot |P|$$

# Einfluss der Selektion

Weitere Betrachtungen zur relativen Fitness eines Schemas:

$$\begin{aligned} f_{\text{rel}}(h) \cdot |P| &= \frac{\sum_{A \in P(t), A.G \triangleleft h} A.F_{\text{rel}}(c)}{|\{A \in P(t) \mid A.G \triangleleft h\}|} \cdot |P| \\ &= \frac{\sum_{A \in P(t), A.G \triangleleft h} \frac{A.F}{\sum_{B \in P(t)} B.F}}{|\{A \in P(t) \mid A.G \triangleleft h\}|} \cdot |P| \\ &= \frac{\frac{\sum_{A \in P(t), A.G \triangleleft h} A.F}{|\{A \in P(t) \mid A.G \triangleleft h\}|}}{\frac{\sum_{B \in P(t)} B.F}{|P|}} = \frac{\overline{f_t(h)}}{\overline{f_t}} \end{aligned}$$

$\overline{f_t(h)}$  mittlere Fitness der in  $P(t)$  zu  $h$  passenden Chromosomen

$\overline{f_t}$  mittlere Fitness aller Chromosomen der  $t$ -ten Generation

Die mittlere Anzahl Nachkommen wird durch das Verhältnis der mittleren Güte eines Schemas zur Gesamtdurchschnittsgüte berechnet

# Einfluss der Mutation

Um zu bestimmen, unter welcher Wahrscheinlichkeit die Passung eines Schemas verloren geht bzw. erhalten bleibt, werden spezielle Maße benötigt.

## Definition (Ordnung (für die 1-Bit- und Binär-Mutation))

Die **Ordnung** eines Schemas  $h$  ist die Anzahl der Nullen und Einsen in  $h$ , also  $\text{ord}(h) = \#0 + \#1 = L - \#*$  ( $\#$ : Anzahl des Auftretens von).

**Beispiel:**  $\text{ord}(**0*11*10*) = 5$

# Einfluss der Mutation

Passung zu Schema  $h$  bleibt erhalten...

- ▶ durch Binär-Mutation mit W'keit  $(1 - p_m)^{\text{ord}(h)}$
- ▶ durch 1-Bit-Mutation mit  
W'keit  $1 - \frac{\text{ord}(h)}{L}$ , falls Bit umgekehrt wird,  
W'keit  $1 - \frac{\text{ord}(h)}{2L}$ , falls neues Bit zufällig bestimmt wird

## Erläuterung:

- ▶ Binär-Mutation invertiert jedes Bit wird mit W'keit  $p_m$  und mit W'keit  $(1 - p_m)$  nicht
- ▶ 1-Bit-Mutation wählt eines der  $L$  Gene eines Chromosoms der Länge  $L$  mit gleicher W'keit

# Einfluss des Crossover

## Definition (Definierende Länge (für das 1-Punkt-Crossover))

Die **definierende Länge** eines Schemas  $h$  ist die Differenz zwischen der Positionsnummer der letzten 0/1 und der Positionsnummer der ersten 0/1 in  $h$ .

**Beispiel:**  $dl(**0*11*10*) = 9 - 3 = 6$

# Einfluss des Crossover

- ▶ 1-Punkt-Crossover: Schnittpunkt liegt mit  $W$ 'keit  $\frac{dI(h)}{L-1}$  so, dass 2 Nicht-Jokerzeichen voneinander getrennt werden

## Erläuterung:

- ▶ 1-Punkt-Crossover: Chromosomen der Länge  $L$  bieten  $L - 1$  mögliche Schnittpunkte (alle gleichwahrscheinlich)
- ▶  $dI(h)$  dieser Schnittpunkte liegen so, dass im Schema festgelegte Gene in verschiedene Nachkommen gelangen
- ▶ Passung könnte, muss aber nicht zwangsläufig, verloren gehen

# Outline

Motivation

## Herleitung des Schematheorems

Selektion

Crossover

Mutation

Das Schematheorem

Baustein-Hypothese

Zusammenfassung

No-Free-Lunch-Theorem



# Definitionen

## Definition (Erwartungswert passender Chromosomen)

$N(h, t)$  ist der Erwartungswert der Anzahl Chromosomen, die in der  $t$ -ten Generation zum Schema  $h$  passen.

