

TUD Computer Poker Challenge

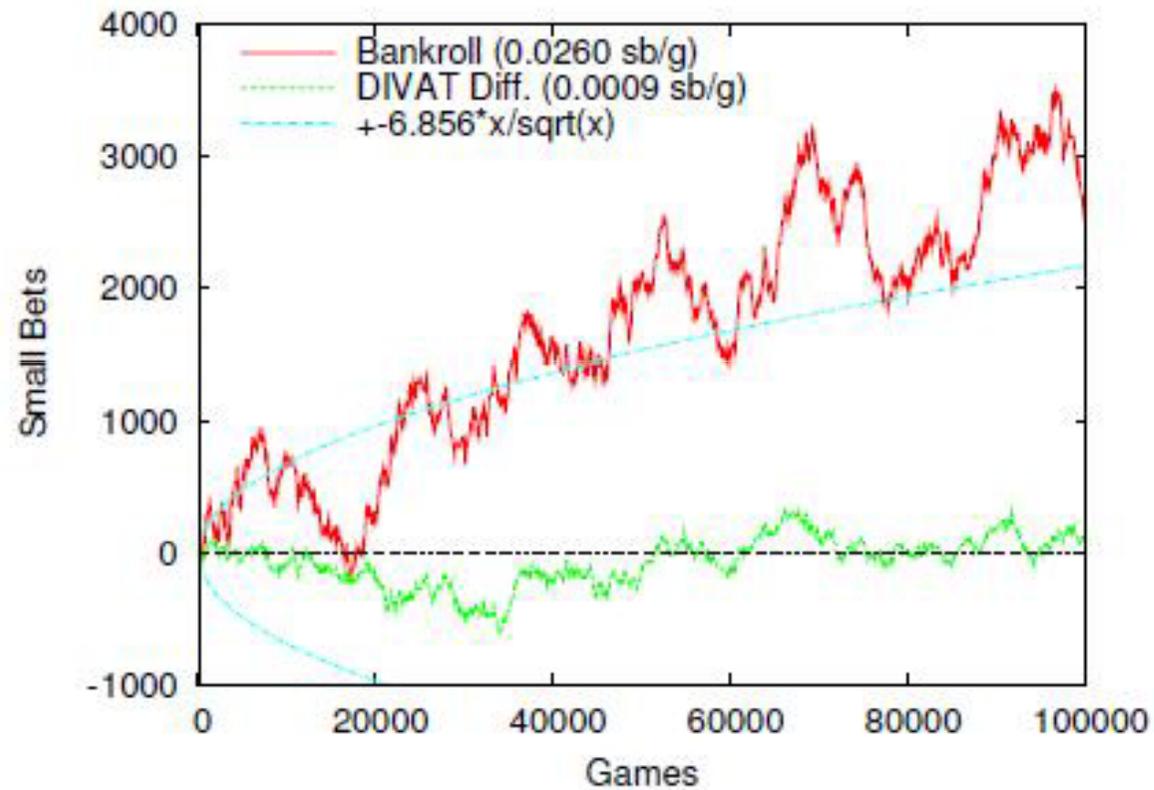


TECHNISCHE
UNIVERSITÄT
DARMSTADT

Strategy Evaluation -
oder: „Wie stark spielt eigentlich mein Pokerbot?“

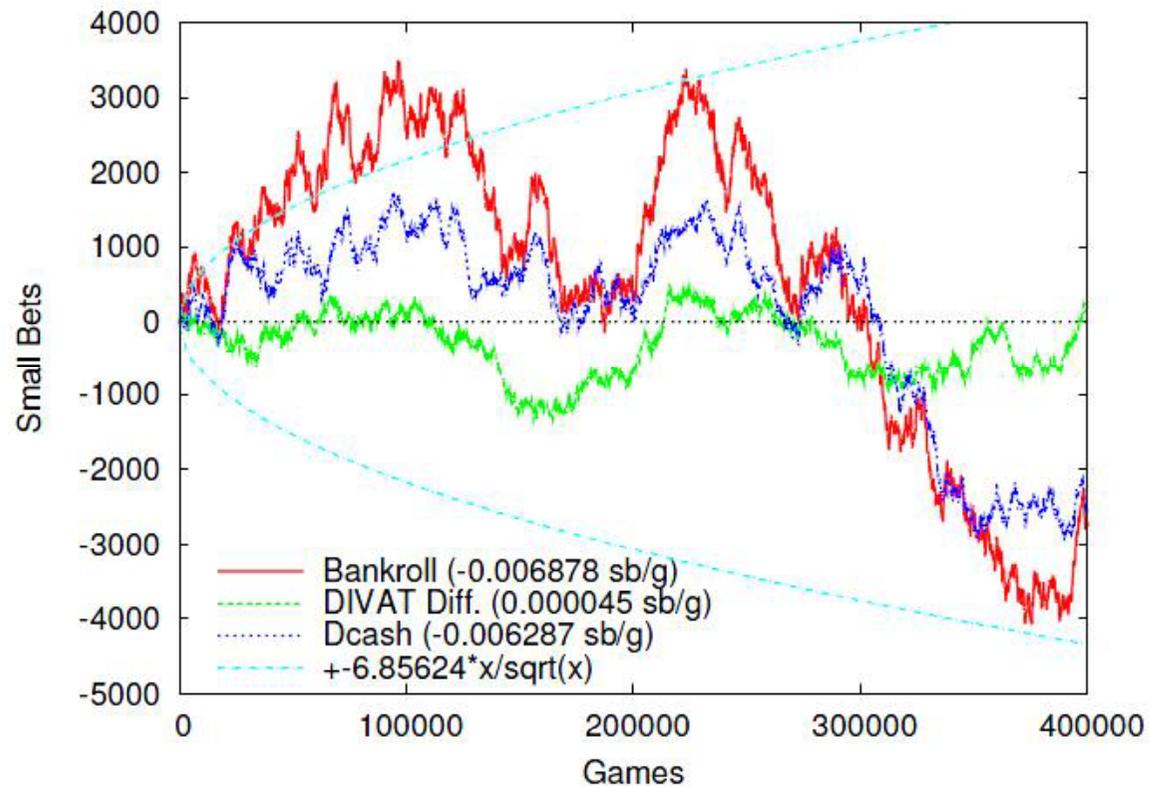


Motivation



Winningplayer ???

