Poker Challenge

TECHNISCHE UNIVERSITÄT DARMSTADT

Evolutionäre Algorithmen

Michael Herrmann und Arno Mittelbach

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

Evolutionäre Algorithmen

Warum seid Ihr hier?

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

Evolutionäre Algorithmen

1. Ansatz

Habe ich zwei Asse auf der Hand, dann gehe ich mit!

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

Evolutionäre Algorithmen

2. Ansatz

Habe ich ein Asspärchen oder zwei Bilder
auf der Hand, dann gehe ich mit!

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

Evolutionäre Algorithmen

42. Ansatz

Bin ich in später Position, habe Pott-Odds von weniger als 30%, mindestens 8 Outs und mindestens ein 7er Pärchen, dann erhöhe um die Pott-Size.

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

Evolutionäre Algorithmen

1824. Ansatz

Bin ich vor drei Spielen auf dem Turn rausgegangen, habe Karo und Herz auf der Hand, wovon eines ein Bild ist und mein zweiter Nachbar von links eben einen neuen Whiskey bestellt hat, dann erhöhe ich um 25 Chips, falls ich noch über mehr als 752 Chips verfüge, gehe jedoch aus dem Spiel, falls die mir gegenüber sitzende Person einen roten Pullover trägt.

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

Evolutionäre Algorithmen

Und nun ... ?

Übersicht

TECHNISCHE UNIVERSITÄT DARMSTADT

- Einfache lernbasierte Ansätze
- Evolutionäre Algorithmen
- Pareto-Koevolution
- Fazit
- Fragen und Diskussion

Lernbasierte Algorithmen - Übersicht

TECHNISCHE UNIVERSITÄT DARMSTADT

- Einführung
- Definition
- Artikel „An investigation of an Adaptive Poker Player“
  - Poker Variante
  - Datenstruktur
  - Algorithmus
  - Gegenspieler
  - Ergebnis

Lernbasierte Algorithmen – Einführung

TECHNISCHE UNIVERSITÄT DARMSTADT

- Wie in der Motivation beschrieben brauchen wir einen Mechanismus, der uns hilft Poker zu ***lernen***.
- Das führt uns zu den ***Lern***basierten Algorithmen.
- Wenn ein Programm lernen könnte, welches gute Karten beim Poker sind, so hätte es schon mal einen wesentlichen Punkt gelernt.
- Dieser Ansatz verfolgt der Artikel, der im Folgenden vorgestellt wird.

Lernbasierte Algorithmen - Definition

TECHNISCHE UNIVERSITÄT DARMSTADT

- Für uns gilt im folgenden diese Definition:

Ein lernbasierter Algorithmus ist ein Algorithmus, der während seiner Laufzeit Teile seiner Daten, mit denen er Spielentscheidungen trifft, dahingehend ändert, dass er bessere Spielentscheidungen treffen kann.

Lernbasierte Algorithmen - Pokervariante

TECHNISCHE UNIVERSITÄT DARMSTADT

- Im Artikel „An investigation of an Adaptive Poker Player“ wird folgende Form von Draw Poker verwendet.

1. Pokereinsatz von jedem Spieler.
2. Jeder Spieler bekommt fünf Karten.
3. Eine Runde Wetten.
4. Zwei Karten dürfen getauscht werden.
5. Eine weitere Runde Wetten.

Lernbasierte Algorithmen - Datenstruktur

TECHNISCHE UNIVERSITÄT DARMSTADT

- Jede mögliche Hand bekommt einen Lernwert (learning value) zugeteilt.
- Das ist eine Kommazahl zwischen 0 und 10.
- Eine hohe Zahl bedeutet, dass die Hand als gut angesehen wird.
- Jede Hand startet mit einem Wert von 10,0.

Lernbasierte Alogorithmen – Algorithmus 1/2

TECHNISCHE UNIVERSITÄT DARMSTADT

- Neben dem eben vorgestellten Lernwert (lv) gibt es noch weitere Kriterien:
  - Potgröße (pv)
  - Anzahl der Spieler die im Spiel geblieben sind. (np)
  - w ist gewichteter Faktor abhängig vom Lernwert (w aus {1,3,4,5})

- Pseudocode des Algorithmus

```
If lv < 6 then FOLD
        elseif lv >= 6 AND lv < 8 then CALL
        elseif lv >= 8 then
                ac = (lv / LOG(pv)) / (np / lv)
                if ac < 10 then CALL
                else RAISE by SQRT(ac) * w
```

Lernbasierte Alogorithmen – Algorithmus 2/2

TECHNISCHE UNIVERSITÄT DARMSTADT

- Daraus ergiben sich z.B. folgende Spielentscheidungen:
  - lv = 8, np = 5, pv = 50, ac = 7,53
    - Spieler wird mitgehen (Call).
  - lv = 8, np = 3, pv = 50, ac = 12,57
    - Spieler wird erhöhen (Raise) um 11 Einheiten. (Weniger Spieler)
  - lv = 9, np = 3, pv = 50, ac = 15,89
    - Spieler wird erhöhen (Raise) 16 Einheiten. (Bessere Erfahrungen mit der Hand)
- Wird eine Hand gespielt und der Spieler gewinnt, wird der Lernwert um 0,1 erhöht. Verliert der Spieler wird der Lernwert um 0,1 herabgesetzt.

