

Eine Analyse des Effektes von Lernen auf Populationsfitness und Diversität in einer NK-Fitnesslandschaft

Lars Melchior

Theoretische Grundlagen

Theoretische Grundlagen

- Genetik
- Genetische Algorithmen
- NK - Modell

Genetik

Genom

- Gesamtheit der vererbaren Information eines Organismus
- Bei Lebewesen hauptsächlich die Abfolge der Basen Adenin, Guanin, Cytosin und Thymin in der DNA
- Zu Lebzeit erlangte Eigenschaften werden nicht auf das Genom übertragen
- Durch zufällige Veränderungen können Basen verändert, hinzugefügt oder entfernt werden (Mutation)

Genetik

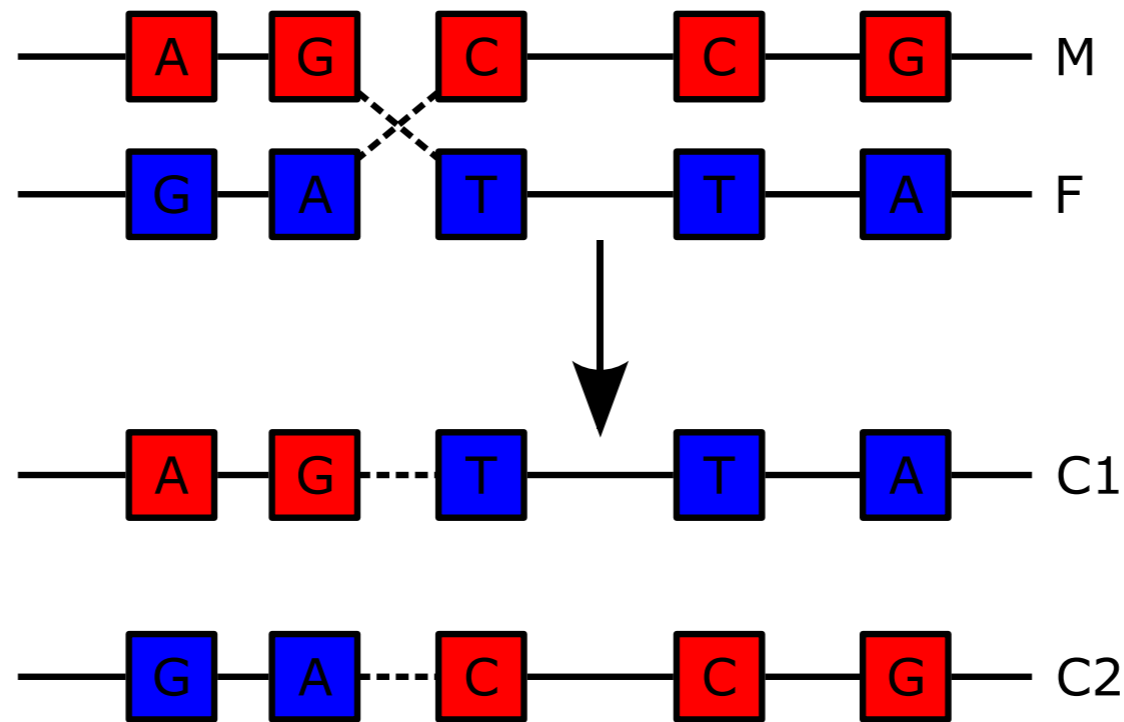
Vererbung

- Bei asexueller Reproduktion wird das Genom direkt an die Nachkommen vererbt
- Bei sexueller Reproduktion:
 - Interchromosomale Rekombination
 - Intrachromosomale Rekombination

Genetik

Intrachromosomale Rekombination

„Cross-Over“



Genetische Algorithmen

- Angelehnt an den Genetik und Evolution
- Wirken auf Populationen
- Lösungen werden nur anhand eines Skalars, der Fitness bewertet
- Arbeiten in diskreten Lösungsräumen

Genetische Algorithmen

Ablauf

- I. Erzeuge Anfangspopulation aus zufälligen Genomen aus dem Lösungsraum (Generation 0)
 - a. Bestimme für jedes Genom die Fitness
 - b. Ersetze die Population durch eine neue mit gleicher Anzahl an Genomen, wobei neue Genome zufällig mit bestimmten Wahrscheinlichkeit durch Cross-Over, Mutation oder Reproduktion aus selektierten Genomen der alten Population erzeugt werden

Genetische Algorithmen

Ablauf

1. Erzeuge Anfangspopulation aus zufälligen Genomen aus dem Lösungsraum (Generation 0)
2. Wende iterativ folgende Operationen auf die Population an, bis eine Abbruchbedingung erfüllt ist (Generation):
 - a. Bestimme für jedes Genom die Fitness
 - b. Ersetze die Population durch eine neue mit gleicher Anzahl an Genomen, wobei neue Genome zufällig mit bestimmten Wahrscheinlichkeit durch Cross-Over, Mutation oder Reproduktion aus selektierten Genomen der alten Population erzeugt werden
3. Das Genom mit der höchsten Fitness wird als Ergebnis der Simulation ausgegeben

