

Zeitreihenmodellierung mit GEP

Evolutionäre Algorithmen
Matthias Jauernig (B. Sc.), 26.06.07

Überblick

1. Kurzvorstellung von GEP
2. Problem der Zeitreihenanalyse
3. Eingesetzter Algorithmus
4. Praktischer Teil, erste Ergebnisse



(1) Kurzvorstellung von GEP



GEP kurz erklärt

- ▶ Spezieller genetischer Algorithmus
- ▶ 2001 von Cândida Ferreira entwickelt
- ▶ Evolution von Programmen/Algorithmen

- ▶ Besonderheiten:
 - ▶ Spezieller Aufbau von Genotypen und Phänotypen
 - ▶ Genotypen fester Länge erlauben Phänotypen variabler Länge
 - ▶ Genotypen erfordern gesonderte Übersetzung in Phänotypen

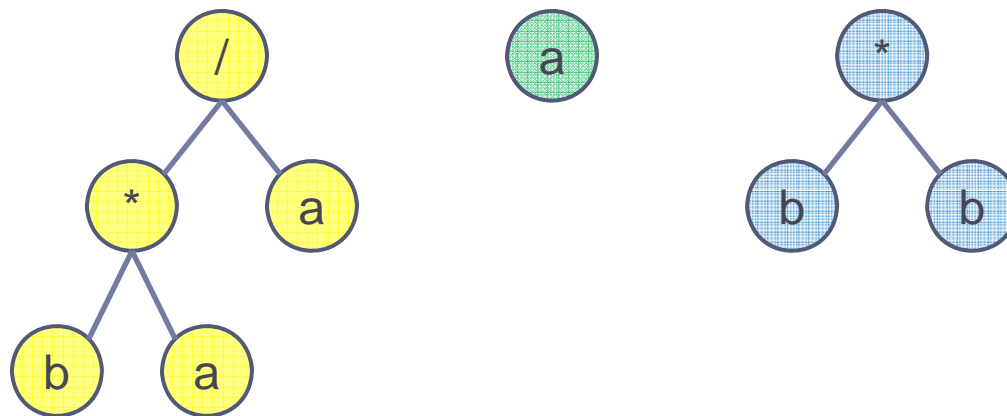


(1)

GEP kurz erklärt

Chromosom (Genotyp)														
Gen 1					Gen 2					Gen 3				
Kopf		Rest			Kopf		Rest			Kopf		Rest		
/	*	a	b	a	a	*	b	b	a	*	b	b	a	a

Syntaxbäume
(Phänotyp):

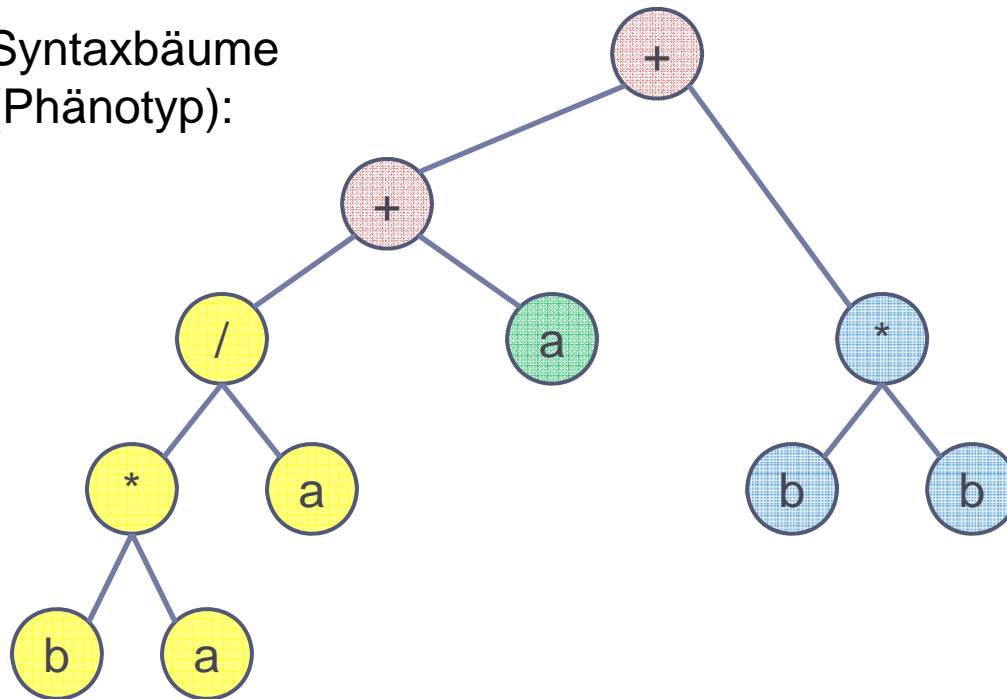


(1)