Motivation



- nicht so einfach zu beantworten

Inhalt



-
- Motivation
 - Inhalt
 - Einleitung
 - Beispielhand
 - Metrik
 - EVAT und LFAT
 - IHR, 7cHR und EHR
 - AIE und ROE
 - DIVAT
 - Analyse der Beispielhand
 - Beispielgraphen
 - Fazit
 - Quellen

Einleitung

- Poker unterscheidet sich von bisherigen Forschungsspielen wie Schach und Go
 - beides 2-Spieler Spiele mit perfekter Information und keinen Zufallsereignissen
 - bei Poker: Multiplayer Spiel mit unvollständigen Informationen und starke Beeinflussung durch Zufallsereignisse
 - durch Zufallsereignisse und hohe Varianz:
 - Spieler kann gut spielen, aber dennoch über viele tausend Hände verlieren
- schwer zu entscheiden wie gut ein Spieler ist oder ob ein Spieler besser ist als ein Anderer
- wird immer schwerer zu beurteilen, da die Programme immer stärker werden und sich die Spielstärke immer mehr angleicht

Einleitung

Vergleich mit Spielen mit perfekter Information wie Backgammon oder Black Jack:

- jede Entscheidung hat einen objektiven EV (expected value - Erwartungswert)
- es existiert meistens genau ein Zug mit maximalem EV
- dieser „best practical move“ lässt sich bestimmen durch
 - Enumeration über alle erreichbaren Spielzustände (perfekte Vorhersage)
 - Monte Carlo Simulation (Abschätzung)
- vergleich einer Spielerentscheidung mit diesem besten möglichen Zug ermöglicht Spielerbewertung

Einleitung



Probleme bei Poker

- Pokerprogramme sind noch nicht in der Lage verwertbare Abschätzungen zu liefern
 - es existiert meist keine (Einzel-)Entscheidung mit maximalem EV
 - ausbalancierte Spielweise nötig
 - spielt man nur starke Hände wird man leicht lesbar
→ bluffing und trapping nötig
 - Strategie stark von der Gegnereinschätzung abhängig
 - falls Alice zu oft blufft sollte Bob öfter callen und raisen mit mittelstarken Händen
- man muss „den Gegner spielen“, nicht nur das Board (wie z.B. beim Schach)

Einleitung



-
- Gegner kann nicht objektiv eingeschätzt werden
→ nicht möglich einen objektiv besten Zug zu bestimmen