Lernbasierte Algorithmen – Gegenspieler

TECHNISCHE UNIVERSITÄT DARMSTADT

- Die Gegenspieler werden statisch programmiert.
- Sie beachten lediglich ihre eigene Hand und treffen darauf hin statisch immer wieder die gleichen Entscheidungen.
- Einteilung dabei in *loose aggressive*, *loose passive*, *tight aggressive*, *tight passiv*.
- Haben ein unterschiedliches Wettrundenverhalten.

- Beispiel *loose aggressive* Gegenspieler (Erstrundenverhalten):
  - Fold → Bei allem bis einschließlich zu einem Achter-Pärchen.
  - Call → Neuner- bis Ass-Pärchen
  - Raise (verschiedene Höhen) → bei allem anderen
  - In der zweiten Wettrunde gleich, nur höherer Raise.

Lernbasierte Algorithmen – 1. Experiment 1/3

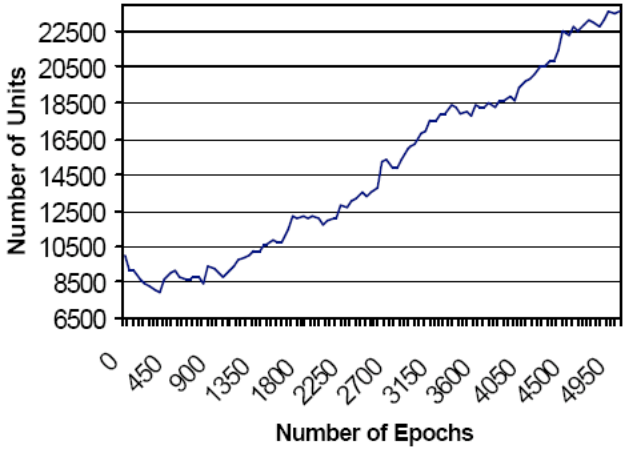
TECHNISCHE UNIVERSITÄT DARMSTADT

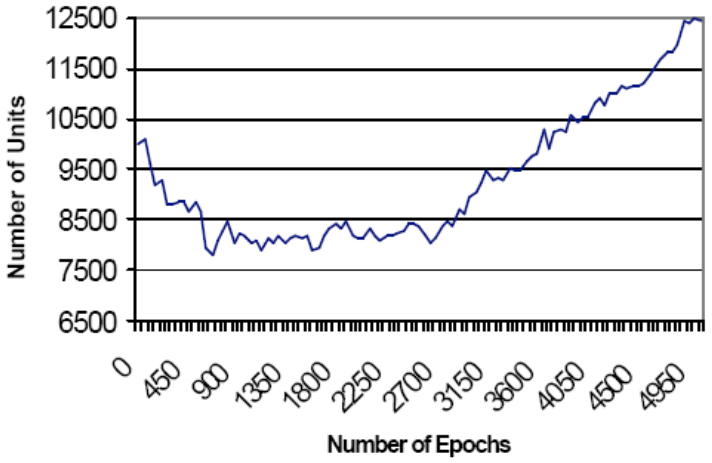
- Experminentaufbau
  - 1. Tisch: Ein sich entwickelnder Spieler, vier *loose aggressive* Spieler.
  - 2. Tisch: Ein sich entwickelnder Spieler, vier *loose passive* Spieler.
  - 3. Tisch: Ein sich entwickelnder Spieler, vier *tight aggressive* Spieler.
  - 4. Tisch: Ein sich entwickelnder Spieler, vier *tight passive* Spieler
- Jeder Spieler bekommt 10.000 Geldeinheiten.
- Grober Ablauf:
  - Zuerst macht ein sich entwickelnder Spieler großen Verlust, weil er jede Hand für eine Tophand hält.
  - Nach einer gewissen Lernphase holt er den Verlust aber wieder rein.
  - Von da an erhöht er stetig seinen Gewinn.

Lernbasierte Algorithmen – 1. Experiment 2/3

TECHNISCHE UNIVERSITÄT DARMSTADT

- Gewinnübersicht des sich entwickelnden Spielers. Im Vergleich gegen *loose aggressive (oben)* und *tight aggressive* Spieler (unten).





Lernbasierte Algorithmen – 1. Experiment 3/3

TECHNISCHE UNIVERSITÄT DARMSTADT

- Weitere Auffälligkeit: Teilnahme an den Spielen….

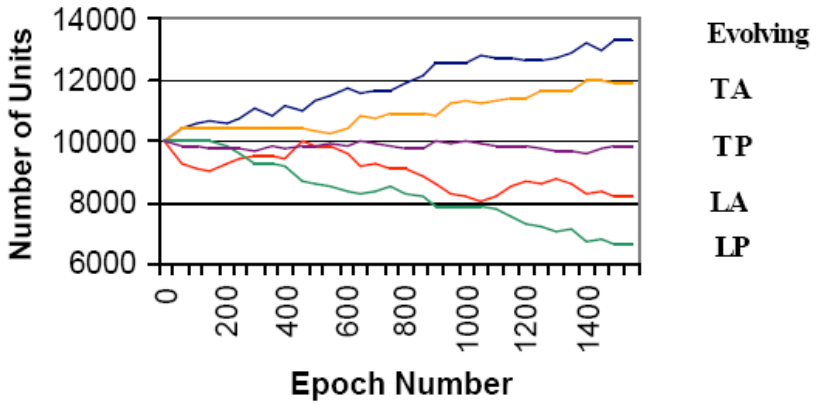
	25	**50**	**75**	**100**	**200**	**300**	**400**	**500**	**1000**	**1500**	**2000**	**2500**	**3000**
Hands Played	25	50	75	100	195	257	324	393	647	925	1187	1421	1664
Hands Won	5	12	19	28	60	81	102	123	216	312	409	484	560
% Played	100	100	100	100	97.5	85.6	81.0	78.6	64.7	61.6	59.3	56.8	55.5
% Won	20.0	24.0	25.3	28.0	30.7	31.5	31.5	31.3	33.3	33.7	34.5	34.0	33.6

