

Die Reiz-Phase im Bridge



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Learning to bid in bridge


von Asaf Amit und Shaul Markovitch (2006)

Vortrag von Patrick Metzler

- Bieten in Bridge
 - Regeln
 - Probleme für Algorithmen
- Algorithmus von Amit und Markovitch: BIDI
 - Übersicht über die verwendeten Methoden
 - Entscheidungsfindung
 - zufällige Zustände generieren
 - Strategie-Funktion
 - Nutzen-Funktion
- Evaluation von BIDI

Reizung

- Ansage
 - Gebot
 - Kontra des letzten Gebots
 - Rekontra des letzten Kontras
 - Pass



The screenshot shows a Bridge game interface with the following details:

- Board 70** (IMP)
- North (NKA)**: J1085 J K 9 7 2 A J 7 2
- South (viyo)**: K Q 7 6 4 3 2 K 9 7 5 Q 10
- West (DiMPLE_)**: A A Q 6 3 8 5 4 Q 10 8 6 4
- East (GAS_)**: 9 10 8 4 2 A J 6 3 K 9 5 3
- Bidding Table**:

DIMPLE	0	NKA	GAS	viyo
1♥	Pass	Pass	Pass	1♥
2♣	Pass	Pass	Pass	3♣
3♣	Pass	Pass	Pass	4♥
4♥	Pass	Pass	Pass	Pass
- Score**: (NS) NS: -1.9 IMPs, (EW) EW: 1.9 IMPs
- Contract**: 4♥ South, NS 0 EW 0

Bridge Base Online

Ansagen



- Gebot
 - Anzahl der Stiche minus 6 und Trumpf oder ohne Trumpf (NT)
 - Kontra
 - Kann nur auf ein gegnerisches Gebot folgen.
 - Erhöht die Punkte der Gegner (Gewinn) bzw. die Eigenen Punkte (Verlust).
 - Rekontra
 - Kann nur auf ein gegnerisches Kontra folgen.
 - Pass
 - Das letzte Gebot ändert sich nicht.
- | | | | | | |
|------|----|----|----|-----|-----------|
| ▪ 1♣ | 1♦ | 1♥ | 1♠ | 1NT | 7 Stiche |
| ▪ 2♣ | 2♦ | 2♥ | 2♠ | 2NT | 8 Stiche |
| ▪ 3♣ | 3♦ | 3♥ | 3♠ | 3NT | 9 Stiche |
| ▪ 4♣ | 4♦ | 4♥ | 4♠ | 4NT | 10 Stiche |
| ▪ 5♣ | 5♦ | 5♥ | 5♠ | 5NT | 11 Stiche |
| ▪ 6♣ | 6♦ | 6♥ | 6♠ | 6NT | 12 Stiche |
| ▪ 7♣ | 7♦ | 7♥ | 7♠ | 7NT | 13 Stiche |

Wikipedia

Bietsysteme (Kommunikation über Ansagen)



Jede Ansage gibt dem Mitspieler Informationen über die eigene Hand:

- Figurenpunkte (HCP)
- Losing-Trick-Count
- Längen der Farben

Regeln (z.B. Eröffnungsgebote abhängig von Figurenpunkten):

13 - 14 HCP: biete 1 in einer Farbe, danach 1NT

18 - 19 HCP: biete 1 in einer Farbe, danach 2NT

22 - 24 HCP: biete 2♣, danach 2NT

28 - 30 HCP: biete 2♣, danach 3NT

31 - 32 HCP: biete 2♣, danach 4NT

Wikipedia

Wertung



Anzahl der gewonnenen Stiche entscheidet über Gewinn/Verlust.

Punkte setzen sich zusammen aus:

- Stichpunkte
- Prämien

Höhe ist abhängig von der Art des Kontrakts.

Prämien	nicht in Gefahr	in Gefahr
Gewinn eines Teilkontraktes	50	50
Gewinn eines Vollspiels	300	500
Gewinn eines Kleinschlemms	500	750
Gewinn eines Großschlemms	1000	1500
Gewinn eines beliebigen kontrierten Kontrakts	50	50
Gewinn eines beliebigen rekontrierten Kontrakts	100	100
für jeden Überstich in einem kontrierten Kontrakt	100	200
für jeden Überstich in einem rekontrierten Kontrakt	200	400

Wikipedia

Komplexität

- ca. 20 bis 30 mögliche Ansagen
(es gibt insgesamt 38 Ansagen)

$$\binom{52}{13} \approx 6.35 \cdot 10^{11} \text{ mögliche Hände}$$

$$\binom{39}{13} \cdot \binom{26}{13} \approx 8.45 \cdot 10^{16} \text{ Möglichkeiten, eine Hand zu vier Händen zu erweitern}$$

$$\binom{52}{13} \cdot \binom{39}{13} \cdot \binom{26}{13} \approx 5.36 \cdot 10^{28} \text{ mögliche Kartenverteilungen}$$

