

TUD Computer Poker Challenge



TECHNISCHE
UNIVERSITÄT
DARMSTADT

The Challenge of Poker

Björn Heidenreich

Anforderungen an einen guten Poker-Spieler

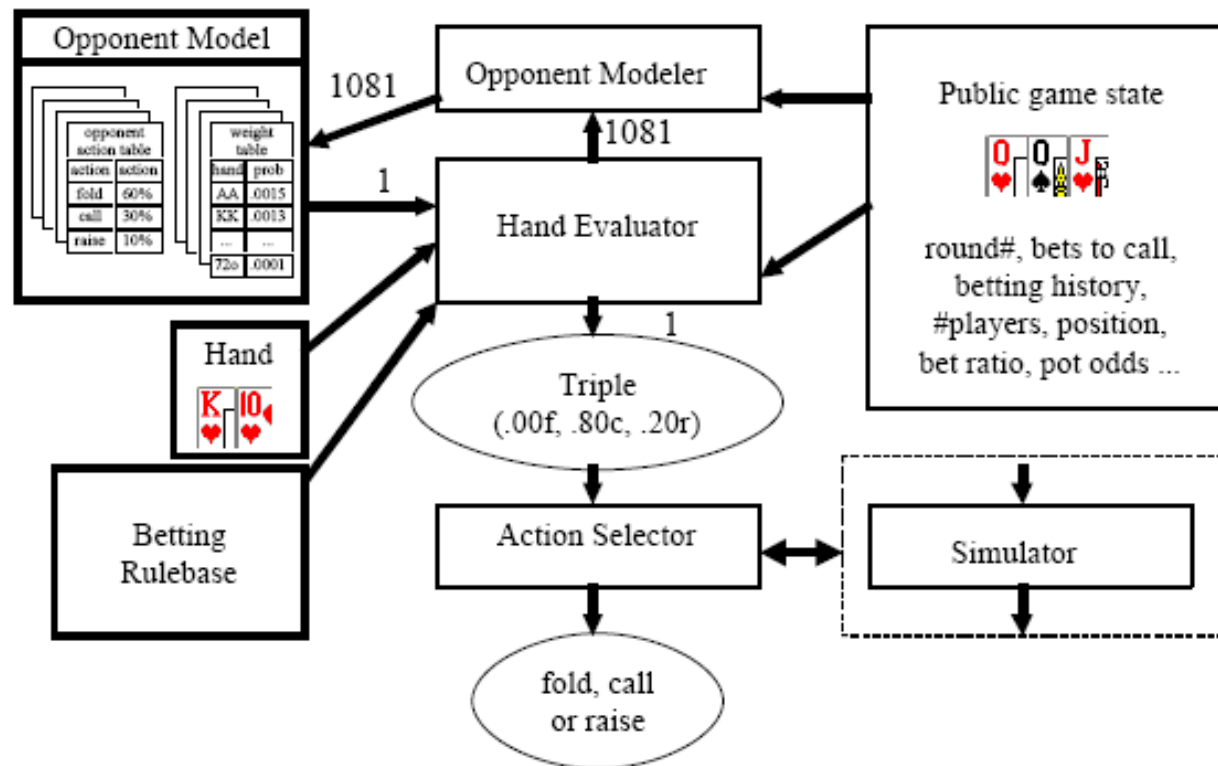


TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Hand Strength
- Hand Potential
- Bluffing
- Unberechenbarkeit
- Gegner-Modellierung

Pokis Architektur

- entwickelt an der University of Alberta





- Pre- und Post-Flop-Strategien unterscheiden sich stark:
 - Vor dem Flop sind nur wenige Informationen bekannt.
 - Post-Flop-Strategien werden durch mehr Faktoren determiniert.