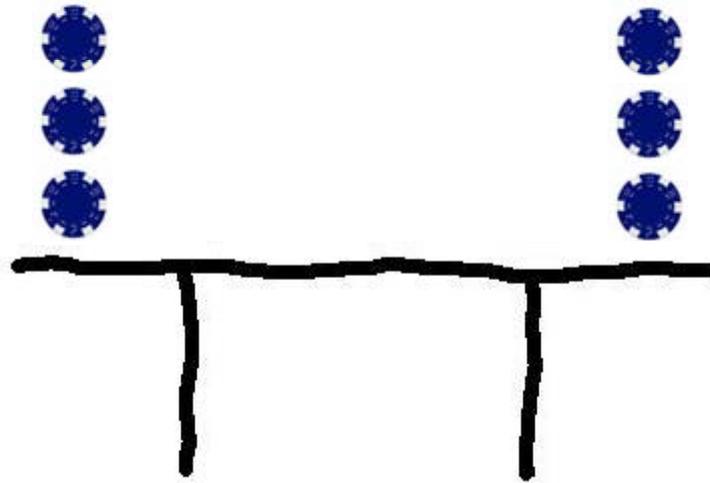
Ein Bewertungssystem sollte deshalb

- einfach
- konsistent
- stark varianzreduzierend und
- erwartungstreu sein

Beispielhand



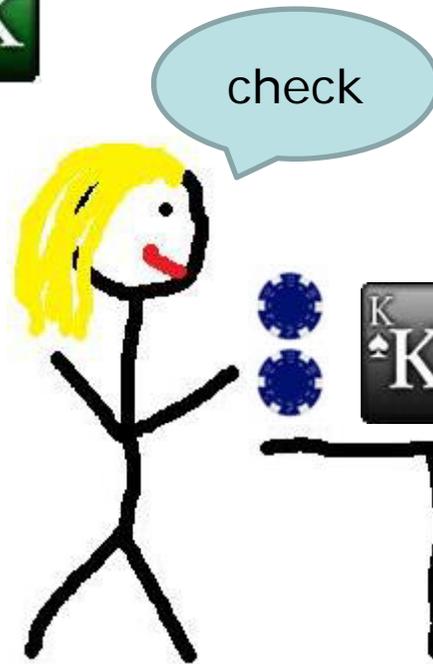
Alice



Bob

Beispielhand

Potsize: 6 SB



Alice



Bob

Beispielhand

Potsize: 10 SB



Alice



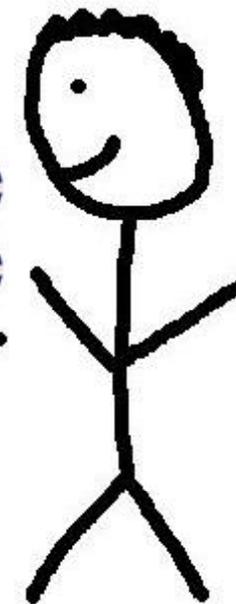
Bob

Beispielhand

Potsize: 14 SB



Alice



Bob

Beispielhand



- Spielzusammenfassung:
- Alice: AcKc Bob: 7h6h Board: Ks5h3dTc4h
- Betting: SIRrC / kBrC / bC / bRc
- s – small blind, l – large blind, k – check, b – bet, f – fold, r – raise
- Groß-/Kleinschreibung ermöglicht leichte Unterscheidung zwischen den Spielern (Alice klein, Bob groß)
- Final Pot: 22SBs
- Alice verliert 11SBs und Bob gewinnt 11SBs

Beispielhand

-
- Wer war nun der bessere Spieler?
 - Alice hat gut gespielt, hatte aber auch viele leichte Entscheidungen
 - Schwer zu beurteilen ob Sie immer gut spielt
 - Beurteilung von Bobs Spiel noch schwerer
 - stark von der Einschätzung von Bob abhängig
 - welche Überlegungen hat Bob über Alice angestellt?
 - History der beiden wichtig

EVAT und LFAT

- EVAT (Expected Value Assessment Tool) beruht auf dem „Fundamental Theorem of Poker“ (David Sklansky 1992):

“Every time you play a hand differently from the way you would have played it if you could see all your opponents' cards, they gain; and every time you play your hand the same way you would have played it if you could see all their cards, they lose. Conversely, every time opponents play their hands differently from the way they would have if they could see all your cards, you gain; and every time they play their hands the same way they would have played if they could see all your cards, you lose.”

- Sklansky sagt, dass der EV eines Spielers gleich der Differenz zu der optimalen Strategie laut FToP ist
- EVAT vergleicht die Spielerentscheidung mit der Entscheidung die getroffen werden würde, wenn alle Karten bekannt wären

EVAT und LFAT



- Problem: Unrealistische Beurteilung, da EVAT vollständiges Wissen voraussetzt
- Alice müsste laut EVAT in der Beispielhand auf dem River check/fold spielen
- Spieler mit 2. bester Hand bekommt eine schlechte Bewertung wenn er gegen die beste Hand spielt und eine gute wenn er gegen die drittbeste Hand spielt
- von Alice wird aber erwartet, dass sie auf dem River weiteres Geld verliert

EVAT und LFAT



-
- LFAT: Luck Filtering Analysis Tool
 - vergleicht die pot equity vor und nach Zufallsereignissen
 - bewertet wie viel „Glück“ die Spieler mit einem Zufallsereignissen hatten

 - EVAT bewertet also zwischen den Setzrunden und LFAT bewertet zwischen Zufallsereignissen

IHR – Immediate Hand Rank

- IHR bewertet die aktuelle Handstärke
- Gegen wie viele Hände sind wir vorne, hinten oder gleichauf

- $$\text{IHR} = \frac{\textit{ahead} + \frac{\textit{tied}}{2}}{\textit{ahead} + \textit{tied} + \textit{behind}}$$

- Beispiel: 
$$\text{IHR} = \frac{1147 + \frac{3}{2}}{1225} = 0,9376$$

7cHR– 7-card Hand Rank



- komplette Enumeration über alle möglichen zukünftigen Community Cards
 - berechnet auf allen Straßen den IHR
 - kombiniert negatives Potential (Chance zu verlieren wenn man aktuell vorne liegt) und positives Potential (Chance zu gewinnen wenn man aktuell hinten liegt)
 - schlechte Kombination beider Werte, da davon ausgegangen wird, dass alle Hände bis zum Showdown gespielt werden
 - überbewerten schwache non-pair-Hands, die evt. schon am Flop zum folden gebracht werden, da beim 7cHR eine 2 Karten Zukunft angenommen wird
 - 6cHR eine Möglichkeit das auszugleichen, ist aber oft zu defensive, da z.B. ein ankommender flushdraw eine 2 Karten Zukunft realistisch macht
-

EHR – Effective Hand Rank



- $EHR = \max(IHR, 7cHR)$
- zu beachten: alle möglichen Gegnerhände werden ungewichtet in die Berechnung mit einbezogen
- d.h. eine Hand wie AKs wird als genauso wahrscheinlich angenommen als 72o

AIE – All In Equity

- misst den Anteil am Pot, den ein Spieler im Schnitt gewinnen wird, falls alle „All In“ sind („fair share“)

- Beispiel:  gegen 

- beide Spieler setzen preflop alles

→  hat eine AIE von 60.363%

AIE – All In Equity

- Problem: unrealistische Messung der equity, da zukünftige bettings nicht mit einberechnet werden (implied odds)
 - Beispiel:
 - Alice: 3h2h
 - Bob: AsTs
 - Board: KsTh7c Ah
 - Alice hat eine AIE von ca. 20% (9/44)
 - bei einem Herz auf dem River kann sie aber davon ausgehen zusätzliche SBs von Bob zu gewinnen
 - kommt kein Herz, verliert sie keine weiteren SBs
- true equity ist höher als die AIE

ROE – Roll Out Equity



- bezieht zukünftige Setzrunden mit ein
- abhängig von einer Standardstrategie
- einfache Abstraktion: eine small bet / Runde
- besser: sehr schlechte Hände folden, starke Hände raisen
- noch bessere Strategie wird später vorgestellt