Wahl des nächsten Zuges

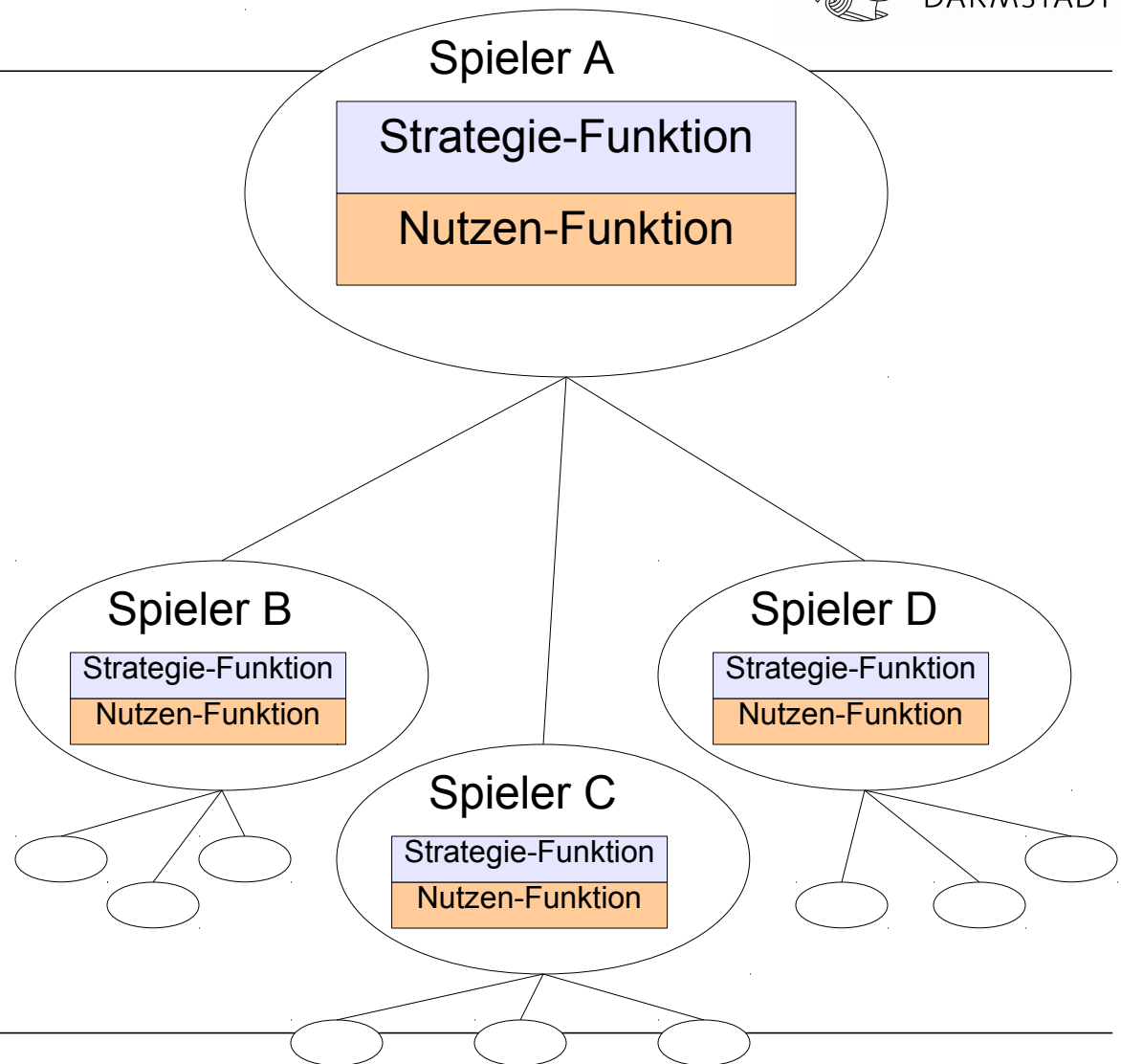
- Rückschlüsse auf die Karten der anderen Spieler nur durch Verlauf der Ansagen möglich
- Bietsysteme bieten eine gute Grundlage, um eine Entscheidung zu treffen.
- Aber: Bietsysteme sind mehrdeutig und unvollständig.
- Evaluation während der Reizung schwierig

Versuch mit einem regelbasierten
System in Kombination mit
vorausschauender Suche