Genetische Algorithmen

Selektion

- Entscheidend für die Evolutionsrate
- Einfachste Möglichkeit:
Auswahlwahrscheinlichkeit \sim Fitness
- Jedoch:
schnell schwindende Diversität
erreichen eines lokalen Fitness Maximums

Genetische Algorithmen

Fitnessfunktion

- Ordnet jedem Genom ein Fitnesswert zu
- Soll durch Genetischen Algorithmus Maximiert werden
- Ist auf den Raum der Genome eine Metrik definiert, so ergibt sich eine Fitnesslandschaft
- Im allgemeinen besitzt diese viele lokale „Hügel“ und „Täler“

NK Modell

$$F_{\text{NK}} : \mathbb{A}^N \rightarrow \mathbb{R}$$

- „Rauheit“ lässt sich anhand des Parameters K einstellen

NK Modell

Die Fitness F_{NK} einer Zeichenkette S ergibt sich aus der Summe der

Beiträge der einzelnen Zeichen $S_i \in \mathbb{A}$:

$$F_{\text{NK}}(S) = \sum_{i=1}^N f(S_i)$$

Der Einzelbeitrag eines Zeichens S_i hängt selbst von K weiteren Zeichen ab:

$$f(S_i) = f(S_i, S_1^i, \dots, S_K^i)$$

NK Modell

Beispiel I

$$N = 5, K = 1$$

$$\mathbb{A} = \{1, 0\}$$

Als Einzelbeitrag eines Zeichens definieren wir:

$$f(S_i) = \begin{cases} f(S_5, S_1), & \text{falls } i = 5 \\ f(S_i, S_{i+1}), & \text{sonst} \end{cases}$$

und

$$f(0, 0) = 0, f(0, 1) = 1, f(1, 0) = 2, f(1, 1) = 0$$

NK Modell

Beispiel 1

$$f(S_i) = \begin{cases} f(S_5, S_1), & \text{falls } i = 5 \\ f(S_i, S_{i+1}), & \text{sonst} \end{cases}$$

$$f(0, 0) = 0, f(0, 1) = 1, f(1, 0) = 2, f(1, 1) = 0$$

Für die Zeichenketten 11000 und 01010 ergibt sich so die Fitness:

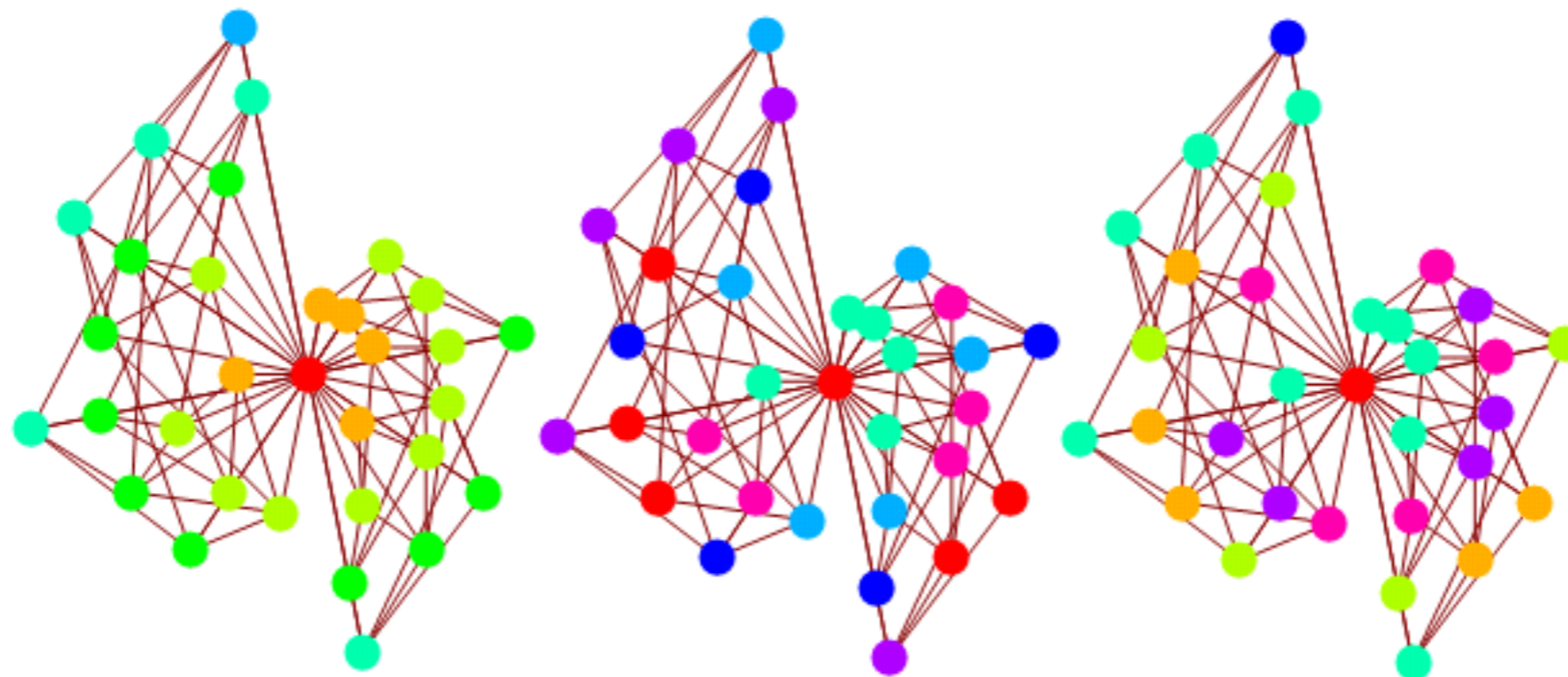
$$\begin{aligned} F_{\text{NK}}(11000) &= f(1, 1) + f(1, 0) + f(0, 0) + f(0, 0) + f(0, 1) \\ &= 0 + 2 + 0 + 0 + 1 = 3 \end{aligned}$$

