

Berufsakademie Sachsen  
Staatliche Studienakademie Leipzig  
Studiengang Informatik  
Evolutionäre Algorithmen

# **Binäre und reelle Kodierung im Vergleich bei der näherungsweisen Berechnung des globalen Minimums der Griewank-Funktion**

Eingereicht von: Georg Andrassy  
Martin Braun

Matrikelnr: 5000593  
Matrikelnr: 5000562

Leipzig, 28. 11. 2016

# 1. Einführung

Zur näherungsweisen Berechnung des globalen Minimums der Griewank-Funktion:

$$f(x) = 1 + \sum_{i=1}^n \frac{x_i^2}{400n} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right)$$

wird ein evolutionärer Algorithmus eingesetzt. Dieser Algorithmus wird vor allem im Hinblick auf den Einfluss unterschiedlicher Kodierungsarten sowie stabiler und wachsender Populationsgrößen auf den Fitnesswert der Individuen untersucht. Zusätzlich wird die Differenz zwischen dem besten und dem schlechtesten Individuum in der Population als Maß der Streuung betrachtet. Die Minimierung der Griewank-Funktion wird für drei verschiedene n (n=5, 20, 50) im Wertebereich von -512 bis 511 getestet:

## 2. Umsetzung

Für die Umsetzung wird folgender Algorithmus verwendet:

1. Erzeugung der Startpopulation
  - fixe Ausgangspopulation mit X Anfangsindividuen
2. Elternselektion
  - Anzahl Elternpaar = e mit zufälliger Selektion
  - e wachsend bzw. stabil
3. Rekombination
  - Reelle Kodierung: intermediäre Rekombination
  - Binärkodierung: Ein-Punkt- und Zwei-Punkt-Rekombination (Zwei-Punkt: Mittelteil austauschen)
4. Mutation
  - Mutationswahrscheinlichkeit M [0,1]
  - von jedem Individuum soll zufällig ein Gen ausgewählt werden
  - für Mutation eines Gens Zufallszahl z [0,1] ermitteln
  - Reelle Kodierung: wenn  $z < M$ , dann mutiere: addiere zum gewählten Allel einen festen Wert W
  - Binärkodierung: zufällig von dem gewählten Binär-Allel ein Bit auswählen, wenn  $z < M$ , dann switche Bit von 0 zu 1 bzw. 1 zu 0

## 5. Umweltselektion

- aus Eltern + Kindern soll neue Population gewählt werden
- zuerst die besten A Individuen nehmen (deterministische Selektion) Gesamtpopulation soll wachsen (A wächst linear von Generation zu Generation um +1 an)
- dann aus Rest zufällig B Individuen auswählen

## 6. gehe zu Punkt 2 bis Abbruchkriterium K erreicht

Verwendete Parameter

| <i>Parameter</i>                       | <i>Wert</i>        |
|--|--------------------|
| Größe der Startpopulation X Individuen | 10                 |
| Anzahl der Elternpaare e               | 10 (+1 pro Zyklus) |
| Rekombinationswahrscheinlichkeit       | 0.9                |
| Mutationswahrscheinlichkeit M          | 0.1                |
| reeller Mutationswert W                | +5.0               |
| Auswahl bester Individuen a            | 10 (+1 pro Zyklus) |
| Auswahl zufällige Umweltselektion B    | 3                  |
| Abbruchkriterium K Zyklen              | 1500               |

## 3. Ergebnisse und Interpretation

Die Ergebnisse wurden für vier verschiedene Aspekte ermittelt (Mittelwerte von 100 Versuchen, Abbruch nach 500 Zyklen):

1. Vergleich zwischen Reeller und Binärer Kodierung (1-Punkt- und 2-Punkt-Rekombination)
2. Vergleich zwischen stabiler und wachsender Population
3. Vergleich zwischen verschiedener Anzahl an Genen (für  $n = 5, 20, 50$ )
4. Vergleich zwischen Streuung innerhalb der Kodierungsarten

Beim Vergleich der Kodierungsarten in Abbildung 1 fällt auf, dass die beiden binären Kodierungsarten bessere Fitnesswerte bringen. Innerhalb der Binärkodierung bringt zusätzlich die Eins-Punkt-Rekombinationsmethode noch einen schnelleren Abfall zum globalen Minimum.