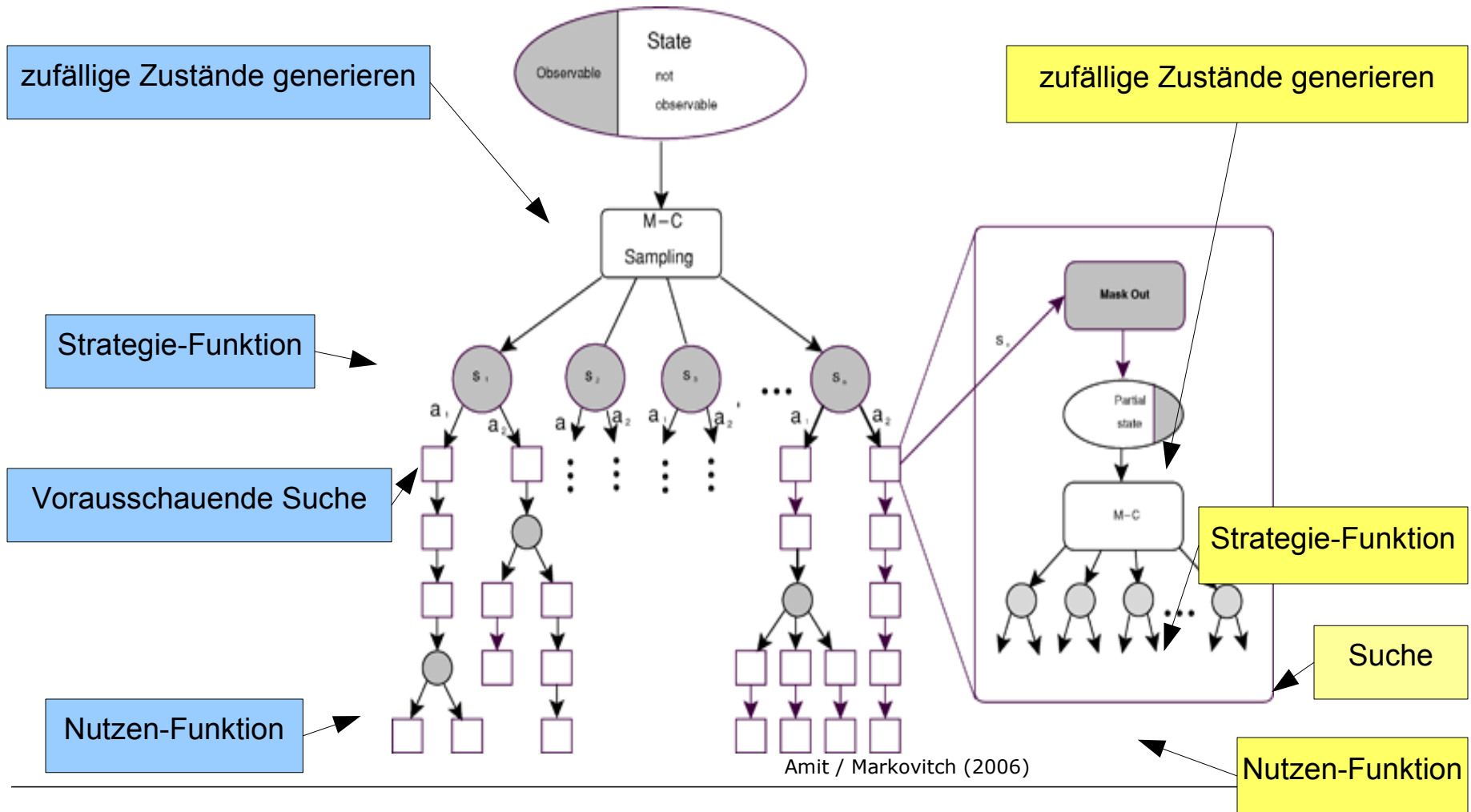
- Bieten in Bridge
 - Regeln
 - Probleme für Algorithmen
- Algorithmus von Amit und Markovitch: BIDI
 - Übersicht über die verwendeten Methoden
 - Entscheidungsfindung
 - zufällige Zustände generieren
 - Strategie-Funktion
 - Nutzen-Funktion
- Evaluation von BIDI

BIDI: Übersicht

- Unvollständige Information
 - Monte Carlo Sampling
- Gegner-Modellierung, Partner-Modellierung
 - Jeder Spieler besitzt Modelle der anderen Spieler
- limitierte Ressourcen aufgeteilt in:
 - Sampling
 - Suche bei eigenen Entscheidungen
 - Suche bei fremden Entscheidungen



