

Abbildung eines 2-Spieler Computerspielers auf 3-Spieler Limit- Poker



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Peter Glöckner

Gliederung

- 1.Motivation
- 2.Poker Spielregeln
- 3.Counterfactual Regret Minimization (CFR)
- 4.Abbildungen
- 5.Erweiterte Abbildungen
- 6.Probleme bei den Abbildungen
- 7.Fazit

- Verschieden Typen von Strategien im Computerpoker
 - Mathematisch faire
 - Gegner ausnutzende (opponent modelling)
 - nash-optimale
 - ...
- CFR
 - Generiert Abbildung von Spielzustand zu Aktionswahrscheinlichkeiten
 - Konvergiert gegen eine nash-optimale Strategie für ein zwei Spieler Spiel
 - Starke aber nicht nash-optimale Strategien für mehr als zwei Spieler
 - Lange Berechnungszeiten

Motivation

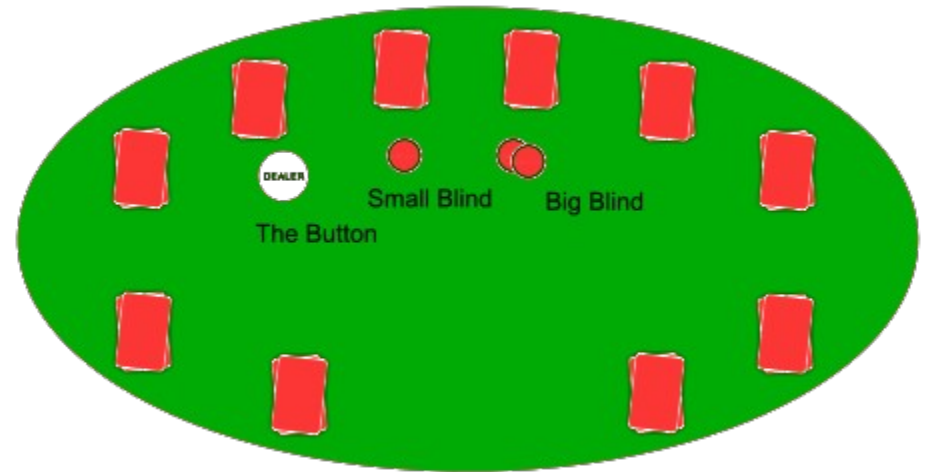
- Entwicklung einer Abbildungen von drei nach zwei Spieler Zuständen
 - Starke Spieler wie CFR Bot können ohne Aufwand in einem Spiel mit mehr Spielern verwendet werden

Poker Spielregeln: Texas Hold'em



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- Positionen im 3 Spieler Spiel
 - Small Blind, Big Blind, Button
- 2 private Handkarten
- 4 Wettrunden
 - Preflop, Flop, Turn, River
- Bis zu 5 Gemeinschaftskarten
- Aktionen
 - Fold, call, raise
- Limit Poker
 - Fest geschriebene Einsatzgrößen
 - Feste Anzahl von Erhöhungen pro Wettrunde



[1]

- Zwei Strategien als abstrahierte Spielbäume (Dealer, Big Blind)
 - Initialisiert mit gleicher Wahrscheinlichkeit für alle Aktionen an den Entscheidungsknoten
- Iterativer Algorithmus
 - Simulation eines Spiel mit jedem möglichen Spielverlauf in einer Iteration
 - Anpassen der Aktionswahrscheinlichkeiten für jeden Spieler an den Entscheidungsknoten, um einen Verlustwert (regret) zu minimieren
 - Volle Kenntnis über die Strategie des Gegners
 - Verlustwerte an den Knoten sind eine obere Schranke für die Verluste des Spielers
- Konvergiert gegen eine nicht auszunutzende Strategie

- In einem drei Spieler Spiel kann CFR schon in einfachen Spielen keine nicht auszunutzende Strategie ermitteln
 - Die Spieler leisten vor jeder Runde einen festen Einsatz
 - Jeder wählt ohne Wissen der anderen Kopf oder Zahl
 - Nach der Wahl werden die Münzen vorgezeigt
 - Bei der gleichen Wahl aller Spieler erhalten die Spieler ihren Einsatz zurück
 - Hat ein Spieler nicht die Wahl der anderen, teilen sich die Gegner seinen Einsatz
- Das Verfahren ermittelt die Strategie $(\frac{1}{2}, \frac{1}{2})$