## Definition (Erwartungswert nach Selektion)

$N(h, t + \Delta t_s)$  ist der Erwartungswert der Anzahl Chromosomen, die in der  $t$ -ten Generation nach Selektion zum Schema  $h$  passen.

# Definitionen

## Definition (Erwartungswert nach Crossover)

$N(h, t + \Delta t_s + \Delta t_x)$  ist der Erwartungswert der Anzahl Chromosomen, die in der  $t$ -ten Generation nach Selektion und Crossover zum Schema  $h$  passen.

## Definition (Erwartungswert nach Mutation)

$N(h, t + \Delta t_s + \Delta t_x + \Delta t_m) = N(h, t + 1)$  ist der Erwartungswert der Anzahl Chromosomen, die in der  $t$ -ten Generation nach Selektion, Crossover und Mutation (und damit in der  $(t + 1)$ -ten Generation) zum Schema  $h$  passen.

- ▶ **gesucht:** Zusammenhang zwischen  $N(h, t)$  und  $N(h, t + 1)$

# Selektion

**Vorgehen:** betrachten schrittweise Auswirkungen von Selektion, Crossover und Mutation anhand mittlerer Fitness, Ordnung und definierender Länge eines Schemas

- ▶ **Auswirkungen der Selektion:** durch mittlere Fitness beschreibbar

$$N(h, t + \Delta t_s) = N(h, t) \cdot f_{\text{rel}}(h) \cdot |P|$$

$N(h, t) \cdot f_{\text{rel}}(h)$  W'keit, dass zum Schema  $h$  passendes Chromosom ausgewählt

$f_{\text{rel}}(h) \cdot |P|$  durchschnittliche Anzahl Nachkommen eines zu Schema  $h$  passenden Chromosoms

- ▶ *beachte:* die relative Fitness  $f_{\text{rel}}(h)$  nicht exakt bestimmt, da die zu  $h$  passenden Chromosomen nur als Erwartungswert bekannt sind

# Crossover

Auswirkungen des Crossover: beschrieben durch

$$N(h, t + \Delta t_s + \Delta t_x) = \underbrace{(1 - p_x) \cdot N(h, t + \Delta t_s)}_{E_A} + \underbrace{p_x \cdot N(h, t + \Delta t_s) \cdot (1 - p_{\text{loss}})}_{E_B} + C$$

$p_x$  W'keit eines Crossover

$p_{\text{loss}}$  W'keit, dass durch 1-Punkt-Crossover Passung eines Chromosoms zu Schema  $h$  verloren geht

$E_A$  Erwartungswert der Anzahl Chromosomen, die zu Schema  $h$  passen und *nicht* am Crossover teilnehmen

$E_B$  Erwartungswert der Anzahl Chromosomen, die am Crossover teilnehmen und deren Passung zu  $h$  dadurch nicht verloren geht

$C$  Gewinne an Chromosomen, die zu Schema  $h$  passen, durch...  
(siehe Übungsaufgabe)

# Betrachtungen zur W'keit $p_{\text{loss}}$

**Beispiele:**

$$h = \text{**0*|1*1*} \quad \text{**0*1*1*} = h$$

$$h \triangleright c_1 = 0000|1111 \rightarrow 00000000 = c'_1 \not\triangleleft h$$

$$h \not\triangleright c_2 = 1111|0000 \rightarrow 11111111 = c'_2 \not\triangleleft h$$

$$h = \text{**0*|1*1*} \quad \text{**0*1*1*} = h$$

$$h \triangleright c_1 = 0000|1111 \rightarrow 00001010 = c'_1 \triangleleft h$$

$$h \triangleright c_2 = 1101|1010 \rightarrow 11011111 = c'_2 \triangleleft h$$

**folglich:**

$$p_{\text{loss}} \leq \underbrace{\frac{dl(h)}{L-1}}_{Pr_A = p_x^*(h)} \cdot \left( 1 - \underbrace{\frac{N(h, t + \Delta t_s)}{|P|}}_{Pr_B} \right)$$