Net / Pot Equity



- $\text{Net} = (\text{AIE} * \text{potsize}) - \text{invested}$
- durchschnittlicher Gewinn bzw. Verlust an einem Pot
- Beispiel:
 - Pot auf dem Turn ist 10 small bets groß (5 von jedem Spieler)
 - 44 mögliche Karten können auf dem River noch kommen (52 Karten im Deck – 4 Karten der beiden Spieler – 4 bekannte Boardkarten)
 - $\text{Net} = (0,25 * 10) - 5 = -2,5 \text{ SB}$
 - wir werden also nach aktuellem Stand (Turn) in der Hand im Schnitt 2,5 SB verlieren

DIVAT



- Analyse einer Hand mit allen Informationen
- Spielrunden werden getrennt voneinander betrachtet
- Geschichte der Spieler wird ignoriert
- ebenso die Umstände wie eine Situation zustande gekommen ist
- Problem: Aktionen wie z.B. ein freecard raise können nicht berücksichtigt werden
- DIVAT System basiert auf *baseline betting sequence*
- diese wird bestimmt durch *folding policies* und *betting policies*
- *baseline betting sequence* repräsentiert eine „vernünftige“ bettingline ohne deceptive plays wie bluffing oder trapping
- gemessen wird die Differenz zwischen der baseline equity und der erreichten Equity eines Spielers

DIVAT

- Beispiel:
- Alice versucht gegen Bob einen checkraise auf dem Flop
- Falls der checkraise gelingt werden durch das trapping mehr SBs in den Pot gebracht als es durch die baseline (bet/call) berechnet wurde
- Alice hat also mehr Gewinn gemacht als erwartet
- daraus resultiert eine bessere Bewertung von Alice
- schlägt der checkraise Versuch fehl, werden weniger SBs als erwartet im Pot sein und Alice bekommt eine schlechtere Bewertung
- Bewertung wird aufgrund der gewonnenen bzw. verlorenen Equity gemacht

DIVAT – folding policy

- entscheidet, ob ein Spieler auf eine Bet des Gegners folden sollte
- basiert auf der optimalen spieltheoretischen fold frequency
- nur abhängig von der potsize und der zu bringenden Bet
- optimale fold frequency = $\frac{\text{betsize}}{\text{potsize} + \text{betsize}}$
- Beispiel: angenommen wir sind auf dem River bei einer potsize von 8 SB. Die betsize beträgt 2 SB
- wir berechnen eine fold frequency von $\frac{2}{8 + 2} = 0,2$
- wir sollten also die schlechtesten 20% aller Hände folden

DIVAT – folding policy

- da der 7cHR einen Wert von 0 bis 1 annimmt und halbwegs gleichmäßig verteilt ist eignet er sich um zu bestimmen, ob wir in der foldingrange sind
- allerdings ist die Berechnung für die pre-River Straßen nicht ausreichend
- Foldschwelle sollte etwas höher liegen, da der 7cHR Hände auf den frühen Straßen teilweise überbewertet
 - z.B. haben kleine suited connectors einen relativ hohen 7cHR, müssen aber oft gefoldet werden
- deshalb wird Schwellenwert auf $\frac{\text{bet size}}{\text{pot size} + \text{bet size}} + \text{offset}$ gesetzt
- offset empirisch bestimmt, abhängig von der Straße

DIVAT – betting policy



-
- entscheidet wie viele bets und raises eine Hand wert ist
 - nur „bets for value“, keine bluffbets o.ä.
 - es entsteht direkt eine positiver EV
 - die Setzhöhe in direkte Abhängigkeit zu der Handstärke zu stellen eigentlich keine gute Idee
 - man wird durch die Gegner leicht lesbar
 - liefert aber eine gute Abschätzung wie viel ein Spieler in einer gegebenen Situation im Schnitt investieren sollte
 - wird gemessen anhand des EHR

DIVAT – Standard Einstellungen



	Offset	Make1	Make2	Make3	Make4
Preflop	0.000	0.580	0.825	0.930	0.965
Flop	0.075	0.580	0.825	0.930	0.965
Turn	0.100	0.580	0.825	0.930	0.965
River	0.000	0.640	0.850	0.940	0.970

DIVAT – Standard Einstellungen



-
- die Make2 Schwelle sagt z.B. aus, dass es eine Hand wert ist 2 SBs (freiwillig) zu investieren
 - erhöhte Riverwerte berücksichtigen Tatsache, dass der Gegner nur mit Händen, die wenigstens etwas Showdown value haben callen wird
 - Gegner wird nicht „mit Nichts“ callen
 - etwas moderater betten auf dem River
 - Beispiel:
 - EHR (Alice) = 0,70 EHR (Bob) = 0,90
 - Alice first to act: betting policy wird bet-raise-call annehmen
 - Bob first to act: betting policy wird bet-call annehmen

Analyse der Beispielhand - Preflop



Board: < >			
	IHR	7cHR	EHR
Alice <AcKc>	0,9376	0,6704	0,9376
Bob <7h6h>	0,1604	0,4537	0,4537
AIE = 0,6036			
LFAT change = +0,2073			
DIVAT baseline = SICrC		Aktuelle Seq. = SIRrC	
Round EV = +0,2073 [SICrC]			
Actual Equity = +0,6218 [SIRrC]			
Baseline Equity = +0,4145 [SICrC]			
DIVAT Difference = +0,2073 SB			

Analyse der Beispielhand - Preflop



- Bob raiset preflop
 - wahrscheinlich nicht mit der Absicht Alice direkt zum folden zu bringen
 - raise kann später aber von Alice missinterpretiert werden
 - daraus kann mehr value entstehen
 - raised Bob nie mit schwächeren Händen wird er leicht zu lesen
- Hätte Alice nur gecallt wäre die Actual Equity gleich der Baseline Equity, weil die gleiche Menge an SBs im Pot gewesen wären
- Alice raised aber und bekommt so eine positive Bewertung auf kosten von Bobs Wertung