$$\begin{aligned} F_{\text{NK}}(01010) &= f(0, 1) + f(1, 0) + f(0, 1) + f(1, 0) + f(0, 0) \\ &= 1 + 2 + 1 + 2 + 0 = 4 \end{aligned}$$

NK Modell

Beispiel II

$$N = 5$$



$$K = 0$$

$$K = 1$$

$$K = 2$$

[Quelle: Wikipedia]

An Analysis of the Effects of Lifetime Learning on Population Fitness and Diversity in an NK Fitness Landscape

Dara Curran, Colm O'Riordan and Humphrey Sorensen

How Learning can guide Evolution

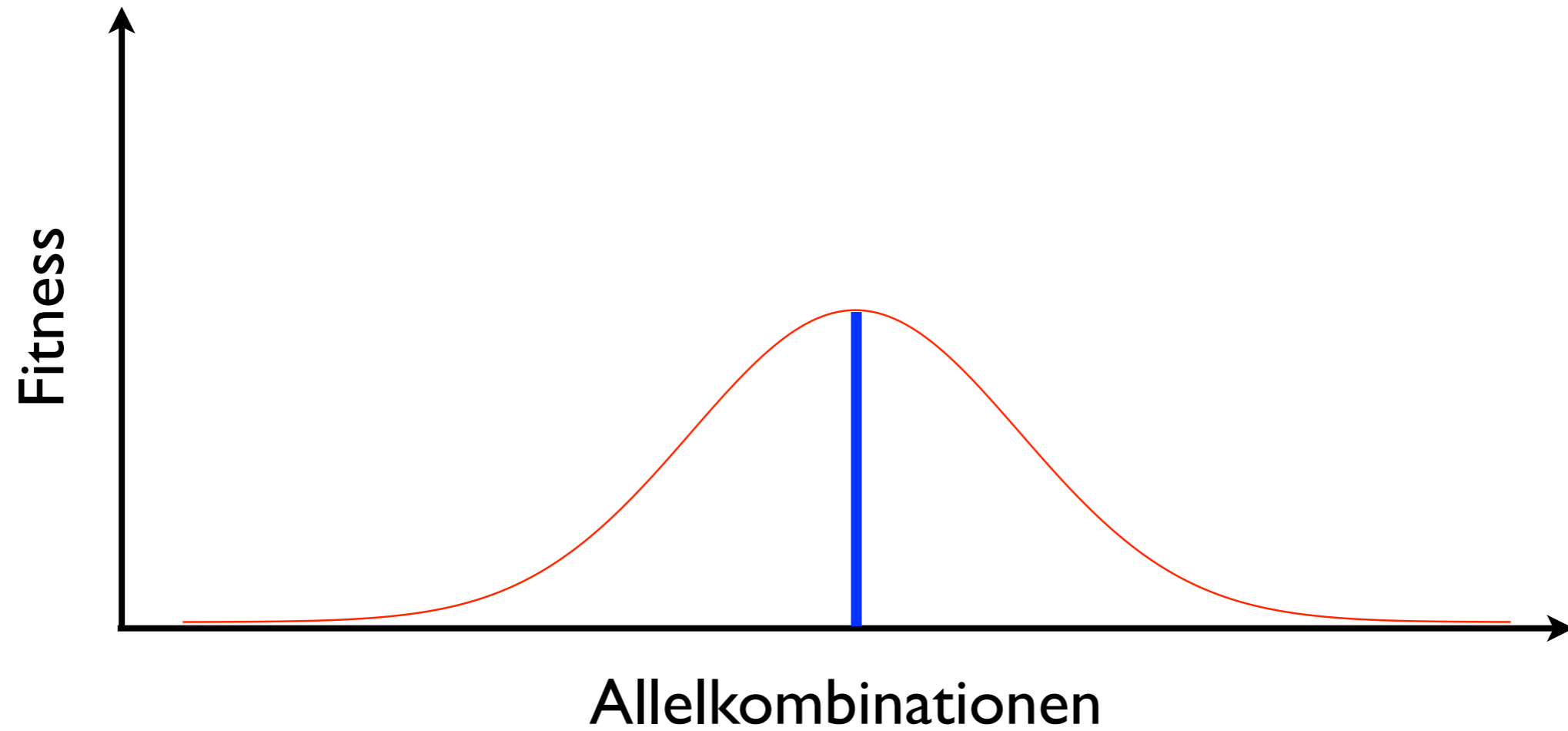
Geoffrey E. Hinton & Steve J. Nowlan

Baldwin Effekt

- Jack Mark Baldwin (1896)
- Zu Lebzeit erlernte Eigenschaften werden nicht auf das Genom übertragen
- Jedoch kann Lernen den Pfad der Evolution steuern

Baldwin Effekt

Extremes Beispiel



An Analysis of the Effects of Lifetime Learning on Population Fitness and Diversity in an NK Fitness Landscape