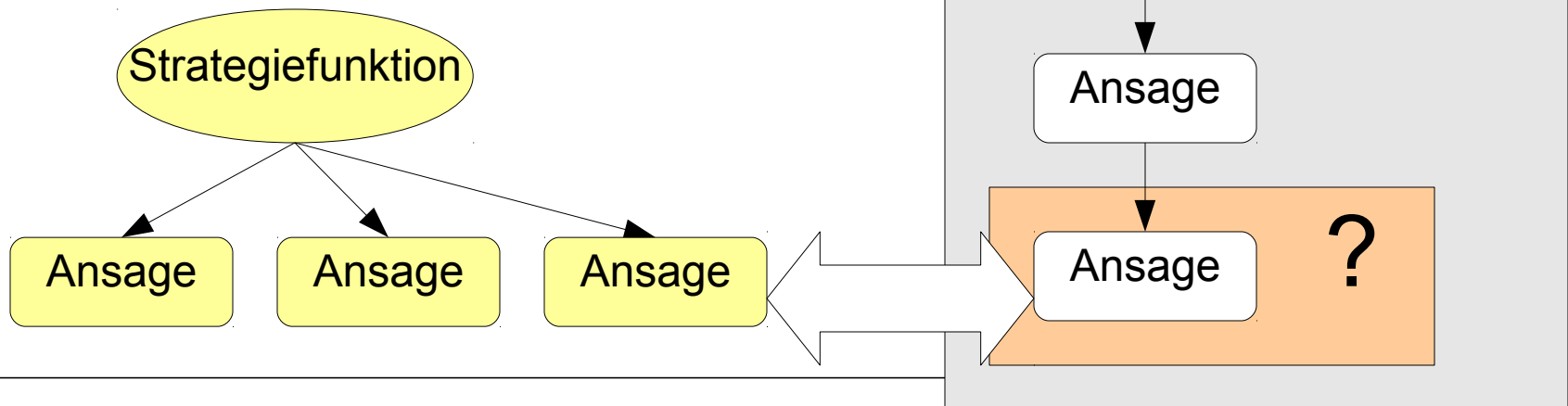
BIDI: Entscheidungsfindung



BIDI: zufällige Zustände generieren

- feste Anzahl zufälliger Zustände wird generiert
- so lange noch Ressourcen:
ersetze schlechtesten Zustand durch besseren

Bewertung der Zustände durch Strategie-Funktionen der Spieler



BIDI: Strategie-Funktion

- Darstellung als Entscheidungsnetzwerk
- Liefert zu einem Zustand eine Menge an sinnvollen Ansagen
- Klassifizierung der Zustände durch ca. 250 Eigenschaften

- Knoten:
 - Menge an Zuständen
 - Menge an Ansage
- Kanten:
 - Verfeinerung

BIDI: Lernen der Strategie- Funktion

- Ziel: jedem Zustand so wenig Ansagen wie möglich zuordnen
- Vorgehen: Offline-Lernen
 - Basiswissen: (regelbasiertes) Bietsystem → Entscheidungsnetzwerk
 - zufällige Zustände generieren
 - optimale Ansagen finden
 - Verfeinerung des Netzwerks