- Möglichste einfach Abbildung eines Spielzustandes
 - Hero ist der Button mit A♠K♠
 - Flop Q♠J♠10♠, Turn Q♣, Turn J♣
 - Aktionen: [r,f,c], [c, r, r, r, c], [r, r, c], []
- Abbildung 1
 - Möglichste einfache Abbildungen eines Spielzustandes S
 - Löschen eines Spielers
 - Hand- und Gemeinschaftskarten können übernommen werden
 - [r,f,c] → ?
 - [c,r,r,r,c] → ?
 - [r,r,c] → ?
 - [] → ?

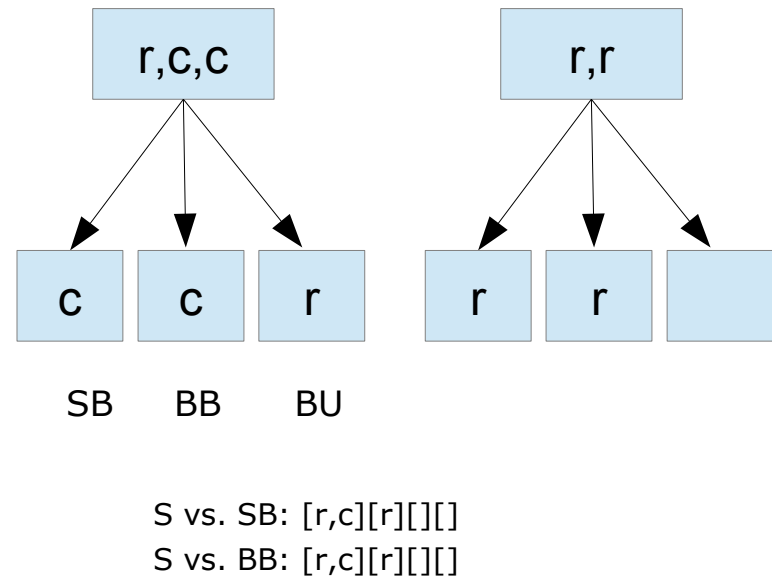
- Anforderungen an die Abbildung der Wettrunden
 - Keine Abbildung für abgeschlossene Spiele
 - Abgeschlossene Wettrunden bleiben abgeschlossen, nicht abgeschlossene nicht abgeschlossen
 - Wettrunden mit zu Anfangs 2 aktiven Spielern werden nicht verändert
 - Umformen der Wettrunden über das minimale Löschen von Einträgen

- $[r,f,c] \rightarrow [r,c]$
- $[c,r,r,r,c] \rightarrow [c,r,r,r,c]$
- $[r,r,c] \rightarrow [r,r,c]$
- $[] \rightarrow []$

- Probleme der ersten Abbildung
 - 6.378 zwei Spieler, 12.827.520 drei Spieler Spielverläufe
 - Es wird nicht berücksichtigt wer eine Aktion ausgeführt hat
- Abbildung 2
 - Für jeden verbleibenden Gegner einen Spielzustand
 - Suche nach dem zwei Spieler Zustand, der den Aktionen des Gegners entspricht

Abbildungen

- Beispiel für die Distanzberechnung beim Spielverlauf: $[r, c, c], [r, r], [], []$
- Hero ist der Button
- Angegeben sind zwei Spielverläufe mit Distanz 0
- Aktionswahrscheinlichkeiten für beide Zustände, die der CFR-Bot liefert, werden gemittelt



- Abbildung 3 ähnlich zu Abbildung 2
 - Aufteilen der Zustände in Klassen
 - Alle Mitglieder einer Klasse besitzen die gleiche Potgröße
 - Bei der Auswahl des zwei Spieler Zustandes wird nur in den Klassen mit der ähnlichste Potgröße zum aktuellen Spielzustand gesucht