Dara Curran, Colm O'Riordan and Humphrey Sorensen

Modell

- Lernende sowie nicht-lernende Agenten
- Jeder Agent besitzt ein Genom
- Über Genetischen Algorithmen werden Populationen von Agenten evolviert
- Fitness wird über ein NK Modell bestimmt

Modell

Genom

Genom eines Agenten i : $G_i \in \mathbb{A}^{20}$

$$\mathbb{A}_{\text{Nicht Lernend}} = \{1,0\} \quad \mathbb{A}_{\text{Lernend}} = \{1,0,?\}$$

Anfangspopulation:

$$p(0) = 0,5$$

$$p(1) = 0,5$$

$$p(0) = 0,25$$

$$p(1) = 0,25$$

$$p(?) = 0,5$$

Modell

Fitnessfunktion

Normierte NK Fitnesslandschaft mit

$$N = 20, K = 15 \text{ und } \mathbb{A} = \{1,0\}$$

Nicht lernender Agent i : $F(G_i) = F_{\text{NK}}(G_i)$

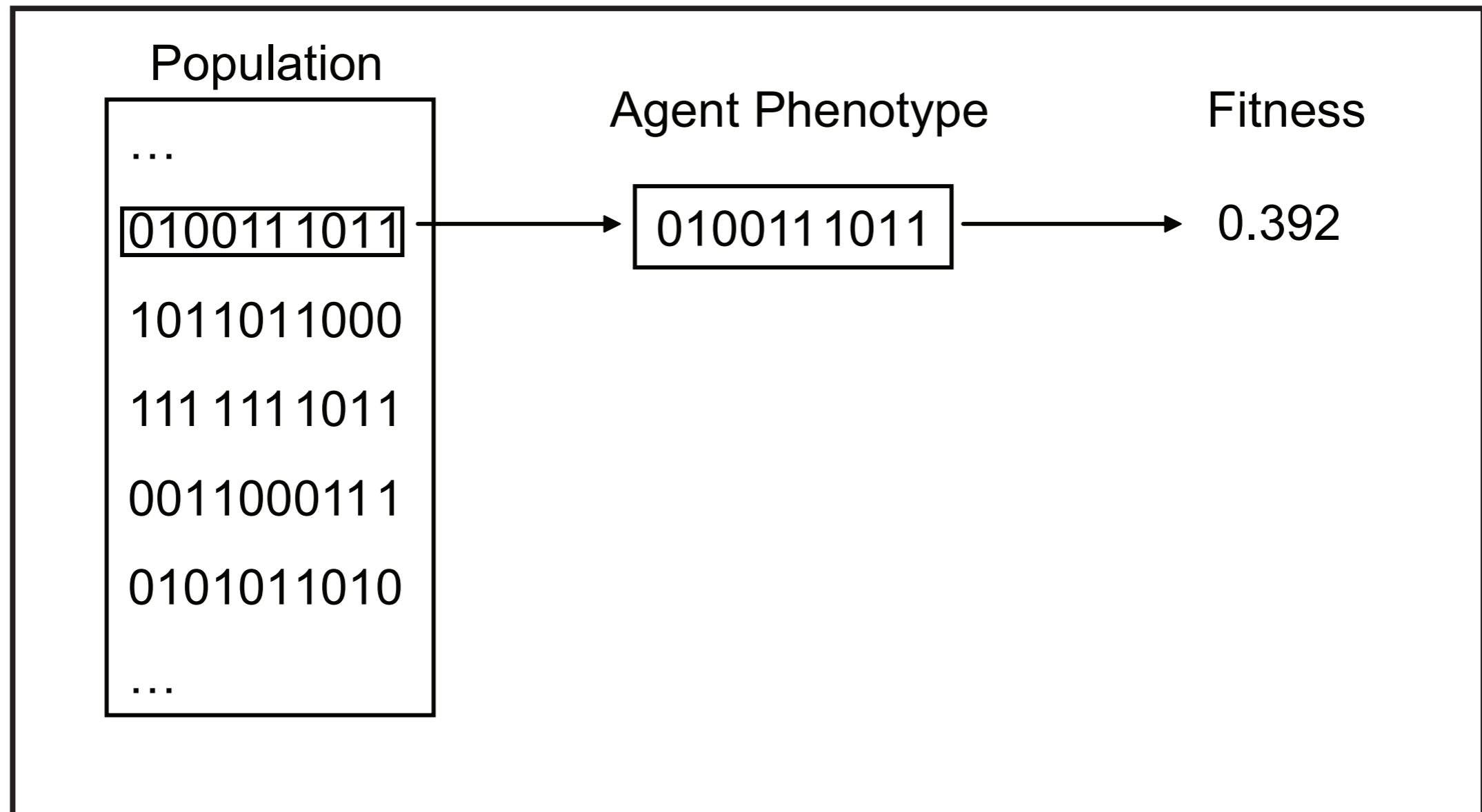
Lernender Agent i : $F(G_i) = \max(F_{\text{NK}}(G_i^1), \dots, F_{\text{NK}}(G_i^L))$