BIDI: Entscheidungsnetzwerk verfeinern

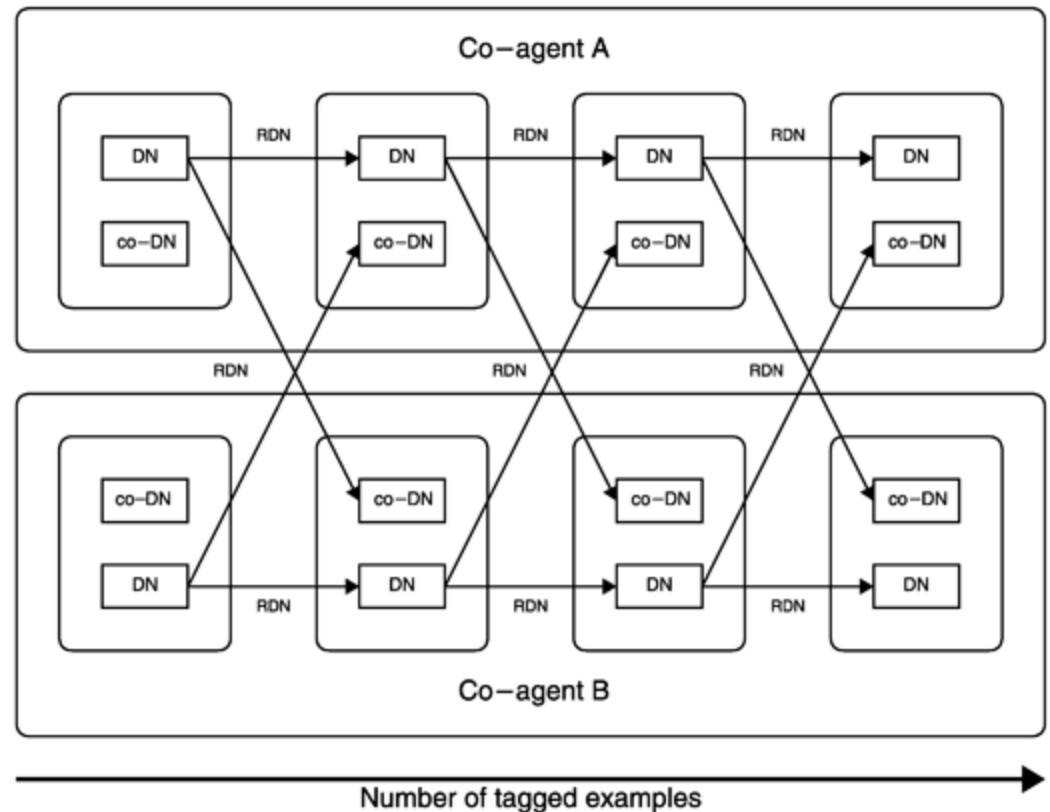


TECHNISCHE
UNIVERSITÄT
DARMSTADT

- zufällige Zustände generieren
 - Interessante Zustände priorisieren
- optimale Ansagen finden
 - Entscheidungsfindung durch aktuelle Strategie-Funktion und Suche (Speedup-Lernen)
- Netzwerk verfeinern
 - kleinere Zustands-Mengen
 - kleinere Ansagen-Mengen

BIDI: kooperatives Lernen

- jeder Partner hat Modelle des anderen
- paralleles Lernen
- Nach jedem Schritt: Austausch der Modelle



Amit / Markovitch (2006)

BIDI: Nutzen-Funktion



- gibt zu erwartende Punktezahl zurück
- Losing-Trick-Count-Methode:
 - Abschätzung der gewonnenen Stiche
 - Handkarten → Anzahl zu erwartender Stiche
- Vergleich mit dem Kontrakt
- nur sinnvoll nach abgeschlossener Reizung

$$18 - (7 + 0 + 2 + 1 + 3) = 5$$

The screenshot shows a Bridge hand with the following details:

- IMP Board:** Board 52
- Hand N (gunmir):** A♠, Q♣, Q♥, J♥, 9♥, 6♥, 3♠, A♠, Q♠, J♠, 10♦, 6♦, 2♦
- Hand W (aadd):** K♣, J♣, 10♣, 9♣, 5♣, 3♣, 8♥, 7♥, 5♥, 5♠, 3♠, A♦, 5♦
- Hand E (patcat):** 8♣, 2♣, A♥, 10♥, 4♥, 9♥, 8♥, 6♠, 4♠, J♦, 9♦, 8♦, 3♦
- Hand S (NKA):** 7♣, 6♣, 4♣, K♥, 2♥, K♠, 10♠, 7♠, 2♠, K♦, Q♦, 7♦, 4♦
- Bidding:** aadd (Pass), gunmir (1♥), patcat (Pass), 0 (Pass), NKA (1♠). Gunmir also bid 3♣ and 3♠.
- Contract:** 3♣x West
- Score:** NS 0, EW 0
- IMP Info:** (NS) NS: 23.9 IMPs, (EW) EW: -23.9 IMPs

Bridge Base Online

Übersicht

- Bieten in Bridge
 - Regeln
 - Probleme für Algorithmen
- Algorithmus von Amit und Markovitch: BIDI
 - Übersicht über die verwendeten Methoden
 - Entscheidungsfindung
 - zufällige Zustände generieren
 - Strategie-Funktion
 - Nutzen-Funktion
- Evaluation von BIDI

Evaluation: Verfahren



- IMP (International Match Points):
 - Vergleich von Gegnern
 - Referenz Punktzahl nötig
- bidding challenge
 - Reizung ohne Spielphase
 - ohne Gegner
 - Referenz-Punktzahl approximiert (LTC) oder aus der Literatur

Punktzahldifferenz → IMPs

von	bis		von	bis		von	bis	
0	10	0	370	420	9	1750	1990	18
20	40	1	430	490	10	2000	2240	19
50	80	2	500	590	11	2250	2490	20
90	120	3	600	740	12	2500	2990	21
130	160	4	750	890	13	3000	3490	22
170	210	5	900	1090	14	3500	3990	23
220	260	6	1100	1290	15	4000	or more	24
270	310	7	1300	1490	16			
320	360	8	1500	1740	17			

Wikipedia

Evaluation: Verfahren

durchgeführte Tests:

- Lernen
 - nach jedem Verfeinerungsschritt:
bidding challenge (automatisch)
- nach dem Lernen:
Spiele gegen GIB (manuell)
 - Team BIDI/BIDI gegen Team GIB/GIB
- Lernen mit Test der Parameter