- Übersicht siehe Artikel

Lernbasierte Algorithmen – 2. Experiment

TECHNISCHE UNIVERSITÄT DARMSTADT

- Nur ein Spieltisch
  - Ein sich entwickelnder Spieler (trainiert) und je ein statische Spieler von jedem möglichen Typ. (Also insgesamt fünf Spieler)
- Ergebnis:
  - Sich entwickelnder Spieler zuerst schlecht, dann besser und dann am besten.
  - Zweitbestes Ergebnis: *tight aggressiv* Spieler
  - Drittbestes Ergebnis: *tight passiv* Spieler
  - Viertbestes Ergebnis: *loose aggressive* Spieler
  - Fünftbestes Ergebnis: *loose passive* Spieler



Evolutionäre Algorithmen

Evolutionäre Algorithmen - Übersicht

TECHNISCHE UNIVERSITÄT DARMSTADT

- Einführung

- Artikel 1 „An Apaptive Learning Model for simplified Poker using Evolutionary Algorithms“.
  - Poker Variante
  - Datenstruktur
  - Algorithmus
  - Experimente

- Artikel 2 „Evolving Vomputer Opponents to play a game of simplified Poker“.
  - Datenstruktur
  - Algorithmus
  - Ergebnisse der Experimente mit obigem Artikel vergleichbar.

Evolutionäre Algorithmen - Einführung

TECHNISCHE UNIVERSITÄT DARMSTADT

- Ein Ansatz um Optimierungsprobleme zu lösen.
- Als Vorbild dient die biologische Evolution.

- Dabei durchlaufen Individuen einer Population folgende Aktionen:
  - Selektion

    Dabei wird eine Richtung vorgegeben in die sich die Population entwickeln soll.
  - Rekombination

    Beispiel: Um noch bessere Ergebnisse zu erreichen werden zwei gute Individuen aus der Population vereinigt in der Hoffnung ein noch besseres Individuum hervorzubringen.
  - Mutation

    Die Aufgabe der Mutation ist es neue Varianten und Alternativen hervorzubringen.

Evolutionäre Algorithmen – Artikel 1 – Poker Variante

TECHNISCHE UNIVERSITÄT DARMSTADT

- Kein vollständiges Texas Hold‘em, sondern vereinfachte Form.
  - Eine Wettrunde.
  - Maximal drei Erhöhungen (Raise) in der Wettrunde.

2. Fünf beste Karten aus zwei privaten und fünf öffentlichen.

Evolutionäre Algorithmen – Artikel 1 – Datenstruktur 1/3

TECHNISCHE UNIVERSITÄT DARMSTADT

- Drei Funktionen, die die Spielfunktionen des Spielern bestimmen.
  - fold(x) = exp(-eval(b) * (x - a))
  - call(x) = eval(c) * exp(-eval(b)^2 * (x - a)^2)
  - raise(x) = exp(eval(b) * (x + a – 1.0))

- Diese drei Funktionen bilden einen **Kandidaten**.
- Wobei:
  - x: Unabhängige Variable zugehörig zur Wahrscheinlichkeit ein Spiel zu gewinnen.
  - a: Definiert für die call Funktion den Mittelpunkt der Gaußklocke.
  - b: Kontrolliert die Breite der Funktion.
  - c: Beeinflusst den maximal erlaubten Wert.
  - eval(c): $[0,1] \rightarrow [0,\infty[$

    eval(c) = 1 / (1-c)

Evolutionäre Algorithmen – Artikel 1 – Datenstruktur 2/3

TECHNISCHE UNIVERSITÄT DARMSTADT

- Aufstellen eines Hypercube´s mit zwei Dimensionen.
  - 1. Dimension: Position des Spielers.
    - Unterteilung in drei weitere Sektionen.
      - Frühe Position
      - Mittlere Position
      - Späte Position
  - 2. Dimension: Risikomanagement
    - Unterteilung in vier weitere Sektionen.
      - Null Wetten zum mitgehen.
      - Eine Wette zum mitgehen.
      - Zwei Wetten zum mitgehen.
      - Drei oder mehr Wetten zum mitgehen.

Evolutionäre Algorithmen – Artikel 1 – Datenstruktur 3/3

TECHNISCHE UNIVERSITÄT DARMSTADT

- Durch diese Aufteilung ergeben sich zwölf Elemente im Hypercube.
- Jedes dieser Elemente besteht aus N Populationskandidaten.
- Jeder Populationskandidat besteht aus sieben Zahlen aus dem Intervall [0,1].
  - Zwei für die Konstanten der fold Funktion.
  - Drei für die Konstanten der call Funktion.
  - Zwei für die Konstanten der raise Funktion.
  - Diese Werte, werden sich mit der Zeit entwickeln.