Analyse der Beispielhand - Preflop



- die Höhe der Bewertung ist gleich der DIVAT Difference
- 2 extra bets sind in den Pot gewandert
- Alice wird im Schnitt ca. 60% (laut AIE) der Hände gewinnen
- Alice „gehören“ also $0,60 * 2SB$ des Pots, 1 SB hat sie selbst investiert
- der Gewinn preflop beträgt also $0,60 * 2 - 1 = +0,20$ SB

- der Einsatz von Alice (2 SBs) hat einen EV von +0,2073 (Round EV), da aber Bob nochmal die gleiche Menge SBs investiert wird von Alice erwartet, dass sie insgesamt einen Gewinn von +0,4145 macht
- sie macht aber einen Gewinn von +0,6218 und erhält daher einen positive Bewertung von +0,2073

Analyse der Beispielhand - Flop

Board: <Ks5h3d>			
	IHR	7cHR	EHR
Alice <AcKc>	0,9685	0,8687	0,9685
Bob <7h6h>	0,0634	0,3798	0,3798
AIE = 0,7636			
LFAT change = +0,9600			
DIVAT baseline = bC		Aktuelle Seq. = kBrC	
Round EV = +0,5273 [bC]			
Actual Equity = +2,6364 [kBrC]			
Baseline Equity = +2,1091 [bC]			
DIVAT Difference = +0,5273 SB			

Analyse der Beispielhand - Flop



- guter Flop für Alice
- Chance auf Sieg steigt auf 76% (AIE)
- wird auch von LFAT-Wert angezeigt
- LFAT = +0,9600 bedeutet, dass das Zufallsereignis einen positiven Effekt im Wert von knapp einem SB für Alice gebracht hat
- wird berechnet durch Pot Equity am Flop - Pot Equity preflop:
- $6 * 0,7636 - 6 * 0,6036 = 0,9600$

Analyse der Beispielhand - Flop



- DIVAT geht von bC aus, baseline equity demzufolge +2,1091 SB
- Alice gelingt ein checkraise und erreicht dadurch +2,6364 SB
- Differenz von +0,5273 wird wieder Alice gutgeschrieben
- Hätte Bob nur gecheckt (kK) wäre die equity von Alice nur +1,5816 gewesen
- daraus wäre eine negative Bewertung für Alice im Wert von -0,5273 entstanden

Analyse der Beispielhand - Turn

Board: <Ks5h3d Tc>			
	IHR	7cHR	EHR
Alice <AcKc>	0,9411	0,8902	0,9411
Bob <7h6h>	0,0662	0,2146	0,2146
AIE = 0,9091			
LFAT change = +1,4545			
DIVAT baseline = bF		Aktuelle Seq. = bC	
Round EV = 0 [bF]			
Fold Equity = +0,9091 [bF]			
Actual Equity = +5,7273 [bC]			
Baseline Equity = +5,0000 [bF]			
DIVAT Difference = +0,7273 SB			

Analyse der Beispielhand - Turn



- Turn bringt noch eine gute Karte für Alice im Wert von +1,4545 (LFAT)
- erhöht die Chance zu gewinnen auf ca. 91%
- nach DIVAT baseline hat die Runde keinen positiven EV für Alice, da Bob folden wird
- Bob steht vor schwerer Entscheidung
- Wir wissen: callen kostet ihn 1,64 SB (da ihm ca. 9% der investierten SBs zustehen), folden kostet ihn aber nur 0,91 SB, da er auf die 9% von den 10SB im Pot verzichtet (fold equity)
- durch implied odds auf dem River lohnt sich aber diese Investition
 - falls auf dem River die 4 kommt, wird Bob von Alice sicher noch zusätzliche SBs gewinnen können

Analyse der Beispielhand - River



Board: <Ks5h3d Tc 4h>			
	IHR	7cHR	EHR
Alice <AcKc>	0,8576	0,8576	0,8576
Bob <7h6h>	0,9955	0,9955	0,9955
AIE = 0,0000			
LFAT change = -12,7273			
DIVAT baseline = bRc		Aktuelle Seq. = bRc	
Round EV = -4,0000 [bRc]			
Actual Equity = -11,0000 [bRc]			
Baseline Equity = -11,0000 [bRc]			
DIVAT Difference = 0 SB			

Analyse der Beispielhand - River



- River bringt perfekte Karte für Bob
- LFAT zeigt an, dass die 4h Alice 12,7273 SBs kostet
- DIVAT erwartet nun, dass Alice hier 4 SBs verlieren wird
- großer Unterschied zu EVAT wird deutlich
- EVAT würde hier eine baseline von check-bet-fold erwarten, bet-raise-call ist aber wesentlich realistischer
- EVAT würde also Alice für bRc schlecht bewerten, DIVAT gibt eine neutrale Bewertung ab
- würde Alice den Braten riechen und check-bet-call spielen würde sie von DIVAT das positiv bewerten, da ein größerer Verlust angenommen wurde