$Pr_A$  W'keit, dass Schnittpunkt zwischen festgelegte Gene fällt

$Pr_B$  W'keit, dass 2. Chromosom zu Schema  $h$  passt

**Frage:** Warum gilt nur  $\leq$  und nicht  $=$ ? (siehe Übungsaufgabe)

# Crossover

Einsetzen des Ausdrucks für  $p_{\text{loss}}$  liefert:

$$\begin{aligned} & N(h, t + \Delta t_s + \Delta t_x) \\ & \geq (1 - p_x) \cdot N(h, t + \Delta t_s) \\ & \quad + p_x \cdot N(h, t + \Delta t_s) \cdot \left(1 - \frac{dl(h)}{L-1} \cdot \left(1 - \frac{N(h, t + \Delta t_s)}{|P|}\right)\right) \\ & = N(h, t + \Delta t_s) \left(1 - p_x + p_x \cdot \left(1 - \frac{dl(h)}{L-1} \cdot \left(1 - \frac{N(h, t + \Delta t_s)}{|P|}\right)\right)\right) \\ & = N(h, t + \Delta t_s) \cdot \left(1 - p_x \frac{dl(h)}{L-1} \cdot \left(1 - \frac{N(h, t + \Delta t_s)}{|P|}\right)\right) \\ & \stackrel{(*)}{=} N(h, t) \cdot f_{\text{rel}}(h) \cdot |P| \cdot \left(1 - p_x \frac{dl(h)}{L-1} \cdot (1 - N(h, t) \cdot f_{\text{rel}}(h))\right) \end{aligned}$$

Schritt (\*): zweimaliges Nutzen der vorher abgeleiteten Beziehung

$$N(h, t + \Delta t_s) = N(h, t) \cdot f_{\text{rel}}(h) \cdot |P|$$

# Mutation

**Auswirkungen der Binär-Mutation:** durch Ordnung beschrieben

$$\begin{aligned}N(h, t + 1) &= N(h, t + \Delta t_s + \Delta t_x + \Delta t_m) \\ &= N(h, t + \Delta t_s + \Delta t_x) \cdot (1 - p_m)^{\text{ord}(h)}\end{aligned}$$

*Erläuterung:* damit Passung nicht verloren geht, darf keines der  $\text{ord}(h)$  Gene verändert werden, die in Schema  $h$  festgelegt

**alternative Modelle** möglich, z.B.:

genau eine Bit-Änderung pro Chromosom  $\Rightarrow$  1-Bit-Mutation

$$\begin{aligned}N(h, t + 1) &= N(h, t + \Delta t_s + \Delta t_x + \Delta t_m) \\ &= N(h, t + \Delta t_s + \Delta t_x) \cdot \left(1 - \frac{\text{ord}(h)}{L}\right)\end{aligned}$$

# Das Schematheorem

Mit Binär-Mutation gilt

$$N(h, t + 1) \geq f_{\text{rel}}(h) \cdot |P| \cdot \left(1 - p_x \frac{dl(h)}{L-1} \cdot (1 - N(h, t) \cdot f_{\text{rel}}(h))\right) \cdot (1 - p_m)^{\text{ord}(h)} \cdot N(h, t)$$

Das Einsetzen des Fitnessverhältnisses liefert

$$N(h, t + 1) \geq \frac{\overline{f_t(h)}}{\overline{f_t}} \left(1 - p_x \frac{dl(h)}{L-1} \left(1 - \frac{N(h, t)}{|P|} \cdot \frac{\overline{f_t(h)}}{\overline{f_t}}\right)\right) \cdot (1 - p_m)^{\text{ord}(h)} \cdot N(h, t)$$

**Interpretation:** Schemata mit

- ▶ überdurchschnittlicher mittlerer Bewertung,
- ▶ kurzer definierender Länge und
- ▶ geringer Ordnung

vermehrten sich besonders stark (etwa exponentiell)