Evolutionäre Algorithmen – Artikel 1 – Algorithmus

TECHNISCHE UNIVERSITÄT DARMSTADT

- Der Spieler ist an der Reihe eine Spielentscheidung zu treffen.

1. Wähle die Population, die sich aus aktueller Spielposition und Anzahl der mitzugehenden Wetten ergibt.
2. Konsultiere die entsprechende Population. Dabei wird nur ein Kandidat befragt, aber es werden alle Kandidaten einer Population der Reihe nach befragt.
3. Das Feedback (wie erfolgreich die Strategie war) wird direkt an die benutzten Kandidaten der Population weitergegeben.

- Evolution geschieht dann, wenn jeder Kandidat eine bestimmte Anzahl oft befragt wurde.
- Es wird eine (1+1) Evolutionsstrategie angewandt.

Evolutionäre Algorithmen – Artikel 1 – Experimentergebnis 1/4

TECHNISCHE UNIVERSITÄT DARMSTADT

- Erstes Experiment - Daten
  - 1000 Einheiten Startkapital
  - 1 Epoch entspricht 500 Händen
  - Jedes Populationsmitglied wird 25 mal befragt bevor Evolution stattfindet.
  - Jede Population besteht aus 20 Kandidaten.
  - Tische *tight*-Tisch und *loose*-Tisch

- Ergbenis
  - Der sich entwickelnde Spieler verliert zu erst.
  - Holt dann wieder seinen Verlust rein.
  - Macht darauf hin wieder Gewinn.
  - Gewinn am *tight*-Table niedriger.

Evolutionäre Algorithmen – Artikel 1 – Experimentergebnis 2/4

TECHNISCHE UNIVERSITÄT DARMSTADT

- Zweites Experiment
  - Die Spieler die am *tight*- und am *loose*-Tisch gelernt haben nennen wir nun A^{tight} und A^{loose}.
  - Wir untersuchen nun den Spielstiel, den sich beide Spieler angeeignet haben.

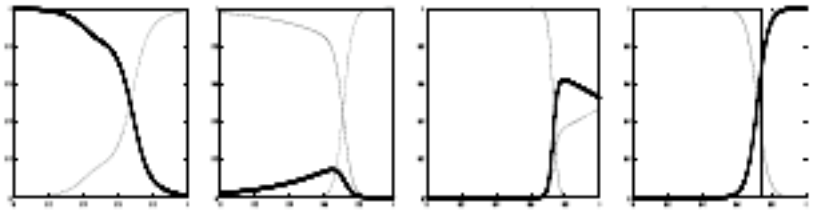
- Wie erwartet A^{tight} spielt eine strenger Strategie. Da er gelernt hat, dass wenn seine Gegner erhöhen, das ein sicheres Anzeichen dafür ist, dass sie eine gute Hand haben.

- A^{loose} spielt 21% seiner Hände verschieden. Gegenspieler können nicht aus dem Spiel geblufft werden. Er erhöht und geht häufiger mit.

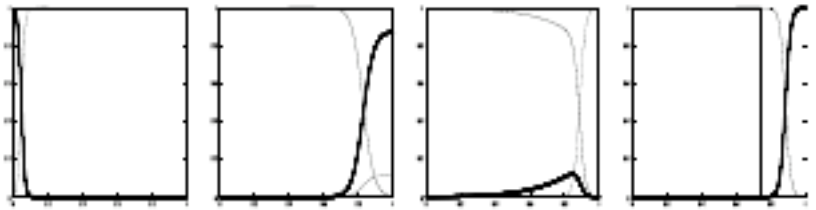
Evolutionäre Algorithmen – Artikel 1 – Experimentergebnis 3/4

TECHNISCHE UNIVERSITÄT DARMSTADT

- Oben: A^{loose} bei später Position und entsprechender Anzahl an Wetten zum mitgehen.
- Unten: A^{tight} ebenfalls in später Position und entsprechender Anzahl an Wetten zum mitgehen.



  - Schwarze dicke Linie: Mitgehen
  - Abszisse: Gewinnwahrscheinlichkeit x
  - Ordinate: Wahrscheinlichkeit der Aktion.