- Besonderheiten im Computerpoker
 - Partie besteht aus 1000 Händen
 - Wiederholungen mit geänderter Startposition der Spieler
 - 6 mögliche Positionierungen der Spieler am Tisch
 - In einer Wiederholung unterscheiden sich die ausgeteilten Karten nicht
 - Jeder Spieler hatte dieselben Karten bei denselben Positionen aller Spieler

0	1	2
0	2	1
1	0	2
1	2	0
2	0	1
2	1	0

Evaluierung

- Turnier gegen Benchmark Bots
 - Jeder Spielt eine Partie gegen jede mögliche Kombination von Gegnern
 - $\binom{7}{3} = 35$ Partien
 - Akuma: Zweiter TUD Poker Challenge 2011 und 2012 (Monte Carlo Simulation mit opponent-modelling)
 - ownbot: Sieger TUD Poker Challenge 2011

ownbot	RaiseBot	akuma	HBB	HBB2	CallBot	HBB3
179020	150998	110394	-8308	-16843	-102782	-312479
109522	-12697	77406	-26042	-39605	-108674	
25283	19020	36038	-43008	-37363		
10391	-10962	19969		-19398		
671	-7230	6559				

- Durch das Erstellen eines Zustandes für jeden Gegner findet keine Anpassung an das drei Spieler Spiel statt
 - Preflop Aktionliste $[r,r]$ erzeugt zwei Spielzustände mit jeweils der Aktionsliste $[r]$
 - Die Wahrscheinlichkeit gegen zwei Gegner zu gewinnen ist allerdings geringer
 - In geeigneten Situation muss ein Bias die Foldwahrscheinlichkeit erhöhen
- Einbeziehen mehrere nächster Zustände beim Errechnen der Aktionswahrscheinlichkeiten
 - Analysieren einer Abbildung durch den Vergleich der zurückgelieferten x nächsten Zustände

- Potodds (Größe des Pots : nötiger Einsatz bei einem call) möglicherweise aussagekräftiger für einen Spielzustand als Größe des Pots
- Es werden Zustände, in denen kein raise möglich ist, auf Zustände, in denen dies möglich ist, abgebildet
- Die Abbildungen erhalten nicht die Eigenschaft wer in Position spielt und wer nicht
 - Ein Spieler ist in einer Wettrunde in Position, wenn er als letzter an der Reihe ist

Erweiterte Abbildungen

- Abbildung 4
 - Es werden nur Zustände betrachtet, die die ähnlichsten Potodds besitzen
 - Ist im aktuellen Spiel ein raise möglich bzw. nicht möglich werden alle zwei Spieler Zustände, in denen dies auch möglich ist bzw. nicht möglich ist, bevorzugt
 - Ansonsten werden Zustände, die einer bevorzugte Positionierung entsprechen, favorisiert
 - Bei Zustände, die in diesen Eigenschaften gleich sind, entscheidet die zuvor beschriebene Distanz in den Aktionen

Bevorzuge Position

Hero / Villan	SB	BB	BU
SB	x	?	BB
BB	?	x	BB
BU	SB	SB	x

Erweiterte Abbildungen



- Für jeden verbleibenden Gegner werden Aktionswahrscheinlichkeiten berechnet und gemittelt
- Abbildung liefert für jeden Gegner mehrere Spielzustände zurück
 - Der CFR-Bot liefert Aktionswahrscheinlichkeiten für die Spielzustände
 - Die Aktionswahrscheinlichkeiten werden gemittelt oder.
 - In Abhängigkeit zu ihrer Distanz gewichtet
 - Numerischer Wert für die Distanzen von Abbildung 4