G_i^j : erzeugt aus G_i durch zufälliges Ersetzen der '?' mit '0' oder '1' (Gleich wahrscheinlich)

Phenotyp: Zeichenkette, die zur höchsten Fitness beigetragen hat

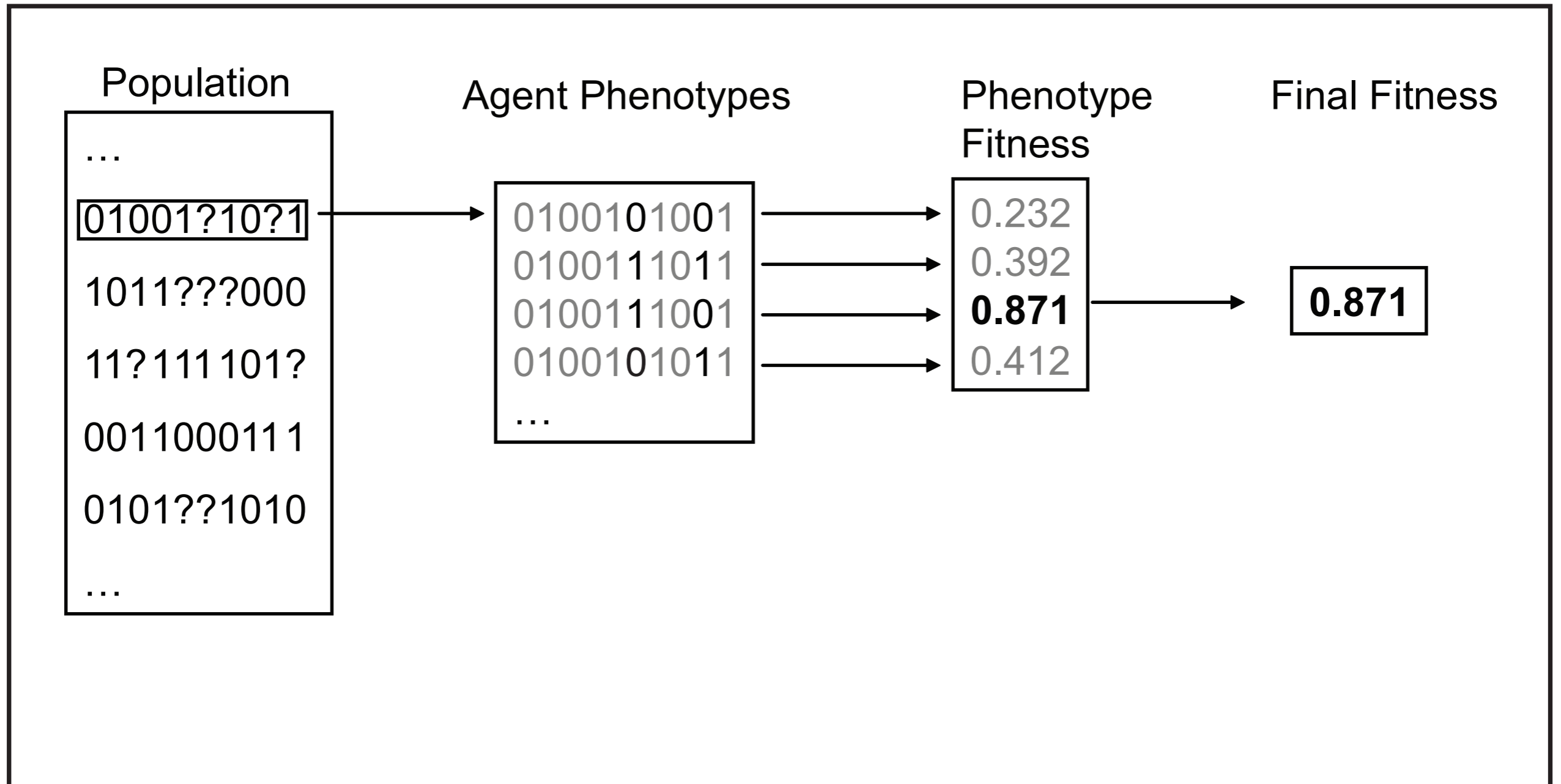
Modell

Fitnessfunktion



Modell

Fitnessfunktion



Modell

Diversität

$$\text{Diversität} := \frac{2}{n(m-1)} \sum_{x=1}^m \sum_{y=x+1}^m h(x, y)$$

n : Zahl der Phenotypen

m : Zahl der einzigartigen Phenotypen