Evolutionäre Algorithmen – Artikel 1 – Experimentergebnis 4/4

TECHNISCHE UNIVERSITÄT DARMSTADT

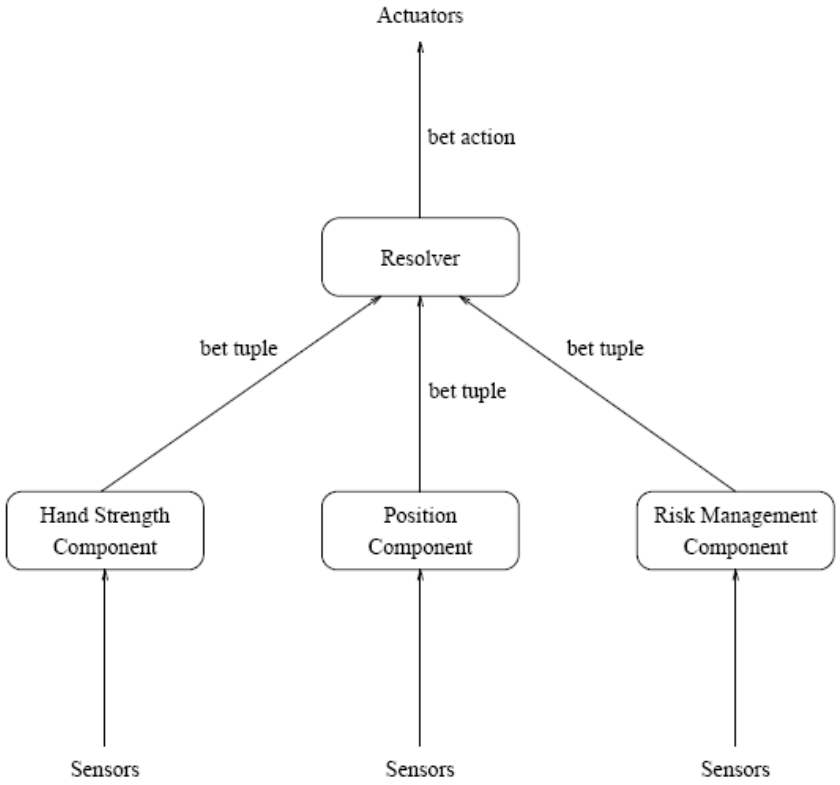
- Drittes Experiment
  - Es wird untersucht ob Spieler, aufgrund ihres Trainings an Tischen gewinnen, die ihrem Training entsprechen.
  - Beispiel: Wird der Spieler A^{tight} an einem tight-Table besser abschneiden als A und A^{loose}.

$W_{tight}(A^{tight})$	2113
$W_{tight}(A)$	-730
$W_{tight}(A^{loose})$	-195
$W_{loose}(A^{loose})$	6255
$W_{loose}(A)$	5400
$W_{loose}(A^{tight})$	4397

Evolutionäre Algorithmen – Artikel 2 – Datenstruktur

TECHNISCHE UNIVERSITÄT DARMSTADT

- Benutzen von Komponenten und einem Resolver.



Evolutionäre Algorithmen – Artikel 2 – Algorithmus

TECHNISCHE UNIVERSITÄT DARMSTADT

- Jede Komponente gibt aufgrund des aktuellen Spielstatus eine Einschätzung mit welcher Wahrscheinlichkeit sie welche Aktion durchführen würde.
- Der Resolver gewichtet diese Meinungen und entscheidet dann eine Spielaktion.
- Sowohl die Einschätzungen der Komponenten sowie die Gewichtung des Resolver werde sich mit der Zeit entwickeln.

- Bsp:

  Bei einem Spiel gegen verschiedene Spieler gewichtet der Resolver 82% Handstärke, 11% Position und 7% Risikomanagement.

  Bei einem Spiel gegen loose-aggressive Spieler gewichtet der Resolver 93% Handstärke, 3% Position und 4% Risikomanagement.

Pareto-Koevolution - Übersicht

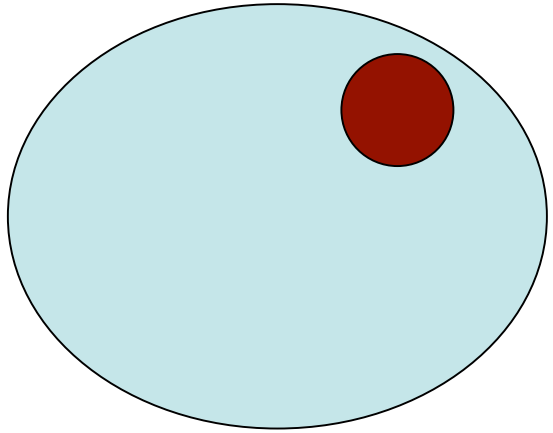
TECHNISCHE UNIVERSITÄT DARMSTADT

- Probleme mit Evolutionären Algorithmen
- Was bedeutet Pareto?
- Pareto-Koevolution in Evolutionären Algorithmen

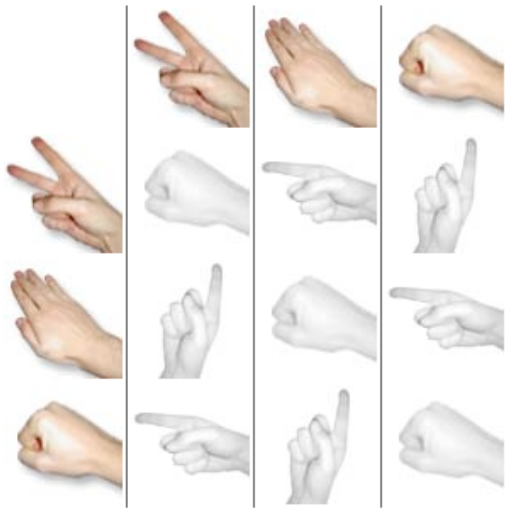
Probleme mit Evolutionären Algorithmen

TECHNISCHE UNIVERSITÄT DARMSTADT

- Problem Evolutionärer Algorithmen
  - Strategien sind häufig auf einen kleinen Ausschnitt des Lösungsraumes spezialisiert.