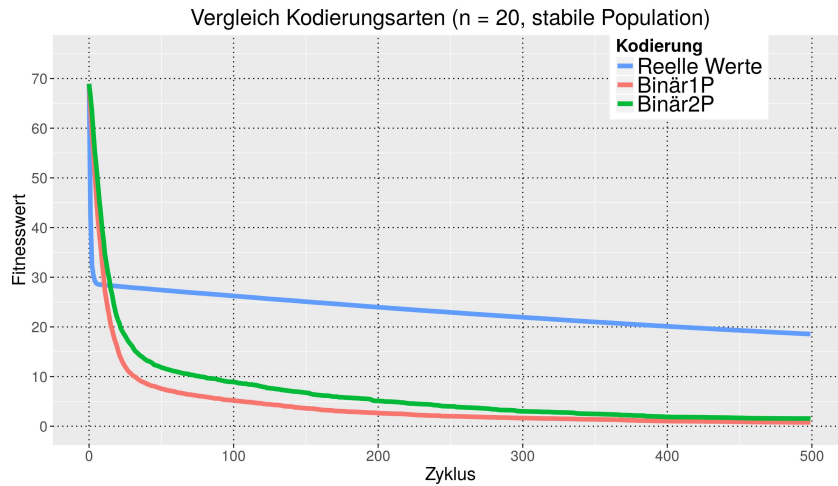


Abb.1: Kodierungen im Vergleich bei n=20 (stabile Population)

Nach Abbildung 2 für  $n = 50$  liefert der Algorithmus mit wachsender Population ein besseres Ergebnis als mit stabiler Population.

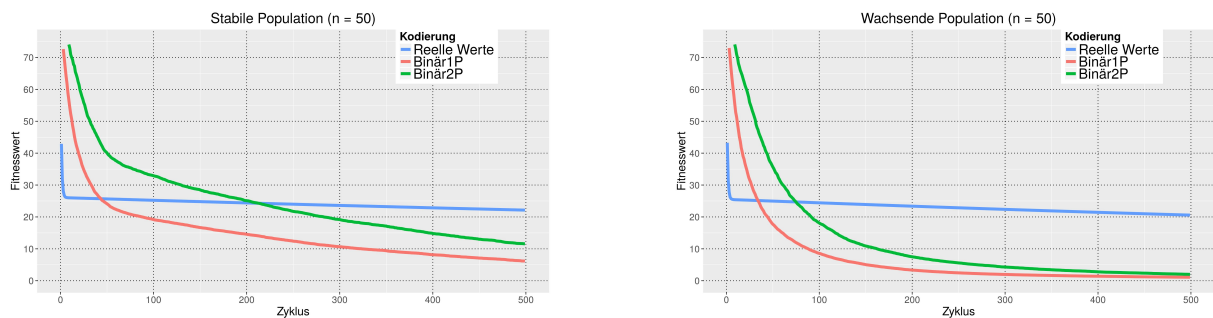


Abb.2: Vergleich stabile und wachsende Population (n=50)

Mit steigender Anzahl der Gene wird das globale Minimum immer verzögerter annähernd erreicht, wie in Abbildung 3 zu sehen ist.

Bei Betrachtung der Streuung der Individuen in Abbildung 4 ist auffällig, dass sich bei der reellen Kodierung das schlechteste und das beste Individuum sehr schnell annähern. Bei der binären Kodierung sind die Unterschiede etwas ausgeprägter. Es wird vermutet dass die Streuung maßgeblich für das Ergebnis verantwortlich ist, so dass die besseren Ergebnisse bei den binären Kodierungen aufgrund der größeren Streuung zwischen den Individuen zustande kommen.