Pre-Flop-Strategie



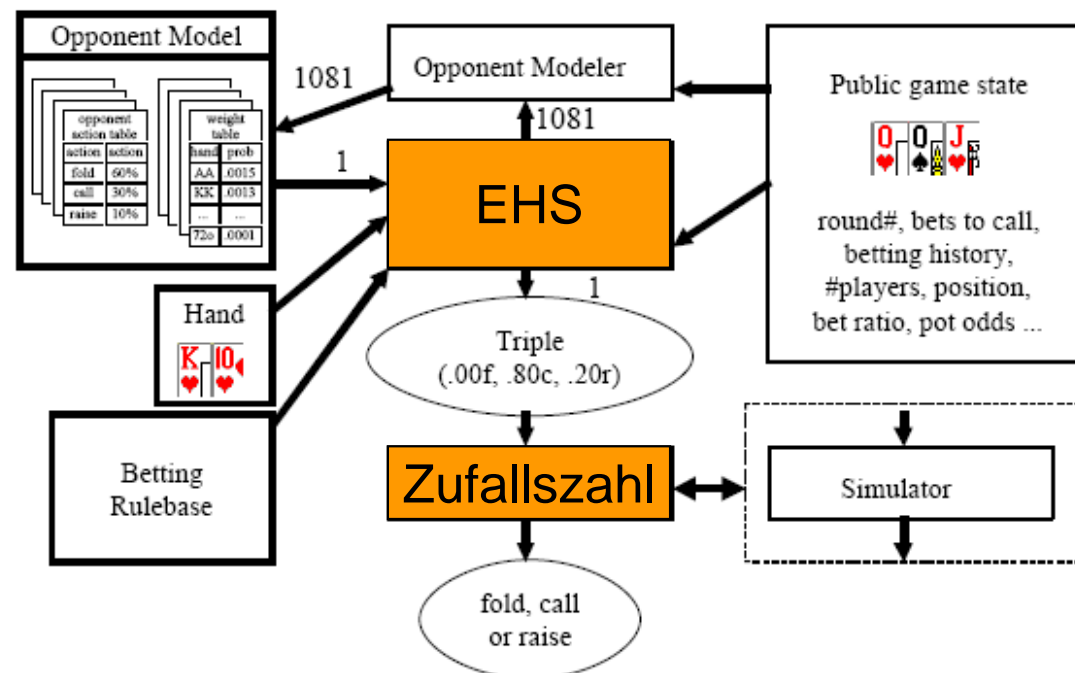
- 1326 Möglichkeiten für die private Hand
- Berechnung der income-rate mittels roll-out Simulation
 - always-call assumption
 - Nur relative Gewichtung in der Pre-Flop-Phase
 - kann für verschiedene Spielsituationen simuliert werden
- Iterierte roll-out Simulation
 - Mehrmalige roll-out Simulation, wobei vorherige Ergebnisse die aktuelle Simulation beeinflussen
 - Realistischere Ergebnisse, da Spieler aus dem Spiel ausscheiden können
 - Simulation so lange, bis sie gegen ein Ergebnis konvergiert
 - weitere Verbesserung durch Noise-Faktor



Post-Flop-Strategien

mögliche simple Strategie:

1. Effective Hand Strength (EHS) von Pokis Hand berechnen
2. EHS in Wahrscheinlichkeitstripel übersetzen
3. Zufallszahl zwischen 0 und 1 generieren und mit ihr die Aktion auswählen



Hand Strength (HS)

- Wahrscheinlichkeit mit der eine Hand besser als die des Gegners ist
- Zählen der Kartenpaare die besser, gleich gut oder schlechter sind als die eigene Hand
- Gewichtung mit
 - 1, falls die eigene Hand besser ist
 - 0,5 falls beide gleich gut sind
 - 0, falls die gegnerische Hand besser ist
- Bei mehreren Gegnern kann die einfache HS mit der Anzahl der Gegner potenziert werden, um die Mehrspieler-HS zu erhalten.

Hand Potential (HP)

- Potential der Hand
 - Positive Potential (PPot):
Wahrscheinlichkeit, dass eine Hand, welche momentan nicht vorne liegt, im Showdown gewinnt
 - Negative Potential (NPot):
Wahrscheinlichkeit, dass eine momentan führende Hand im Showdown verliert

Effective Hand Strength



- Kombiniert Hand Strength und Hand Potential
- $P(\text{Gewinn}) = HS \cdot (1 - \text{NPot}) + (1 - HS) \cdot \text{PPot}$
- $\text{NPot} = 0$
 $\text{EHS} = HS + (1 - HS) \cdot \text{PPot}$
- Mehrere Gegner
 $\text{EHS}_i = \text{HS}_i + (1 - \text{HS}_i) \cdot \text{PPot}_i$