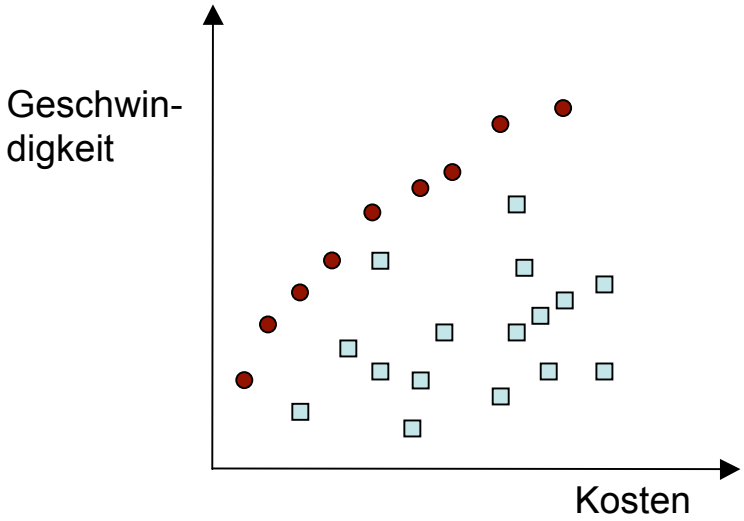
- Erklärungsversuch
  - Nicht transitive Überlegenheitsrelation
    - Strategie A gewinnt gegen Strategie B
    - Spieler 1 spielt immer Stein
    - Spieler 2 spielt immer Schere
    - Spieler 3 spielt immer Papier

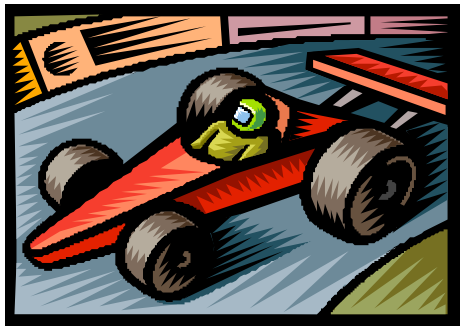


Begriffserklärung - Pareto

TECHNISCHE UNIVERSITÄT DARMSTADT

- Begriff aus der VWL
- Pareto-Optimum
  - Optimaler Kompromiss
  - „Keine Dimension kann verbessert werden ohne eine andere zu verschlechtern“
- Pareto-Dominanz
  - In allen Dimensionen mindestens gleich gut
  - In mindestens einer besser



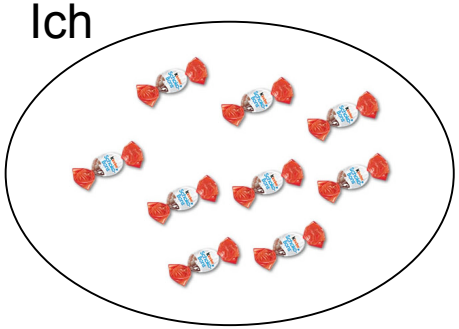


Optimieren mit Hilfe des Pareto-Optimums

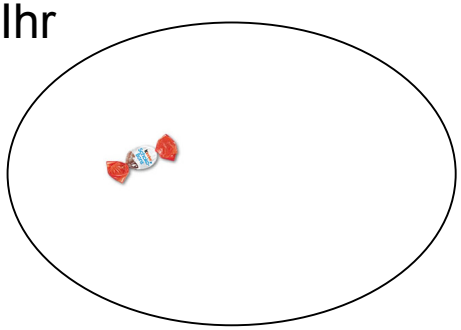
TECHNISCHE UNIVERSITÄT DARMSTADT

- Wir haben eine Tüte Schokobons zu verteilen!

Ich



Ihr



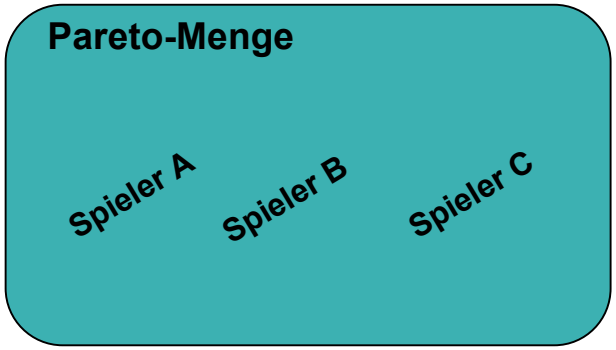
Pareto Optimum (9,1) erreicht!

Aber es existieren sehr viele weitere Pareto Optima.

Pareto-Koevolution in Texas Hold'em Limit

TECHNISCHE UNIVERSITÄT DARMSTADT

- Idee
  - Strategien werden als Dimensionen eines Optimierungsproblemes angesehen.
- Pokervariante
  - Texas Hold'em Limit 2$/4$
- Kodierung der Strategien
  - 2 Wahrscheinlichkeiten (bluff, check-raise)
  - 24 Integer (6 pro Wettrunde)
  - 4 Gruppen mit je 4 Binärwerten
- Aufbau des Experiments
  - Population mit 100 Individuen
  - Tisch mit 10 Spielern für 50 Hände
  - 200 Spiele => Evolution



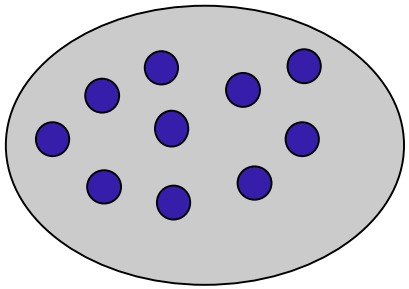
	Spieler A	Spieler B	Spieler C	Spieler D
Spieler A		250$	-100$	-40$
Spieler B	-250$		40$	70$
Spieler C	100$	-40$		30$
Spieler D	40$	-70$	-30$	