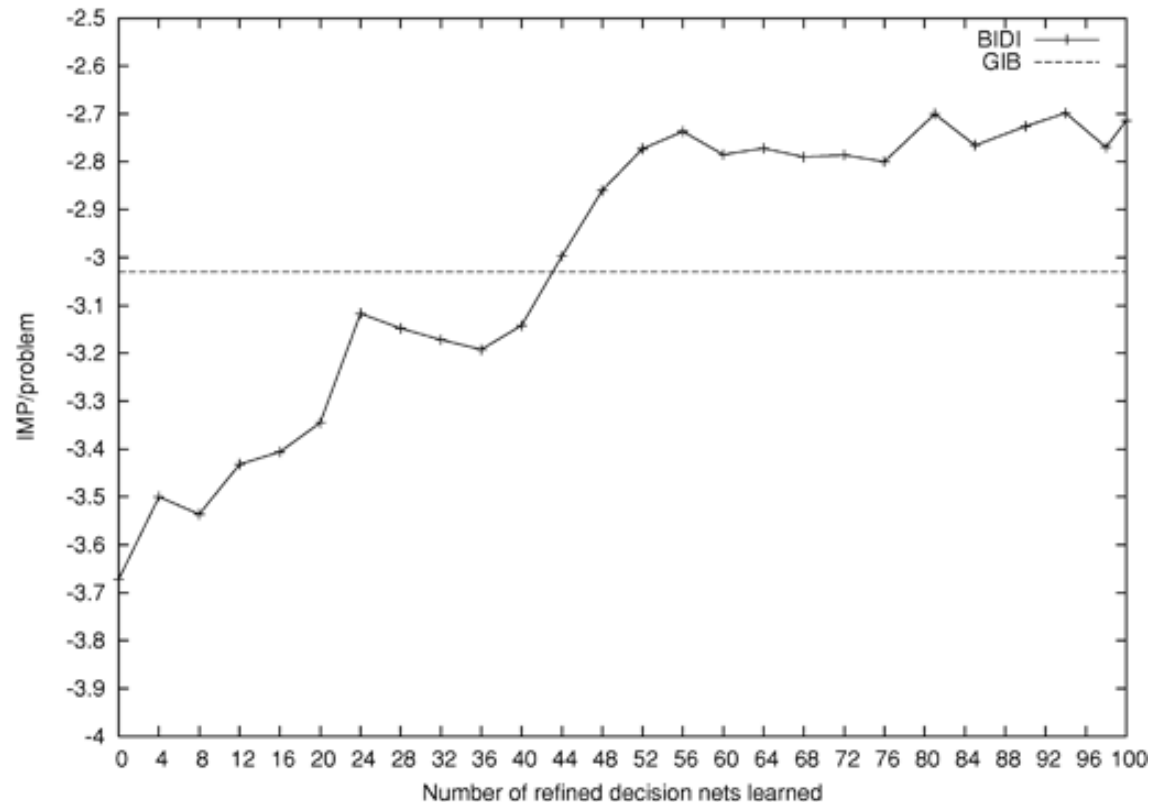
Evaluation: BIDI gegen GIB



- zwei Tische (mit den selben Karten):
 1. Tisch: BIDI N/S, GIB W/O
 2. Tisch: BIDI W/O, GIB N/S
- nach der Reizung:
 - automatisches Spiel durch GIB (für alle vier Spieler)
- 48 Runden
- Ergebnis:
 - BIDI: 103 IMP
 - GIB: 84 IMP
 - 0.4 IMP pro Runde

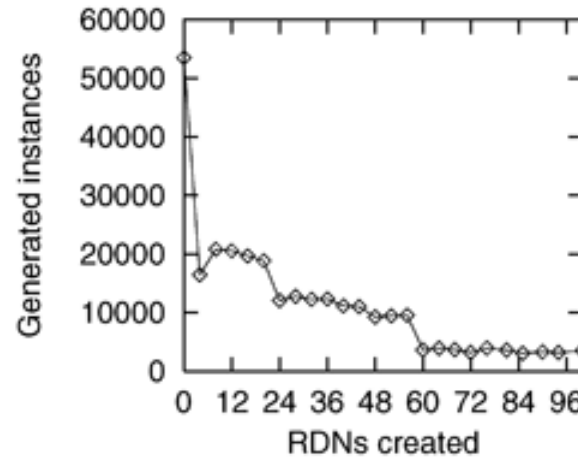
Evaluation: Lernen

- Lernen
 - Pool an 2000 zufälligen Zuständen
 - Priorisierung: Zustände, die viel Rechenzeit benötigen
 - Referenz-Ansage finden
 - 5 Sekunden
 - Sample-Größe 10
- bidding challenge nach jedem Schritt
 - 100 zufällige Kartenverteilungen
 - Kartenverteilungen aus der Literatur