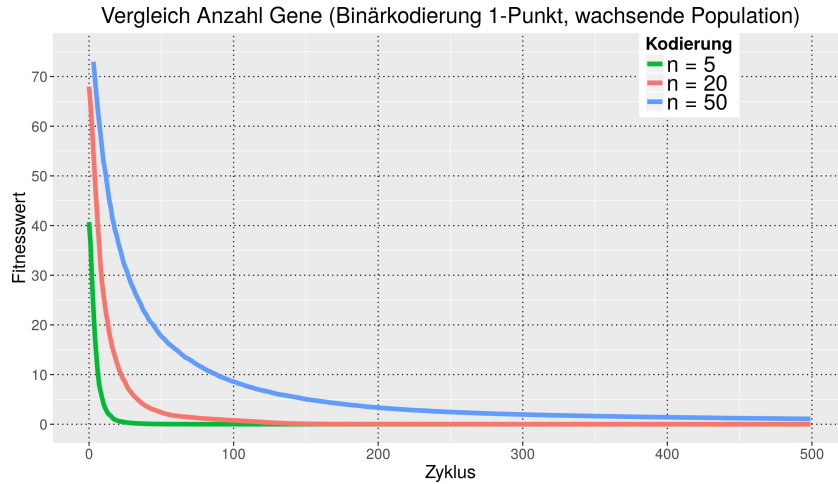


Abb.3: Vergleich verschiedene Anzahl an Genen (Binärkodierung mit 1-Punkt-Rekombination, wachsende Population)

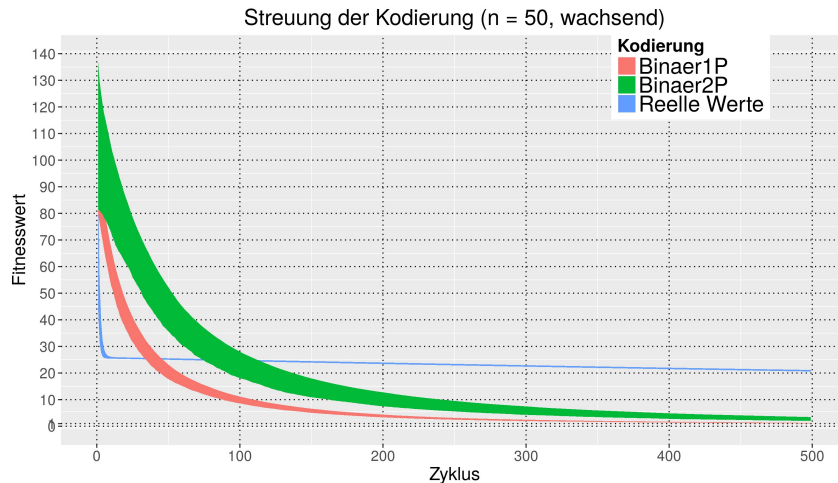


Abb.4: Vergleich der Streuung innerhalb der Kodierungsarten ( $n = 50$ , wachsende Population)

Im untersuchten Algorithmus liefert die reelle Kodierung generell schlechtere Ergebnisse. Ein Ansatzpunkt für bessere Ergebnisse wäre, die Streuung der reell kodierten Individuen zu erhöhen. Durch die Anpassung der weiteren Parameter des evolutionären Algorithmus sind bessere Ergebnisse zu erwarten.

Ein Erklärungsansatz für den Unterschied zwischen reeller und binärer Kodierung stellt die spezielle Anforderung der Griewank-Funktion dar. Deren globales Minimum liegt exakt bei Null. Bei der reellen Kodierung kann entweder ein zufällig anfangs existierendes Individuum mit Allelen nah bei Null für das optimale Ergebnis sorgen. Außerdem ist dies über die Mutation möglich. Währenddessen bietet sich bei der Binärkodierung zusätzlich die Chance auf die Existenz eines solchen Individuums durch Rekombination von Elternpaaren, so dass sich ein Binärstrang ausschließlich aus Nullen ergeben kann.