Approximierte Matrix

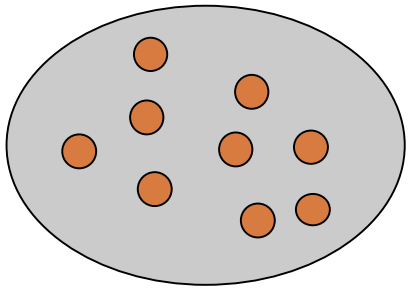
Pareto-Koevloution Ergebnisse Experiment 1

TECHNISCHE UNIVERSITÄT DARMSTADT

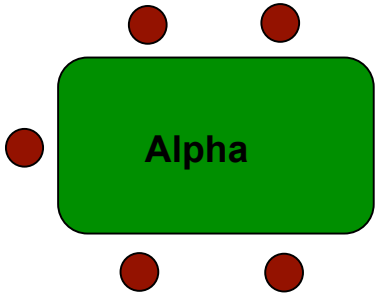
- 2 Referenztische (Alpha & Beta)
  - Je 5 Handkodierte Spieler

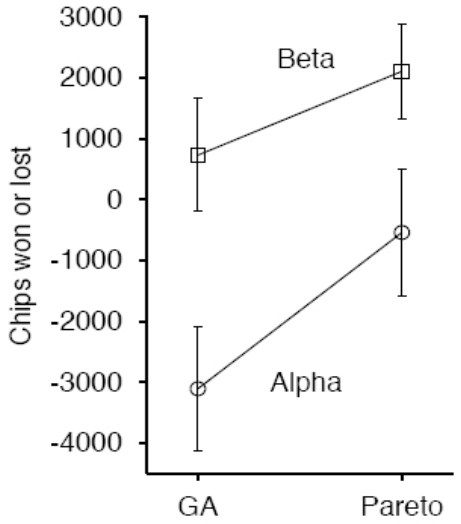


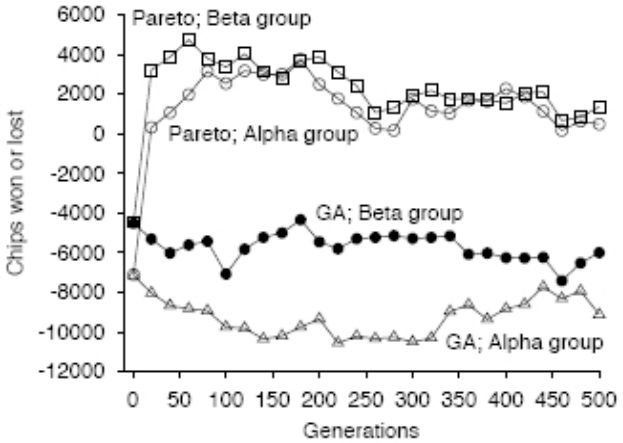
Pareto-Population nach 100 Generationen



Vergleichspopulation nach 100 Generationen



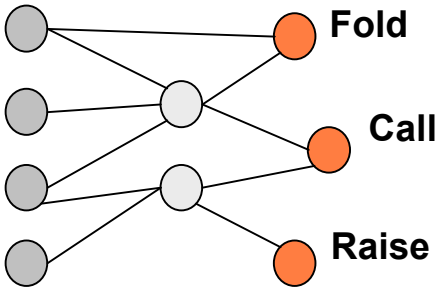




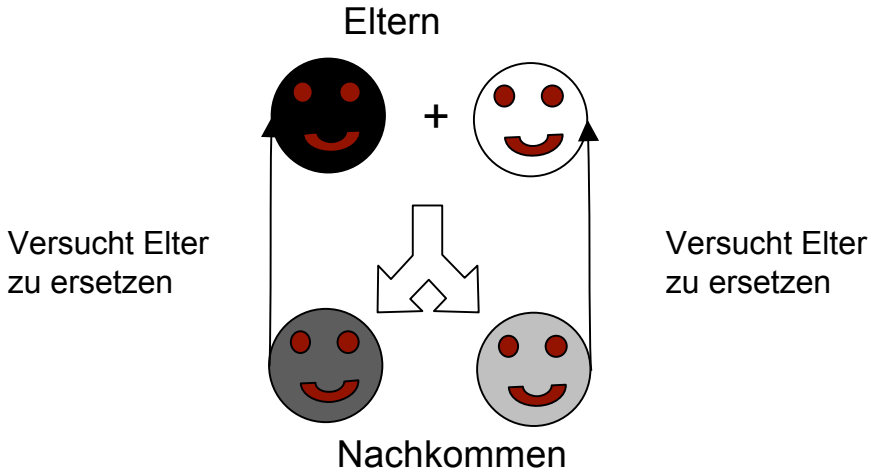
Pareto-Koevolution ein zweiter Versuch

TECHNISCHE UNIVERSITÄT DARMSTADT

- Ähnlich wie 1. Experiment



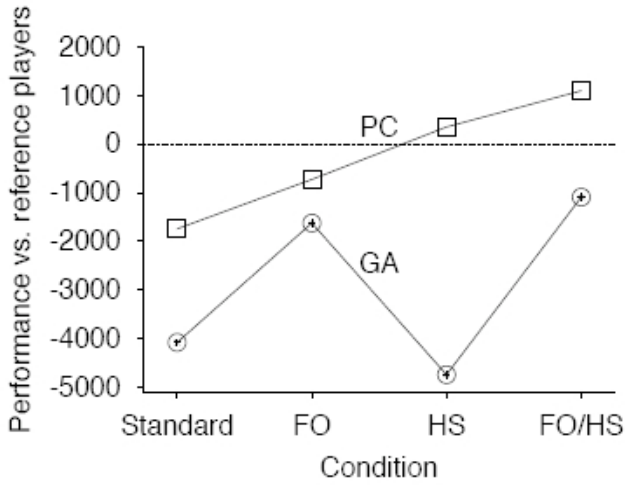
- Darstellung der Strategien
  - 2 Neuronale Netze
    - Pre-Flop mit 69 Eingängen
    - Flop, Turn & River mit 109 Eingängen
- Selektion mit Fehlergrenze von 100$
- Deterministic Crowding
  - „Technik zum Erhalt der Vielfalt“
  - Gene der Eltern werden zufällig auf Nachkommen verteilt.

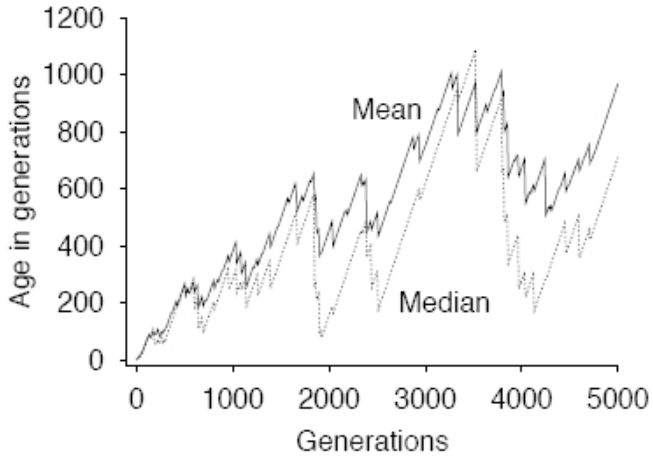


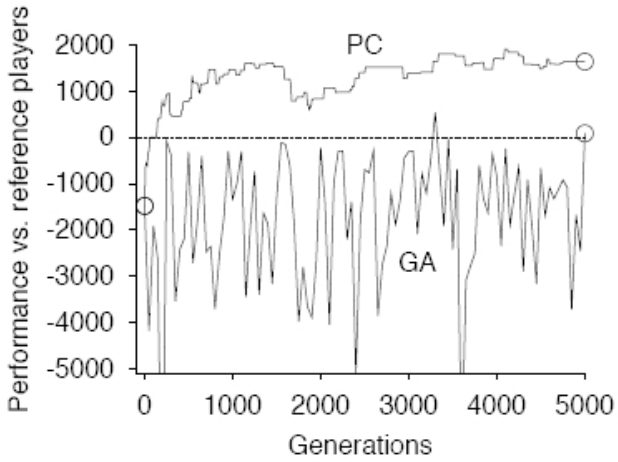
Pareto-Koevolution Ergebnisse

TECHNISCHE UNIVERSITÄT DARMSTADT

- 4 Experimente
  - Standard (Untrainiert)
  - Fixed-Opponent (FO)
  - Head-Start (HS)
  - Fixed-Opponent + Head-Start (FO/HS)







Fazit und Diskussion

TECHNISCHE UNIVERSITÄT DARMSTADT

Vielen Dank

Quellen

TECHNISCHE UNIVERSITÄT DARMSTADT

- Barone, L., & While, L. (1997). Evolving computer opponents to play a game of simplified poker. In Royle, G., & Backman, R. B. (Eds.), 8th University of Western Australia Computer Science Research Conf.