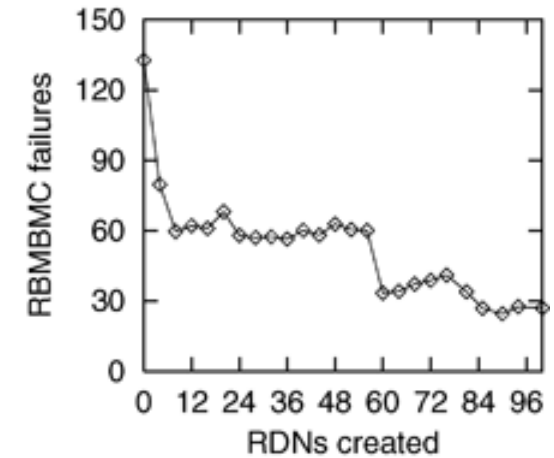


Amit / Markovitch (2006)

Evaluation: Lernen



(b)



(b)

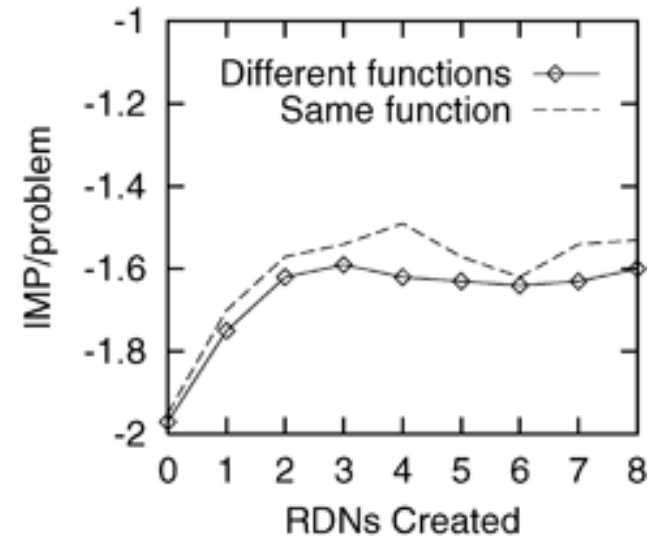
Amit / Markovitch (2006)

weniger Aufrufe der Suche, dadurch:

- weniger Zustände mit maximaler Fehlerzahl
- weniger Zustände generiert: Einsparung von Ressourcen

Evaluation: Nutzen-Funktion

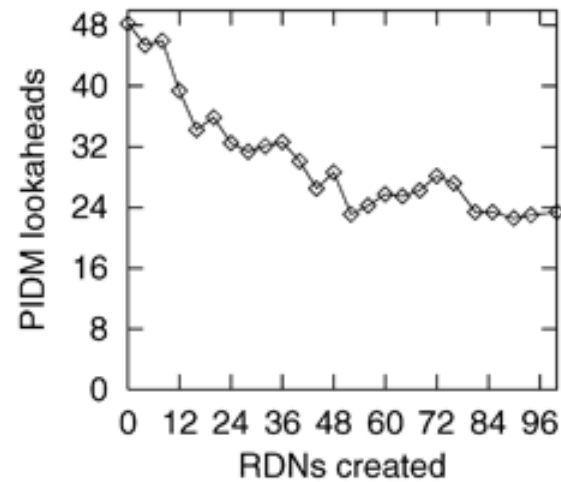
- gestrichelt:
 - beide Partner mit LTC
- durchgezogen:
 - eigene Nutzen-Funktion: LTC
 - Partner-Nutzen-Funktion: vereinfachtes LTC (mit Figurenpunkten)
- kein Vergleich mit simulierter Spielphase



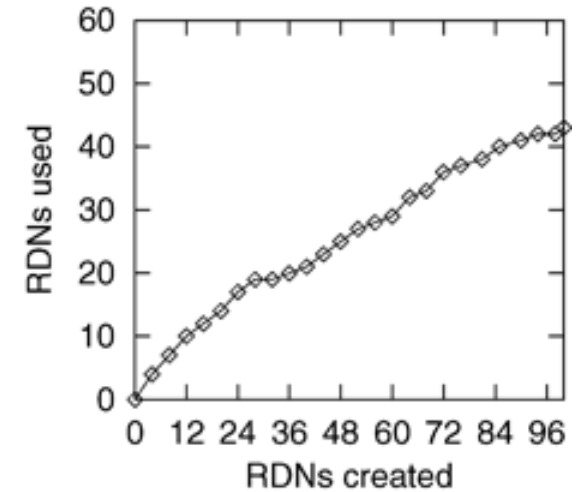
Amit / Markovitch (2006)

- Asaf Amit und Shaul Markovitch:
Learning to bid in bridge (2006)
- Stuart Russel und Peter Norvig:
Artificial Intelligence
- Wikipedia:
 - Contract Bridge
 - Bidding system
 - 2/1 game forcing