Analyse der Beispielhand – AIE vs. ROE



- Alice: AcKc
- Bob: 7h6h
- Board: Ks5h3dTc4h
- Betting: SIRrC / kBrC / bC / bRc

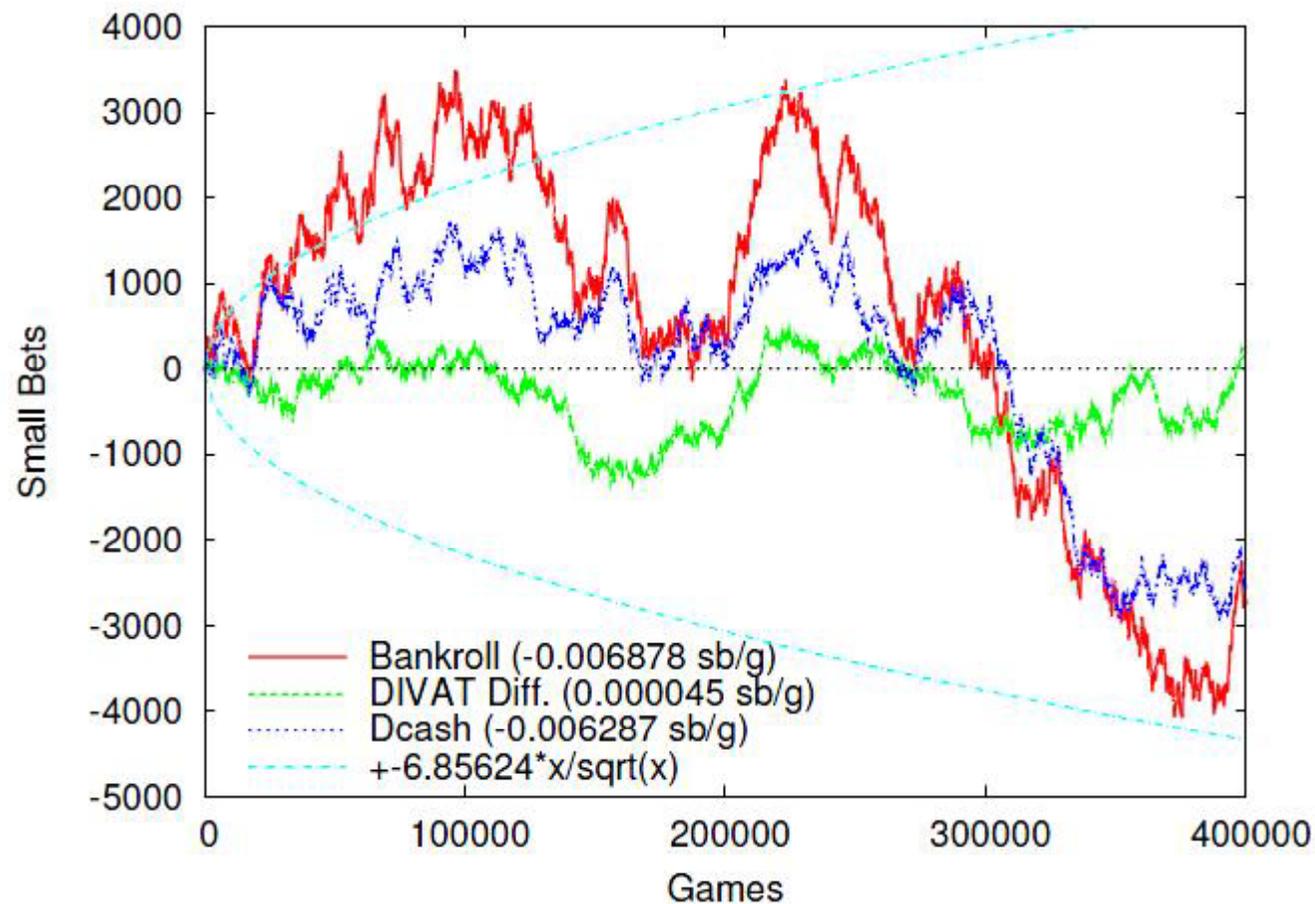
Round	Round EV (AIE)	DIVAT diff
preflop	+0,207	+0,207
flop	+0,527	+0,527
turn	+0,909	+0,727
river	-4,000	0
Total DIVAT difference = +1,462 SB		

Round	Round EV (ROE)	DIVAT diff
preflop	+0,029	+0,029
flop	+0,523	+0,523
turn	+0,909	+0,636
river	-4,000	0
Total DIVAT difference = +1,189 SB		