$h(x, y)$: Hamming-Abstand zwischen zwei einzigartigen Phenotypen

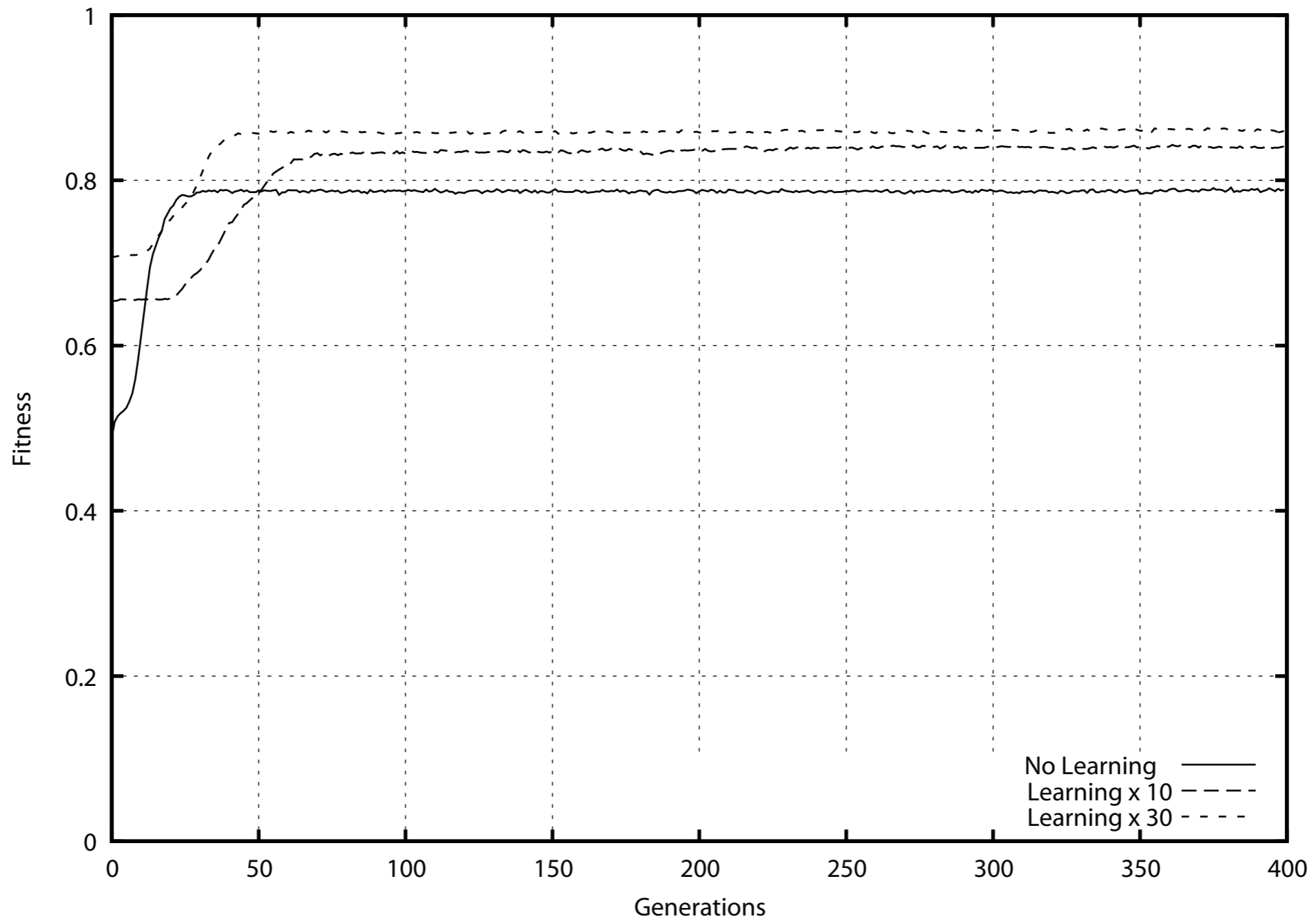
Simulationen

Wahl der Parameter

- 3 Populationen mit jeweils 1000 Agenten
 1. ohne Lernen
 2. $L = 10$
 3. $L = 30$
- Populationen werden über 400 Generationen evolviert
- 1-Point Crossover Wahrscheinlichkeit ist 0,6
- Mutationswahrscheinlichkeit ist 0,01
- $N = 20, K = 15$
- Resultate sind Mittelwert aus 20 Unabhängigen Durchlaufen