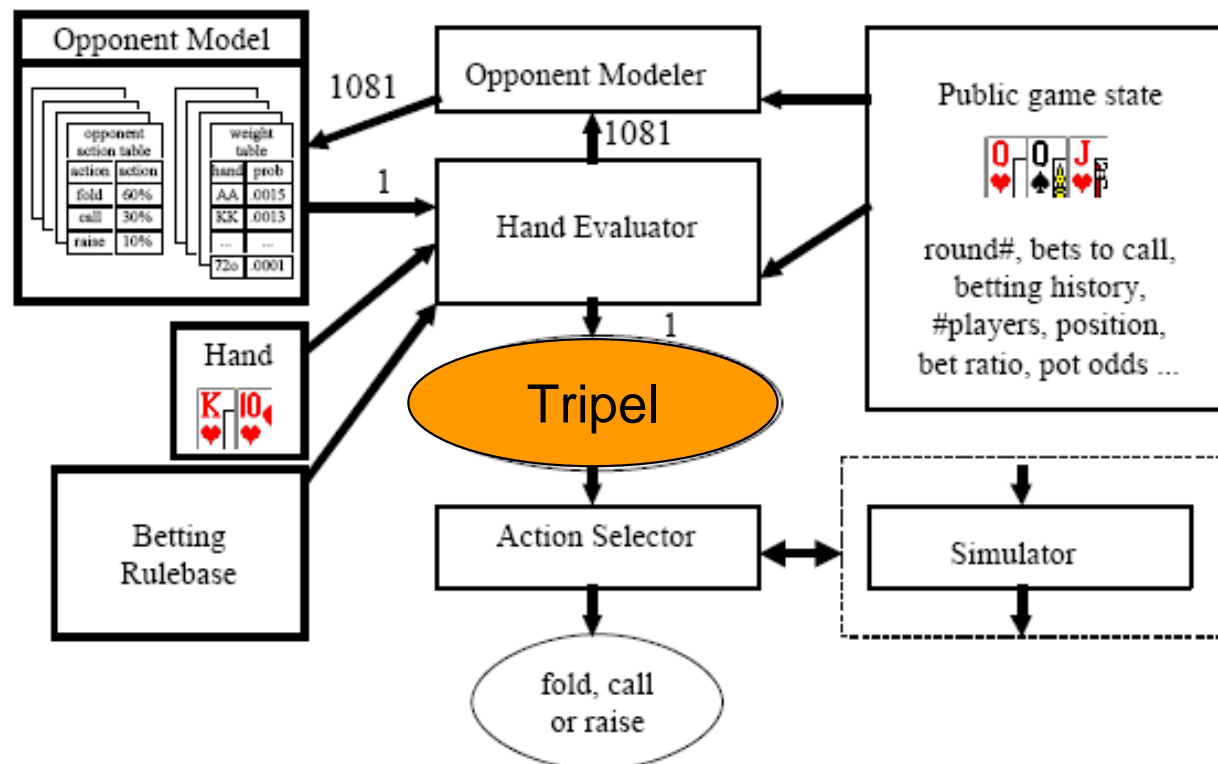
Weight Tables

- Bisherige Berechnung von HS und HP geht davon aus, dass Kombinationen von Kartenpaaren gleich wahrscheinlich sind.
- Die Wahrscheinlichkeit, dass ein Gegner nach dem Flop Ass und König auf der Hand hat, ist größer, als dass er eine sieben und eine zwei besitzt; da die meisten Spieler mit sieben und zwei vor dem Flop folden würden.
- Einführung von weight tables, welche die Wahrscheinlichkeiten der Kartenpaare zu einem Zeitpunkt des Spiels beinhalten:
 - Gewicht zwischen 0 und 1
 - Gewichtung mit 1 am Anfang
 - Werden die Kartenpaare im Laufe des Spiels unmöglich, wird das Gewicht auf 0 gesetzt.
- Laufende Anpassung der Gewichte