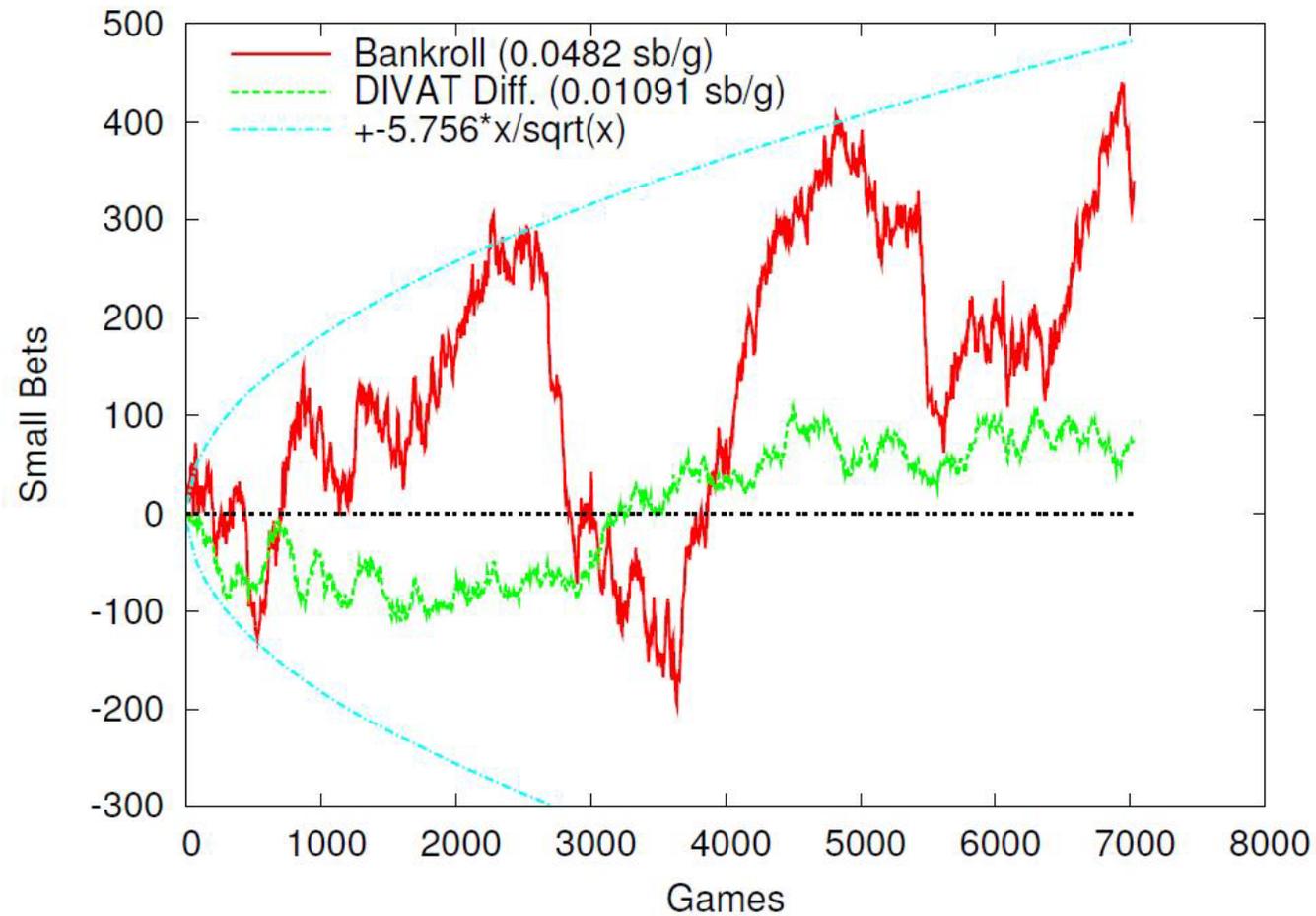
Beispielgraphen – ALWAYS_CALL vs. ALWAYS_RAISE