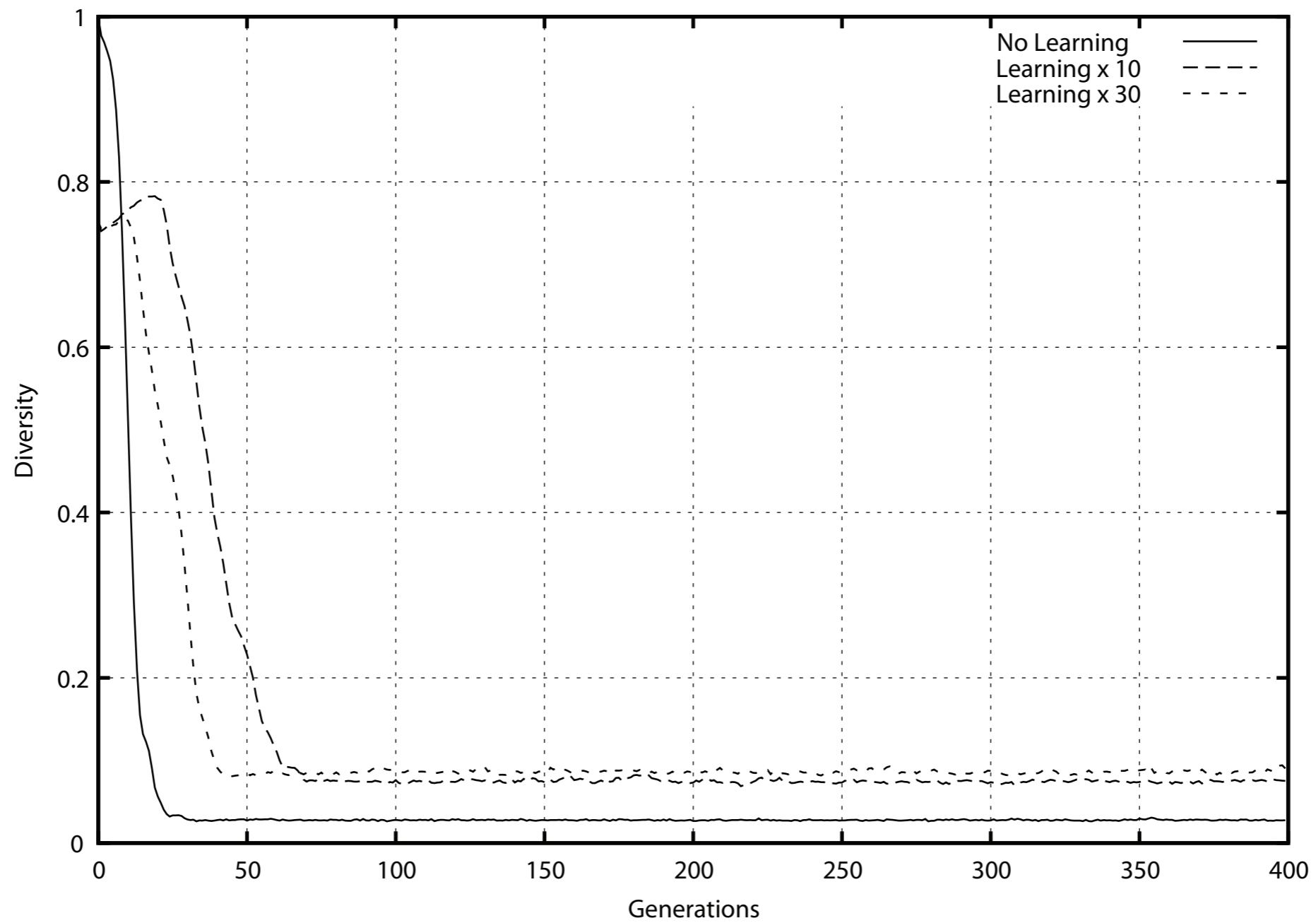
Ergebnisse

Fitness



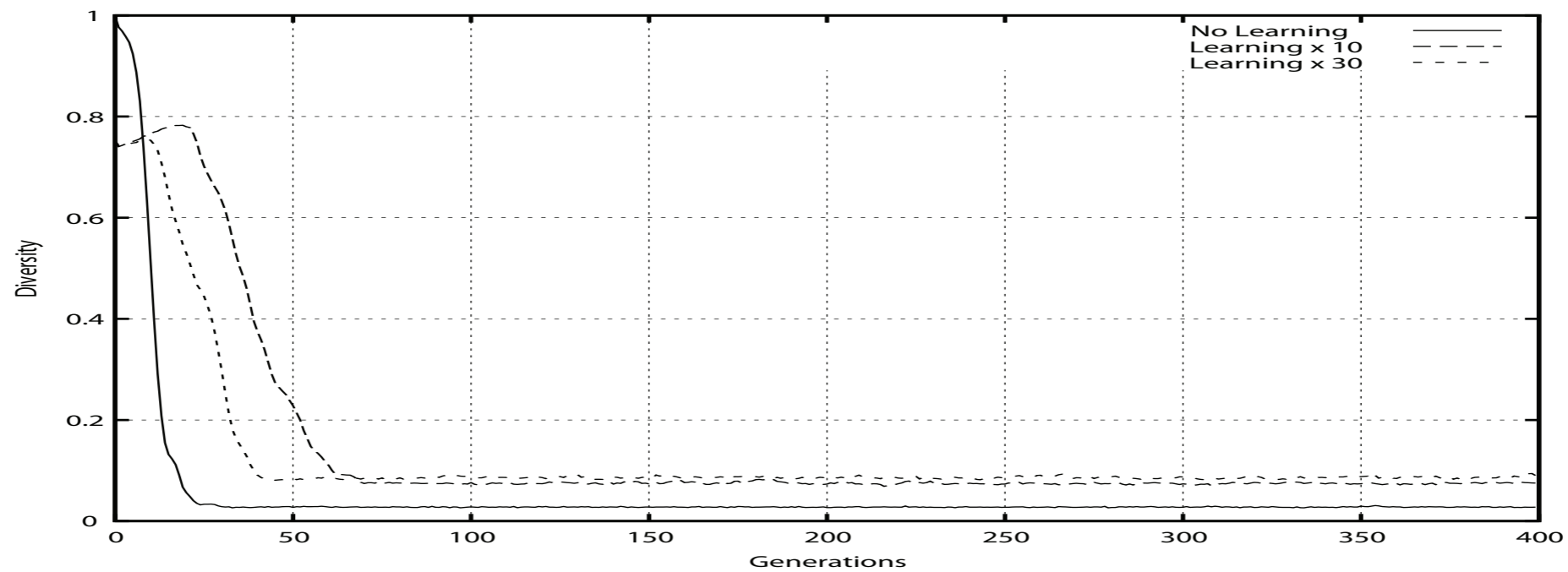
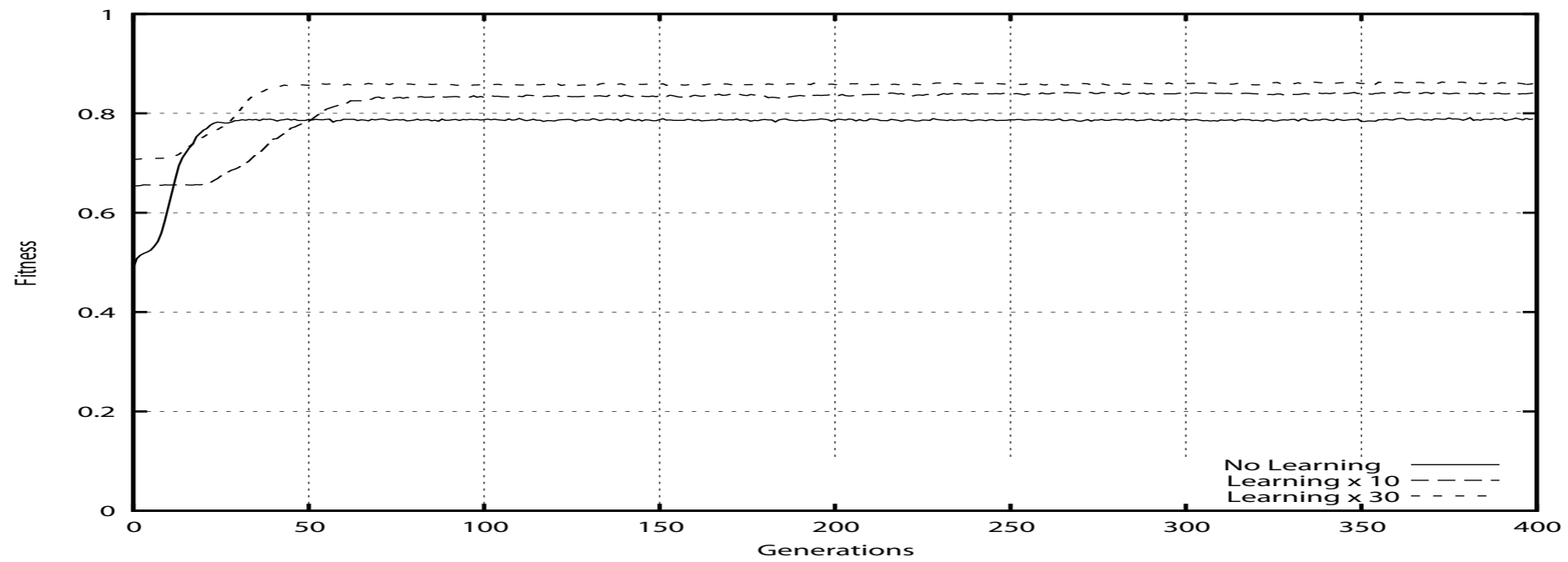
Ergebnisse

Diversität



Ergebnisse

Diversität



Ergebnisse

- Diversität kann über Phenotypen berechnet werden
- Lernen → Höhere Fitness, Höhere Diversität in NK Fitnesslandschaften