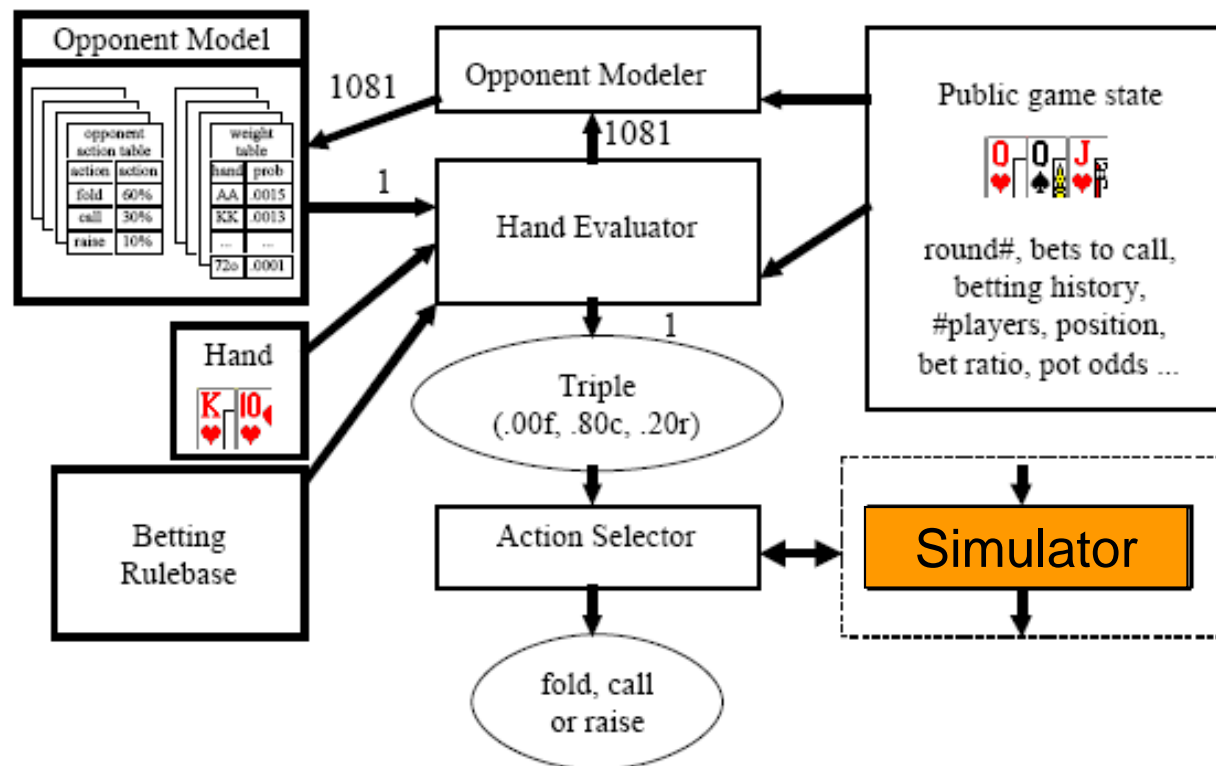
Wahrscheinlichkeits-Tripel

- Wahrscheinlichkeit für Fold, Call und Raise
- $P = \{P(\text{fold}), P(\text{call}), P(\text{raise})\}$
- $P(\text{fold}) + P(\text{call}) + P(\text{raise}) = 1$



Simulation

- Die Simulation ersetzt bei Poki den Action Selector



Simulations-basierte Strategien

- Poker ist sehr komplex und Entscheidungen müssen im jeweiligen Kontext getroffen werden.
- Man benötigt dynamische, adaptive Techniken.
- Verwendung eines Simulators jedesmal wenn eine Entscheidung getroffen werden muss.
- Simulation vieler Szenarien und Berechnung, wieviel Geld gewonnen oder verloren wird:
 - Vom aktuellen Kontext ausgehend werden verschiedene Alternativen bis zum Ende simuliert.
 - Jede Runde wird zwei Mal simuliert um die Konsequenzen von check oder call und bet oder raise zu simulieren.
 - Die gegnerischen Karten werden hierbei entsprechend der weight table angenommen.