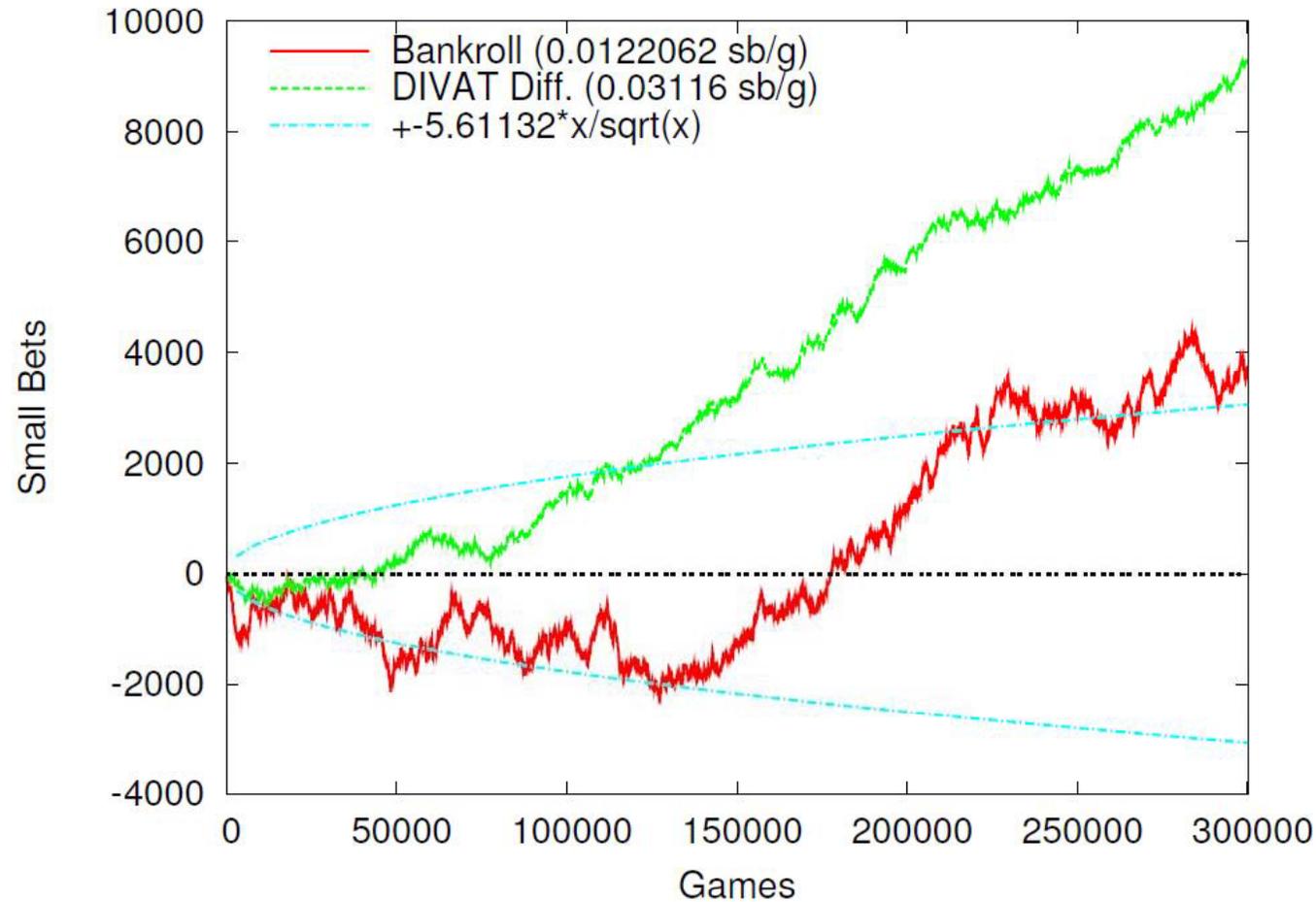
TECHNISCHE
UNIVERSITÄT
DARMSTADT



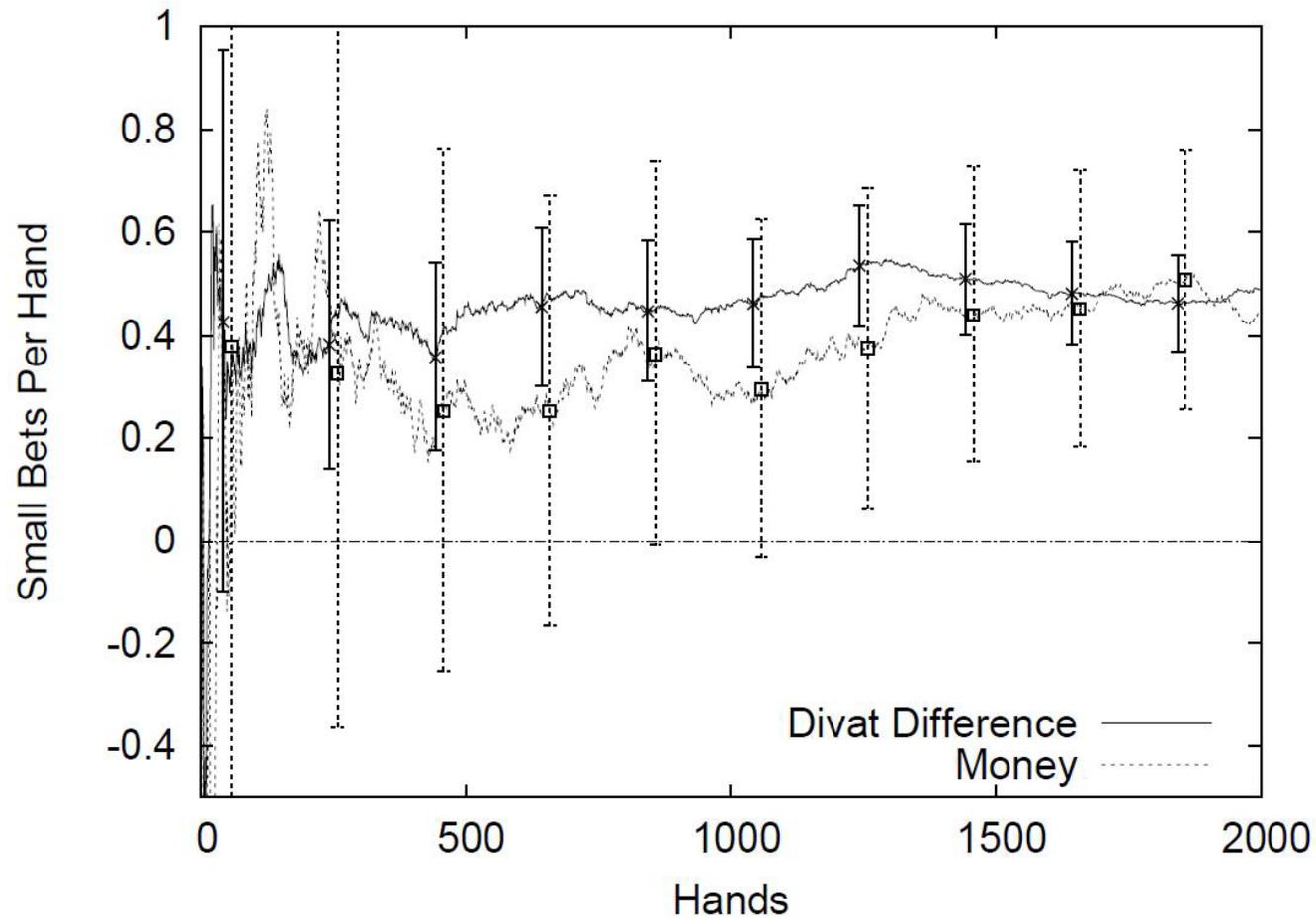
Beispielgraphen – thecount vs. PsOpti1



Beispielgraphen – VEXBOT vs. PsOpti4



Beispielgraphen – Mensch vs. Bot



Fazit



- Messung der Spielstärke anhand des bankroll-Verlaufs kaum machbar
- Bewertungsmethode muss enorm varianzreduzierend sein
- DIVAT stellt ein solches System zur Verfügung
- allerdings nur für heads up-Spiele
- Erweiterung auf 3max Spiele für die Challenge wäre nötig
- Analyse von Spielsituationen während einer Hand mit den gegebenen Metriken ebenfalls denkbar
- Ermittlung der baseline kann evt. auch noch verbessert werden, z.B. sind lines wie check-bet-raise-fold nicht vorgesehen

Danke für Eure Aufmerksamkeit



Quellen



-
- Darse Billings and Morgan Kan. A Tool for the Direct Assessment of Poker Decisions. The International Association of Computer Games Journal, October 2006.
 - Martin Zinkevich, Michael Bowling, Nolan Bard, Morgan Kan and Darse Billings. Optimal Unbiased Estimators for Evaluating Agent Performance. AAI, AAI Press, 2006.
 - Michael Bowling, Michael Johanson, Neil Burch and Duane Szafron. Strategy Evaluation in Extensive Games with Importance Sampling. Proceedings of the 25th Annual International Conference on Machine Learning (ICML), pp 72-70, 2008