GEP kurz erklärt

Chromosom (Genotyp)														
Gen 1					Gen 2					Gen 3				
Kopf	Rest				Kopf	Rest				Kopf	Rest			
/	*	a	b	a	a	*	b	b	a	*	b	b	a	a

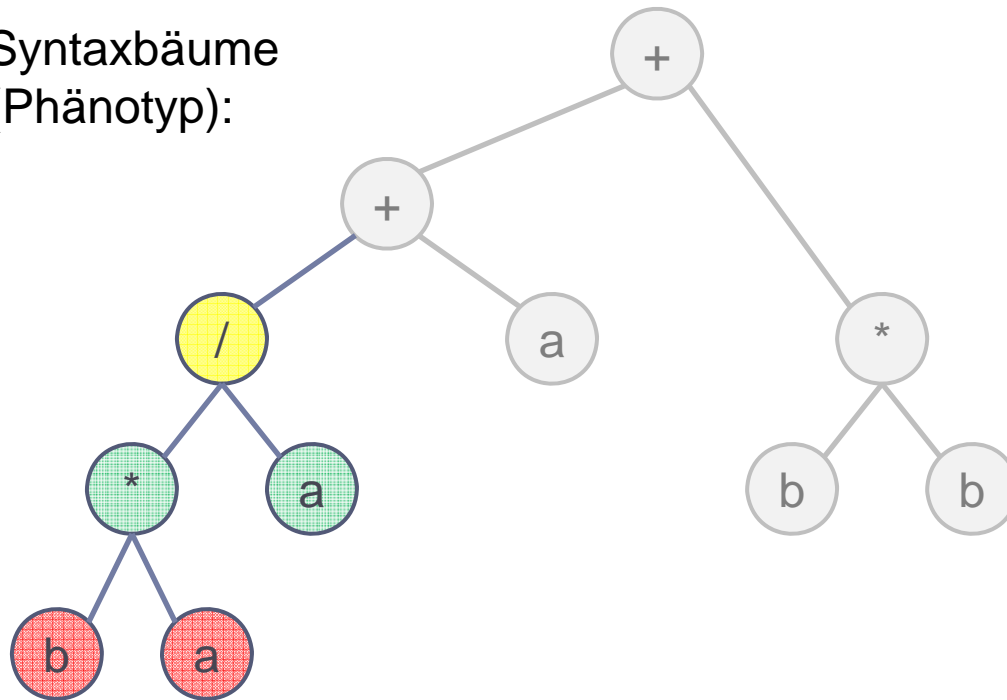
Syntaxbäume
(Phänotyp):



GEP kurz erklärt

Chromosom (Genotyp)														
Gen 1					Gen 2					Gen 3				
Kopf		Rest			Kopf		Rest			Kopf		Rest		
/	*	a	b	a	a	*	b	b	a	*	b	b	a	a

Syntaxbäume
(Phänotyp):