Distanzberechnung Abbildung 4

Raise möglich	t	t	f	f
Position	SB	BB	BB	BB
Distanz	4	2	3	4
	4	6	9	10

$$a = \frac{\sum \frac{1}{1+d_i} * a_i}{\sum \frac{1}{1+d_i}}$$

Erweiterte Abbildungen

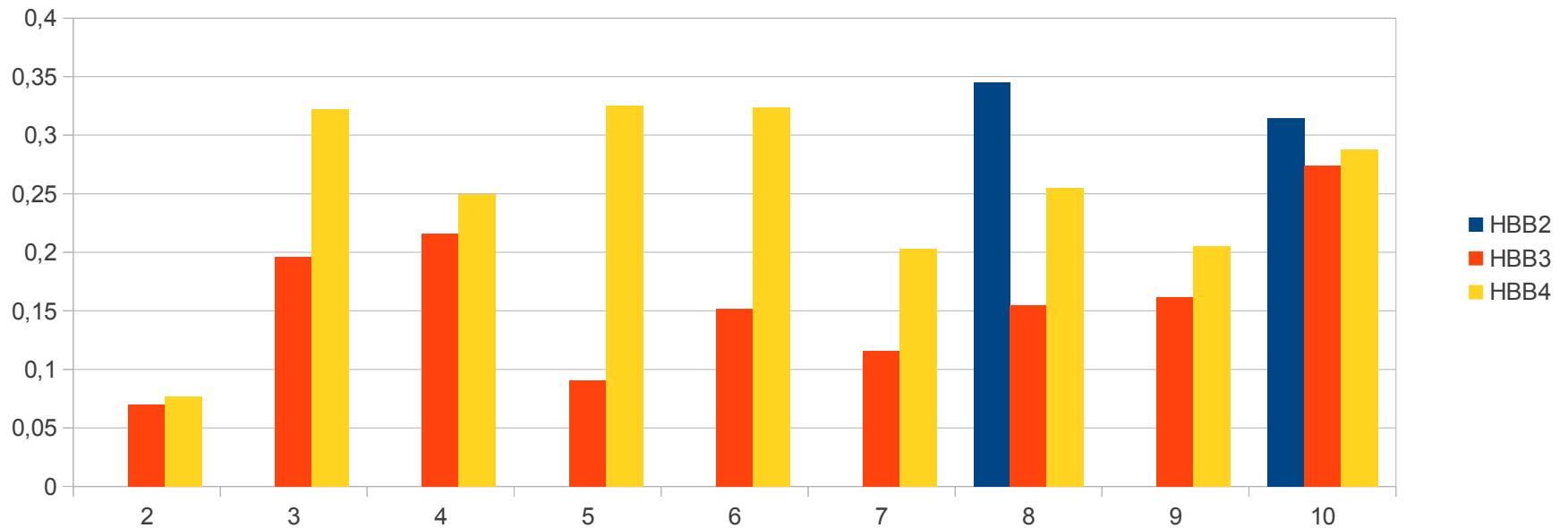
- Tabelle 1: Partien mit x nächsten Spielzuständen gegen die vorherigen Sieger (akuma, ownbot) ohne unterschiedliche Gewichtung der nächsten Spielzustände
- Tabelle 2: Partien gegen zwei mathematisch faire Spieler mit unterschiedlicher Gewichtung