# Outline

Motivation

Herleitung des Schematheorems

**Baustein-Hypothese**

Kritik an der Baustein-Hypothese

Zusammenfassung

No-Free-Lunch-Theorem

# Baustein-Hypothese

- ▶ Ein GA durchsucht  $\Omega$  besonders gut in Hyperebenen (also Schemata mit hoher mittlerer Fitness, kleiner definierender Länge und geringer Ordnung)
- ▶ Chromosomen vermehren sich in solchen Regionen am stärksten
- ▶ Diese Schemata heißen **Bausteine** (engl.: **building blocks**), daher der Name **Baustein-Hypothese**
- ▶ **Beachte:** Diese Form der Hypothese gilt nur für Bitfolgen, fitnessproportionale Selektion, Binär-Mutation und 1-Punkt-Crossover
- ▶ Bei anderen genetischer Operatoren werden Bausteine u.U. durch andere Eigenschaften charakterisiert
- ▶ Eine hohe mittlere Fitness ist jedoch stets charakterisierende Eigenschaft, da jedes Selektionsverfahren Chromosomen mit hoher Fitness bevorzugt

# Baustein-Hypothese

GA arbeitet optimal, wenn aus **kurzen Schemata mit geringer Ordnung** (sog. Bausteine) Schemata mit höherer Fitness entstehen.

- ▶ **aber:** Kombination von Bausteine führt zu höherer Ordnung und definierender Länge des Bausteins
- ▶ dadurch sind neue Bausteine viel mehr gefährdet, durch genetische Operatoren zerissen zu werden
- ▶ Arbeitsweise eines GA ist immer noch nicht ausreichend erforscht

# Outline

Motivation

Herleitung des Schematheorems

Baustein-Hypothese

**Zusammenfassung**

No-Free-Lunch-Theorem

# Zusammenfassung

Schematheorem gilt nur für **ein Schema** unabhängig von anderen Schemata in Population.

- ▶ Andere Schemata werden sich auch vermehren
- ▶ Mit der Zeit: Konvergenz der Population, wodurch der Selektionsdruck sinkt
- ▶ relative Fitness eines Schemas konvergiert gegen  $1/|P|$   
(Gleiche Fitnesswerte für alle Individuen)
- ▶ schließlich: erwartete Anzahl Kopien nimmt ab durch Zerstörung durch genetische Operatoren

# Zusammenfassung

- ▶ implizite Annahme: kaum Wechselwirkungen zwischen Genen (geringe *Epistasie*), also Fitness von Chromosomen, die zu Schema passen, sei sehr ähnlich
- ▶ implizite Annahme: interagierende Gene liegen im Chromosom eng zusammen für kleine Bausteine
  - ▶ Einwand betrifft nur Beschränkung auf 1-Punkt-Crossover und nicht Ansatz an sich
  - ▶ andere Maße als definierende Länge möglich, die operationenspezifisch sind

# Outline

Motivation

Herleitung des Schematheorems

Baustein-Hypothese

Zusammenfassung

## **No-Free-Lunch-Theorem**

Formale Definitionen

Das Theorem

Konsequenzen und Zusammenfassung

# No Free Lunch Theorem

## Voraussetzungen

Suchraum  $\Omega$

$\mathcal{F}$  Raum aller Optimierungsprobleme (Zielfunktionen)

$\Omega$  und  $\mathcal{F}$  seien endlich

Unwissenheit über Optimierungsproblem

- ▶ Gleichverteilung aller dieser Probleme
- ▶ jedes Problem  $F \in \mathcal{F}$  tritt mit Wahrscheinlichkeit  $\frac{1}{|\mathcal{F}|}$  auf
- ▶ weitere Vereinfachung
  - ▶  $\forall F \in \mathcal{F}$  gilt  $F : \Omega \mapsto \mathbb{R}$
  - ▶  $\forall F \in \mathcal{F}$  sind auf selbem Suchraum  $\Omega$  definiert
- ▶  $\mathcal{A}$  sei Menge aller Optimierungsalgorithmen, die auf  $\Omega$  arbeiten



# No Free Lunch Theorem

## Charakterisierung eines Algorithmus

- ▶ welche Ind. werden in welcher Reihenfolge auf  $F \in \mathcal{F}$  betrachtet
- ▶ lediglich  $n$  Auswertungen zur Optimierung möglich:
  - ▶ Optimierung $_{F,n} : \mathcal{A} \mapsto \Omega^n$
- ▶ bei jeder Optimierung bewertet Algorithmus ein Individuum  $1x$
- ⇒ Optimierung $_{F,n}(\text{Alg})$  enthält insgesamt  $n$  untersch. Individuen
  - ▶ jeder Algorithmus Alg sei deterministisch
- ⇒ Optimierung $_{F,n}(\text{Alg})$  auch eindeutig