- Barone, L., & While, L. (1999). An adaptive learning model for simplified poker using evolutionary algorithms. In Angeline, P. J., Michalewicz, Z., Schoenauer, M., Yao, X., & Zalzala, A. (Eds.), Proceedings of the Congress on Evolutionary Computation, Vol. 1, pp. 153{160 Mayower Hotel, Washington D.C., USA. IEEE Press.
- Kendall, G., & Willdig, M. (2001). An investigation of an adaptive poker player. In Australian Joint Conference on Articial Intelligence, pp. 189-200.
- Noble, J. (2002). Finding robust texas hold'em poker strategies using pareto coevolution and deterministic crowding.
- Noble, J., &Watson, R. A. (2001). Pareto coevolution: Using performance against coevolved opponents in a game as dimensions for pareto selection. In Spector, L., Goodman, E. D., Wu, A., Langdon, W. B., Voigt, H.-M., Gen, M., Sen, S., Dorigo, M., Pezeshk, S., Garzon, M. H., & Burke, E. (Eds.), Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001), pp. 493{500 San Francisco, California, USA. Morgan Kaufmann.
- Wikipedia (2008a). Evolutionärer Algorithmus - wikipedia, die freie Enzyklopädie. [Online; Stand 27. März 2008].
- Wikipedia (2008b). Künstliches neuronales Netz - wikipedia, die freie Enzyklopädie. [Online;Stand 26. März 2008].
- Wikipedia (2008c). Koevolution - wikipedia, die freie Enzyklopädie. [Online; Stand 25. März 2008].
- Wikipedia (2008d). Pareto-optimum - wikipedia, die freie Enzyklopädie. [Online; Stand 25. März 2008].