x	HBB2	HBB3	HBB4
1	940	-2247	-2180
2	-2821	-4462	-3624
5	-4421	-6014	-3911
10	-5312	-4945	-5956
15	-6188	-7319	-6166

x	HBB2	HBB3	HBB4
1	-1884	-789	-504
2	-4300	-2558	-2731
5	-4975	-4121	-3065
10	-6276	-4122	-3733
15	-5714	-3864	-3153

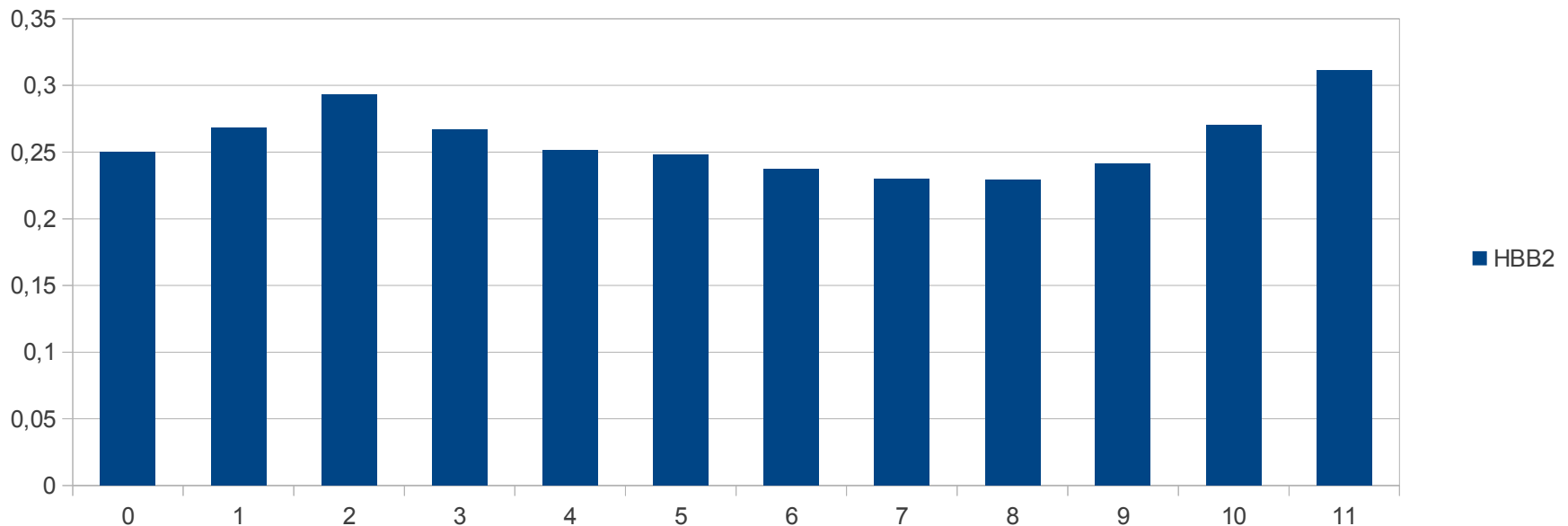
Varianzanalyse

- Varianz in den raise Wahrscheinlichkeiten bei einer gewissen Anzahl von nächsten Zuständen



Distanzanalyse

- Ein Diagramm berücksichtigt von der Abbildung erstellten Spielzustände aus vergleichbare Spielsituationen
- Verhältnis zwischen Abweichung vom Durchschnitt und von der Abbildung berechneten Distanz
 - bei guten Abbildungen sollten Zustände mit geringer Distanz repräsentativ sein



Bias

- Spielzustand mit Hero im Big Blind und $8\spadesuit 10\spadesuit$
 - Button raised
 - Small Blind raised
- Eine mögliche Abbildung würde zwei Spielzustände erstellen mit Hero im Big Blind und raise des Small Blinds
 - es wird nicht beachtet, dass die Wahrscheinlichkeit gegen zwei Gegner den Pot zu gewinnen geringer ist
- Tabelle mit den prozentual Anteilen der Aktionen (fold, call, river) und deren Differenz zu einem starken Spieler (hyperborean) für jede Wettrunde und die Anzahl der aktiven Spieler

HBB2	Preflop	Flop	Turn	River
2	(0.076,0.229,0.694)	(0.065,0.487,0.447)	(0.086,0.508,0.404)	(0.086,0.514,0.398)
	(0.063, 0.282, -0.345)	(0.02, 0.024, -0.043)	(-0.001, 0.056, -0.055)	(-0.002, 0.042, -0.039)
3	(0.122,0.124,0.753)	(0.017,0.347,0.634)	(0.043,0.47,0.486)	(0.054,0.463,0.482)
	(0.423, -0.017, -0.406)	(0.133, 0.196, -0.329)	(0.113, 0.046, -0.158)	(0.136, 0.066, -0.201)

Bias



HBB2	A	K	Q	...
...
4	0.0 0.0 0.1	0.0 0.136 0.961	0.0 0.25 0.75	...
3	0.0 0.0 0.1	0.0 0.04 0.96	0.0 0.222 0.777	...
2	0.0 0.0 0.1	0.0 0.121 0.878	0.0 0.136 0.863	...

Hyp.	A	K	Q	...
...
4	0.075 0.29 0.634	0.971 0.028 0.0	0.965 0.034 0.0	...
3	0.295 0.118 0.585	0.977 0.022 0.0	0.967 0.032 0.0	...
2	0.967 0.042 0.0	0.957 0.042 0.0	0.974 0.025 0.0	...

