



# **Strategien bei der Entwicklung und Modellierung von Poker-Agenten**

***Andreas Eismann***

***TU Darmstadt – Technische Universität Darmstadt***

***Fachbereich Informatik***

***Knowledge Engineering***



## Gliederung

1. Einleitung / Einordnung
2. Pokerstrategien
3. Spielsituationen
4. Modellierungsstrategien
  - Not-To-Lose / Play-To-Win
  - Restricted Nash Response
5. Fazit / Ausblick



## 1. Einleitung

Texas Hold'em Poker kurzgefasst:

- 4 Spielrunden (Preflop, Flop, Turn, River)
- Jeweils 3 Aktionsmöglichkeiten (fold, call, bet/raise)
- Varianten von Texas Hold'em:
  - Limit/No-Limit
  - Heads-Up/Ring
- Wichtige Verhaltensweisen: (semi-)bluff, trapping, value bet



## 1. Einordnung

- Pokerspielen ist Annäherung der Lösung eines Spieltheoretischen Problems
- Versuch ein Nash-Gleichgewicht zum Gegner zu erzielen
- Lösungsansätze aus dem Bereich der Stochastik sowie Operations Research



## 2. Pokerstrategien

- Gegenspielerabhängige Spielstrategie
- In vielen Situationen Abwägung zwischen
  - Durchschauen des Gegners und Ausnutzen dieses Wissens
  - Eigene Durchschaubarkeit durch diese Aktionen
- 2 strategische Ansätze
  - Mit Wahrscheinlichkeiten gewichtete Strategien
  - Start mit Nashgleichgewichtsstrategie mit Verfeinerung



## 2. Pokerstrategien

- Verfahren von Poker-Agenten:
  - Spiel von einigen tausend Spielen pro Sekunde
  - Reduzierung der Varianz durch Wiederholung von Spielen
- Performanz-Maß eines Pokeragenten:
  - Small bets per game (Limit-Variante)



## 3. Spielsituationen

- Erfassung von Spielsituationen als Poker-Baum
  - Jeder Knoten im Baum stellt eine Entscheidung dar
  - Bei einfacher Strategie ruft gleiche Kartenkonstellation selbe Aktion hervor
  - Bei verhaltensabhängiger Strategie wird anhand der Gewichtung eine Aktion ausgewählt
- Sehr große Bäume entstehen
  - $3 \cdot 10^{17}$  Spielzustände,  $3 \cdot 10^{14}$  mögliche Spielzüge
  - Zeit als entscheidender Faktor



## 3. Spielsituationen

- Abstraktion durch Bucketing
  - Handhabbare Größe von Spielsituationen und Strategien
  - Abnehmender Grenznutzen bei großer Anzahl von Eimern
- Gegenzugstrategien teuer, Bsp: „best response“
- Erschaffung mehrerer Agenten (Team), anschließend:  
Auswahl aus dem Team anhand von Erfahrungswerten





# 4. Modellierungsstrategien

- „Nicht verlieren“ mit „Counterfactual Regret Minimization“
- immer bessere Annäherung an Nash-Gleichgewicht
- „Gewinnen“ -> wichtig im Wettbewerb
- Wahl der Abstraktionsebene, je höher, desto größerer Erwartungswert des Ergebnisses (millibets/game)



## 4. Modellierungsstrategien

- Restricted Nash Response (RNR)
- Berücksichtigt Strategieänderungen des Gegners
- Aufgabe des Algorithmus:
  - Durchschauen der Strategie(-dynamik) des Gegners
  - Anpassung der eigenen Strategie
  - Gegnerstrategie wird versucht vorauszusagen
- Weitere Verbesserung: Teams von RNR-Agenten
- Auswahl des Agenten, der in Vergangenheit am besten in ähnlicher Konstellation abgeschnitten hat



## 5. Fazit / Ausblick

- Neue Arbeitsfelder seit AAAI Poker Challenge 2007
  - Parallele Berechnung auf mehreren Rechnern
  - Erhöhung des Abstraktionslevels
- Viele Merkmale müssen bei Modellierung beachtet werden
- Insbesondere ist der Zeitfaktor zu beachten
- Ein Blick in die Ideensammlung der Arbeit lohnt evtl. zur Verbesserung der eigenen Implementierung



## Backupfolie Regret

- Eigene Wkeit: (0.5, 0.25, 0.25) (fold, call, raise)
- Gegner Wkeit: (0.33, 0.33, 0.33)
- Nutzen (-5, 10, 20)
- $E(x) = 0.5 \cdot -5 + 0.25 \cdot 10 + 0.25 \cdot 20 = 5$
- Regret (-10, 5, 15)
- Gewichtung mit Gegner Wkeit: (-1.33, 1.33, 5)
- Neue Wkeiten (0,  $1.33/5=0.2666$ ,  $1-0.2666=0.7333$ )
- ...
- ...