(2) Problem der Zeitreihenanalyse



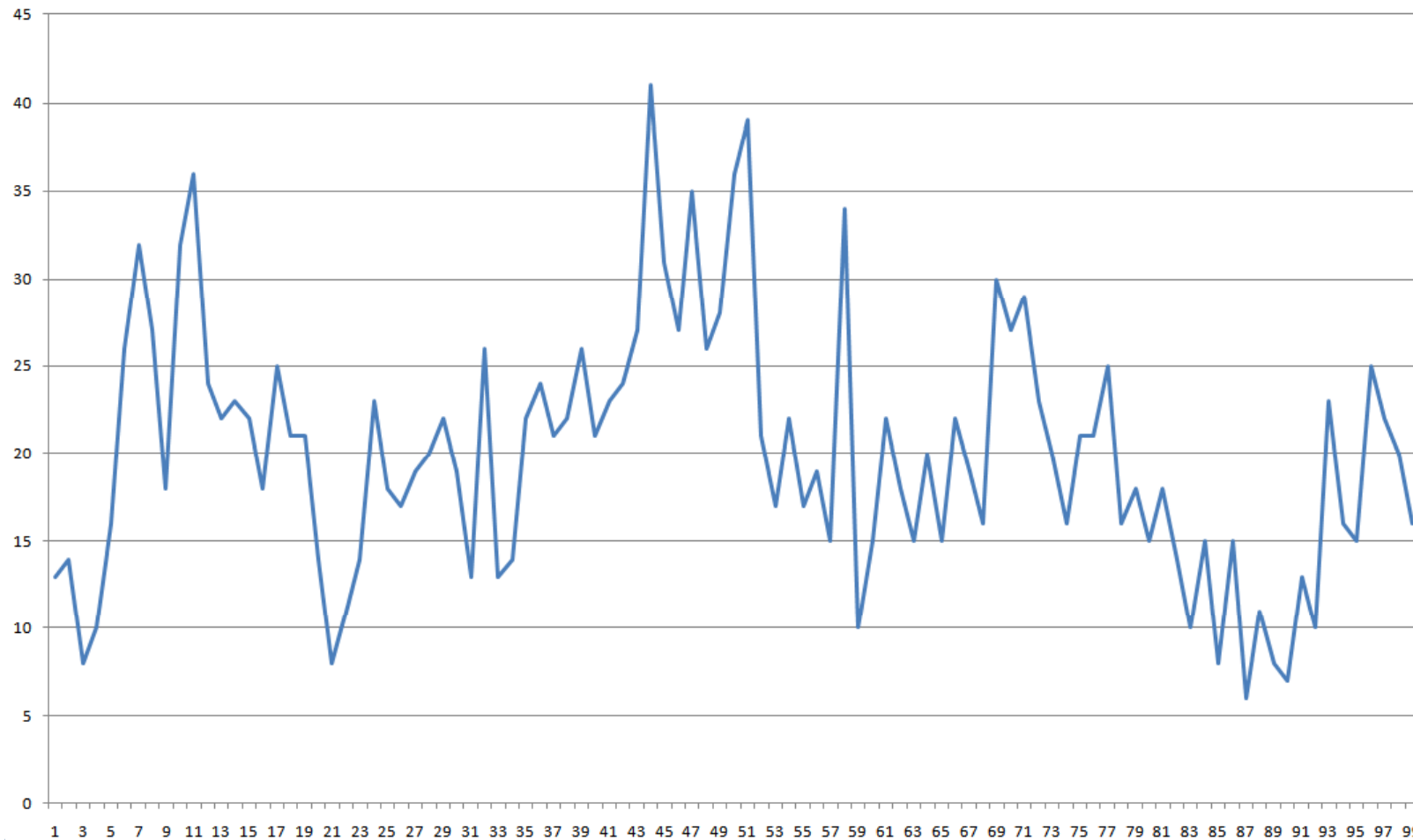
Zeitreihenanalyse

- ▶ Folge von x Zahlen, Zeitpunkte der Messung
gewöhnlich äquidistant:
 - ▶ $x(t) = f [x(t-\tau), x(t-2*\tau), \dots, x(t-(d-1)*\tau)]$
 - ▶ τ ... Zeit zwischen 2 Messungen
 - ▶ d ... eingebettete Dimension
 - ▶ f ... (gesuchte) Funktion zur Beschreibung der Zeitreihe
- ▶ Typische Zeitreihen entstehen sowohl durch
regelmäßige als auch zufällige Ursachen
- ▶ Unterteilung in Trainingsdaten und Testdaten
- ▶ Algorithmus finden, der zu einer Anzahl n von
Parametern (Zahlen $i-n \dots i-1$ der Zeitreihe) den
aktuellen Wert i voraussagt



Zeitreihenanalyse

► Typische Zeitreihe:



(3) Eingesetzter Algorithmus



Eingesetzter Algorithmus

- ▶ **Fitnessfunktion**
 - ▶ Summe über quadratischen oder absoluten Fehler aller Voraussagen über die Trainingsdaten



Eingesetzter Algorithmus

▶ Replikation

- ▶ Eltern werden durch Selektion ausgewählt und „wie sie sind“ in die nächste Generation übernommen → Populationsgröße bleibt stets erhalten

▶ Selektion

- ▶ Duplikate möglich und sogar höchst wahrscheinlich
- ▶ Roulette-Rad (fitnessproportional) mit rangbasierter Verteilungsmethode
- ▶ Turnier-Selektion: herausgreifen eines Teils der Population und dort den Gewinner ermitteln



Eingesetzter Algorithmus

- ▶ Mutation

- ▶ 1-Punkt-Mutation

- ▶ Im Gen-Kopf: gegen beliebige Symbole
 - ▶ Im Gen-Rest: nur gegen Terminalsymbole zur Wahrung der Gen-Struktur



Eingesetzter Algorithmus

- ▶ **Rekombination**
 - ▶ 1-Punkt-Rekombination
 - ▶ 2-Punkt-Rekombination
 - ▶ Gen-Rekombination



Eingesetzter Algorithmus

▶ Transposition

(Verschieben von Genmaterial innerhalb eines Chromosoms)

▶ IS-Element-Transposition

- ▶ „Insertion Sequence“: irgendeine Symbol-Sequenz

▶ RIS-Element-Transposition

- ▶ „Root Insertion Sequence“: starten mit Funktionssymbol, werden an Gen-Anfang gesetzt

▶ Gen-Transposition

- ▶ Gen wird innerhalb eines Chromosoms verschoben
- ▶ Verdrängtes Gen-Material wird „weiter geschoben“



(4) Praktischer Teil, erste Ergebnisse



Ende