- Matrix für die Starthände
 - Obere Hälfte gleichfarbige (suited) , untere Hälfte verschiedenfarbige (offsuited) Starthände
- $\text{Bias}(Q2o) = 0.974 - 0.0 = 0.974$
 - Q2o wird Preflop um 95,7% mehr gefoldet

- Ergebnisse einer Partie gegen zwei mathematisch faire Bots

-	Ohne Bias	Mit Bias
HBB2	-1698	-1245
HBB3	-1378	-716
HBB4	-1318	-737

Probleme bei den Abbildungen

- Beispiele für Abbildungen eines drei Spieler Spielzustandes
 - Hero ist der BU mit A♥K♦, Hero raises, SB call, BB calls
 - Flop [7♠8♠9♠] SB bet, Button raises, Hero ???
 - Pot: 90, Potodds: 90:20
 - HBB2 99% call

Hero Position	Preflop	Flop	Pot	Potodds	CFR-Bot
SB	r,c	r	50	50:10	99% call
SB	r,c	r,r,r	90	90:20	99% call
BB	c,r,c	c,r	50	50:10	99% call

- Durch kein zwei Spieler Zustand wird die Foldwahrscheinlichkeit erhöht

Probleme bei den Abbildungen

- Hero ist der BU mit $K\spadesuit 9\clubsuit$, Hero calls, SB call, BB calls
- Flop [$J\clubsuit 3\diamondsuit 6\diamondsuit$] Hero calls, SB calls, BB calls
- Turn [$A\clubsuit$] SB bets, BB raises, Hero ??
- Pot: 90, Potodds: 90:40
- HBB2 74% call 26% fold

Hero Position	Preflop	Flop	Turn	Pot	Potodds	CFR-Bot
SB	cc	cc	r	40	40:20	73% call 26% fold
SB	cc	rc	r	60	60:20	13% call 86% fold

- Die Foldwahrscheinlichkeit wurde erhöht mit einem Zustand der von keiner unserer Abbildungen erstellt werden würde

Fazit

- Keine Abbildungen konnte das gewünschte Ergebnis erzielen
- Einige weitere Ideen wurden getestet, die nicht im Vortrag enthalten sind
- Viele Weiter Abbildungen sind denkbar
- Es ist fraglich, ob ein zwei Spieler Bot in einem drei Spieler Spiel durch Abbildungen sinnvoll verwendet werden kann

- [1] Annual Computer Poker Competition. Website. <http://www.computerpokercompetition.org/>, Zugriff am 27.2.2013.
- [2] Richard Gibson, Marc Lanctot, Neil Burch, Duane Szafron, and Michael Bowling. Generalized sampling and variance in counterfactual regret minimization. AAI-12, 2012.
- [3] Richard Gibson and Duane Szafron. On strategy stitching in large extensive form multiplayer games. NIPS-11, 2011.
- [4] Michael Johanson, Nolan Board, Marc Lanctot, Richard Gibson, and Michael Bowling. Efficient nash equilibrium approximation through monte carlo counterfactual minimization. AAMAS-12, 2012.
- [5] Michael Bradley Johanson. Robust strategies and counter-strategies: Building a champion level computer poker player. Master's thesis, University of Alberta, 2007.
- [6] Michael Johanson, Nolan Bard, Neil Burch, and Michael Bowling. Finding optimal abstract strategies in extensive-form games. AAI-12, 2012.
- [7] Nicholas Abou Risk. Using counterfactual regret minimization to create a competitive multiplayer poker agent. Master's thesis, University of Alberta, 2009.
- [8] Nick Abou Risk and Duane Szafron. Using counterfactual regret minimization to create competitive multiplayer poker agents. AAMAS-10, 2010.
- [9] Duane Szafron, Richard Gibson, and Nathan Sturtevant. A parameterized family of equilibrium profiles for three-player kuhn poker. AAMAS-13, 2013.
- [10] Wikipedia. Website. http://de.wikipedia.org/wiki/Texas_Hold'em, Zugriff am 27.3.2013.
- [11] Martin Zinkevich and Michael Johanson. Regret minimization in games with incomplete information. NIPS-07, 2007.
- [12] <https://upload.wikimedia.org/wikipedia/commons/thumb/d/d2/PokerTisch.svg/500px-PokerTisch.svg.png>

Fragen



TECHNISCHE
UNIVERSITÄT
DARMSTADT