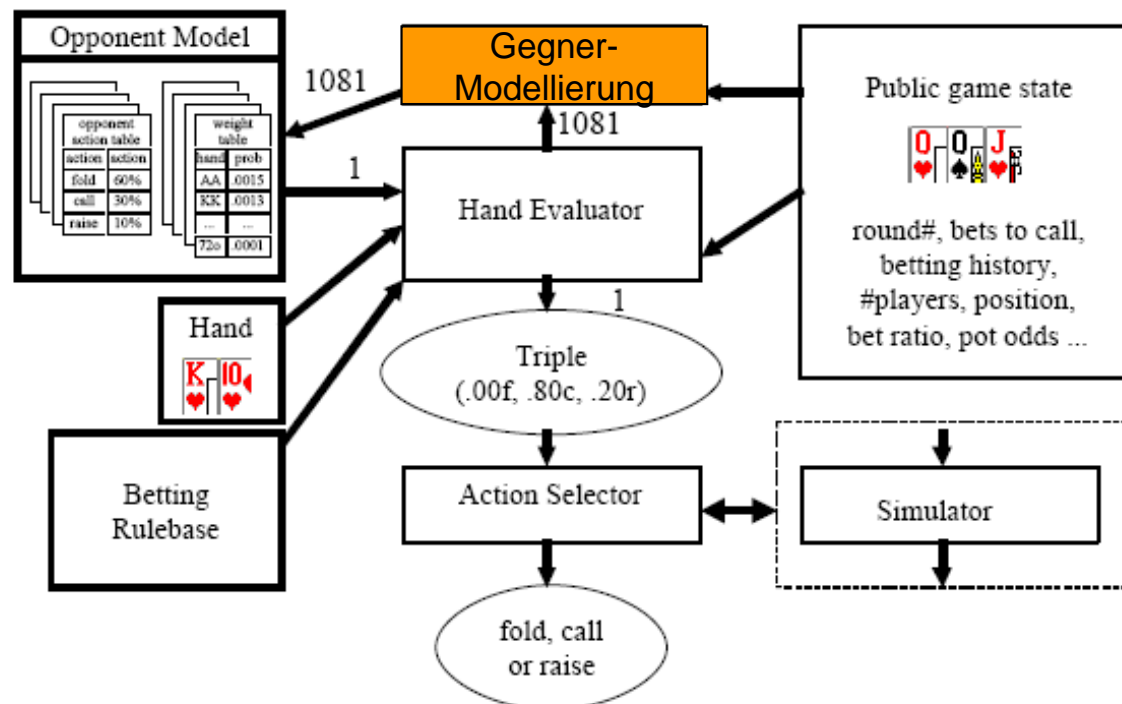
Simulations-basierte Strategien



- Aktuelle Implementierung von Poki wählt die Aktion mit der höchsten Gewinnerwartung aus.
- Bei gleicher Gewinnerwartung wird die aggressivere gewählt.
- Steigerung der Unvorhersehbarkeit des Programms bspw. durch zufällige Auswahl der Aktion bei ähnlich hohe Gewinnerwartungen.
- Problem:
Aufgrund der Komplexität können nicht alle Kartenverteilungen und Aktionen der Gegner simuliert werden.
- Lösung:
Verwendung einer Gegner-Modellierung, um dessen Aktionen vorhersagen zu können.

Gegner-Modellierung

- Gegner-Modellierung ist nötig, um dessen Schwächen identifizieren und ausnutzen zu können.
- Da jeder Gegner andere Strategien benutzt und diese gegebenenfalls im Laufe des Spiels wechselt, muss diese Modellierung adaptiv erfolgen.



Statistisch-basierte Gegner-Modellierung



- generic opponent modeling:
Unterstellung der eigenen Spielstrategie
- specific opponent modeling:
vergangenes Verhalten des Spielers wird unterstellt
- Beispiel:
 - Gegner macht in 40% der Fälle direkt nach dem Flop seinen Einsatz.
 - Es lässt sich annehmen, dass er in dieser Spielsituation mit den Top 40% der Kartenpaare einen Bet ausführt.
- Problematik:
 - Wird der Kontext zu weit definiert, so lernt das System vielleicht nicht alle wichtigen Aspekte des Gegners.
 - Wird er zu eng definiert, so lernt das System nur sehr langsam und manches vielleicht überhaupt nicht, da die zu lernenden Situationen zu selten auftreten.

Gegner-Modellierung mittels neuronalen Netzen



- Poki implementiert ein neuronales Netz zur Vorhersage der gegnerischen Aktionen.
- Input:
Eigenschaften des Spiel-Kontexts, welche sich auf die Spieler-Entscheidungen auswirken können oder mit ihnen korreliert sind
- Output:
drei Knoten, welche für fold, call und raise stehen
- Manche Input-Knoten sind dominant, andere spielen nur eine kleine Rolle.
- Messung der Vorhersagegenauigkeit mittels cross-validation der realen Spielzüge.
- Man erhält so eine relativ kleine Klasse an wichtigen Faktoren, welche die statistische Gegner-Modellierung signifikant verbessert.

Performance Evaluation



- Poki ist ein komplexes System, wo kleine Änderungen große Auswirkungen haben können.
- Vergleich der Software schwer, da Glück eine große Rolle spielt.
- Adaptive Systeme, welche sich auf die Gegner einstellen.
- Neue Software-Features lassen sich testen, indem man die Software gegen ihre alte Version spielen lässt.
- Software kann verschiedenartigste Spielstile nicht abdecken, daher besser gegen menschliche Gegner spielen lassen.
 - Ein gutes Resultat kann dann allerdings darin begründet liegen, dass nur gegen schwache Gegner gespielt wird.

Fazit

- Poker ist ein komplexes Spiel mit vielen Anforderungen. Um erfolgreich zu sein, müssen die Aktionen des Spielers einerseits unvorhersagbar sein, andererseits müssen die gegnerischen Aktionen jedoch bestmöglich vorhergesagt werden.
- Bei der Entwicklung von Poki wurde eine zyklische Vorgehensweise benutzt. Eine Komponente wurde so lange verbessert, bis eine andere den Performanzengpass darstellte.
- Es ist möglich umfangreiche Daten über seine Gegner zu sammeln. Das Problem besteht darin, diese auszuwerten und die nützlichen Features zu finden.
- Momentan besteht in diesem Bereich noch viel Forschungsbedarf.

Ende



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Für Fragen stehe ich gerne zur Verfügung.