Für Problem  $F \in \mathcal{F}$ , Optimierungsproblem  $\text{Alg} \in \mathcal{A}$  und  $n \in \mathbb{N}$  ist

$$\text{Optimierung}_{F,n}(\text{Alg}) = (y_1, \dots, y_n) \in \Omega^n$$

mit  $y_i \neq y_j$  für  $i \neq j$  wobei  $y_k$  Individuum, das Alg mit  $F$  als  $k$ -tes Element untersucht.

## Vergleich von $\text{Alg}_1, \text{Alg}_2 \in \mathcal{A}$

- ▶ mittels Leistungsmaß  $\text{QuAlg}$  (Qualität eines Algorithmus)
  - ▶ definiert mit Optimierung  $f_{F,n}(\text{Alg}) = (y_1, \dots, y_n)$  durch  $q_n : \mathbb{R}^n \mapsto \mathbb{R}$  als

$$\text{QuAlg}_{F,n}(\text{Alg}) = q_n(F(y_1), \dots, F(y_n))$$

- ▶ z.B. durchschnittliche bzw. beste Güte oder
  - ▶ Anzahl der benötigten Auswertungen bis Optimum erreicht
- ▶ zu erwartende Leistung  $E$  der  $n$  ersten Auswertungen von Alg auf beliebigem Problem

$$E \left[ \text{QuAlg}_{F,n}(\text{Alg}) \mid F \in \mathcal{F} \right] = \frac{1}{\#\mathcal{F}} \sum_{F \in \mathcal{F}} \text{QuAlg}_{F,n}(\text{Alg})$$

⇒ Mittel über alle möglichen Probleme

# No Free Lunch

## Satz (No free lunch)

Für je zwei Algorithmen  $Alg_1, Alg_2 \in \mathcal{A}$  und die Klasse aller Probleme  $\mathcal{F}$  gilt bezüglich eines Leistungsmaßes  $QuAlg$ :

$$E \left[ QuAlg_{F,n}(Alg_1) \mid F \in \mathcal{F} \right] = E \left[ QuAlg_{F,n}(Alg_2) \mid F \in \mathcal{F} \right]$$

# Konsequenzen

Kein Alg. ist im Mittel über alle möglichen Probleme einem Anderen überlegen

Gäbe es Algorithmus, der auf  $\mathcal{F}' \subset \mathcal{F}$  überlegen, also

$$E \left[ \text{QuAlg}_{F,n}(\text{Alg}_1) \mid F \in \mathcal{F}' \right] < E \left[ \text{QuAlg}_{F,n}(\text{Alg}_2) \mid F \in \mathcal{F}' \right]$$

- ▶ dann folgt sofort

$$E \left[ \text{QuAlg}_{F,n}(\text{Alg}_1) \mid F \in \mathcal{F} \setminus \mathcal{F}' \right] > E \left[ \text{QuAlg}_{F,n}(\text{Alg}_2) \mid F \in \mathcal{F} \setminus \mathcal{F}' \right]$$

- ▶ für jeden Algorithmus:  $\exists$  Nische im Raum aller Probleme, für die Algorithmus besonders gut geeignet
- ▶ Anwendung: welchen Algorithmus nutzen bei gegebenem Problem?
- ▶ Wissenschaft: welche Klasse von Problemen optimal für bestimmten Algorithmus?

# Zusammenfassung

falls keinerlei Problemwissen vorhanden

⇒ Erwartung von EA gegenüber beliebigem Verfahren nicht höher

falls Problemwissen vorhanden

▶ z.B. Annahmen über gewisses Wohlverhalten der  
Gütelandschaft

⇒ generelle Anwendbarkeit von bestimmten Algorithmen wird  
nahegelegt

Wissen über Struktur des Problems sollte fließt in Auswahl  
oder in Entwurf des Optimierunsalgorithmus einfließen

# Literatur zur Lehrveranstaltung I



Holland, J. H. (1975).

*Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.*

University of Michigan Press.