Anhang: Evaluation



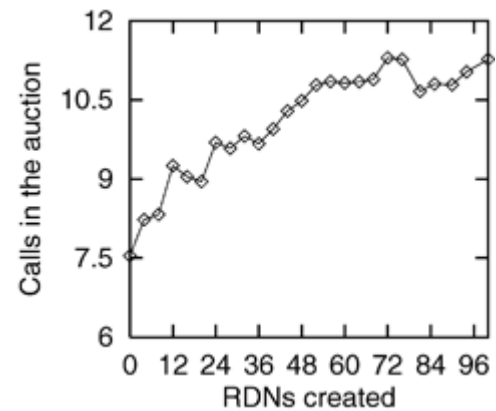
(a)



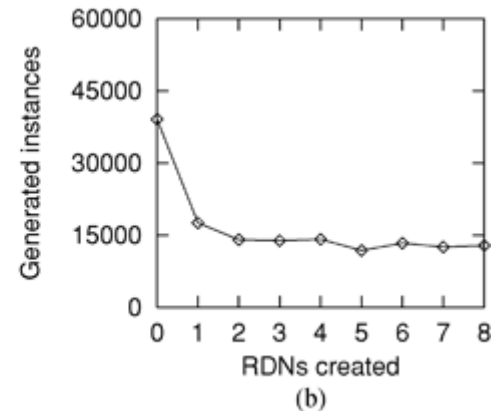
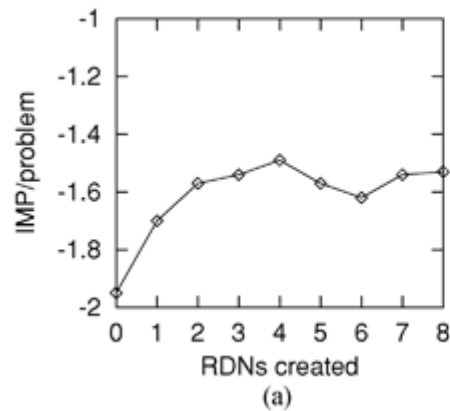
(b)

- Lernen verringert Aufrufe der Suche
- neue Knoten im Entscheidungsnetzwerk werden benutzt

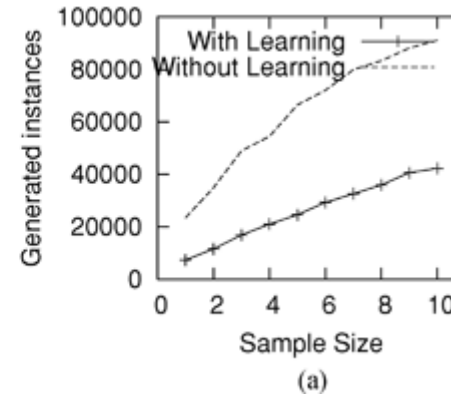
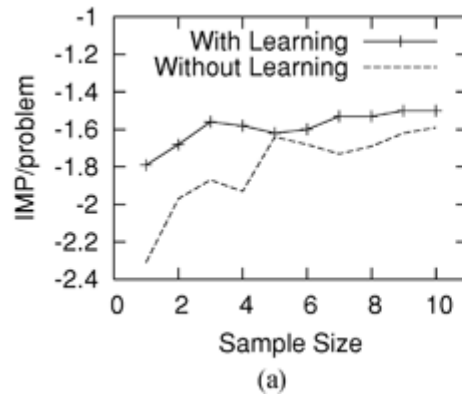
Anhang: Evaluation



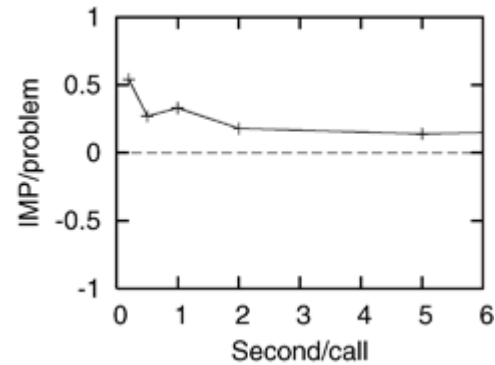
Anhang: Evaluation



Anhang: Evaluation



Anhang: Evaluation



Anhang: Entscheidungsnetzwerk (Attribute)

A.1. Attribute Classes for Evaluating High Cards

Attribute Description

HCPA The weighted count of High Card Points where $A = 4$; $K = 3$; $Q = 2$ & $J = 1$. Range is $[0, 10]$ for each suit and $[0, 37]$ for the whole hand

CONTROLSA The weighted count of aces and kings: $A = 2$ & $K = 1$. Range is $[0, 3]$ for each suit and $[0, 12]$ for the hand

HCA The count of aces, kings and queens. Range is $[0, 3]$ for a single suit and $[0, 12]$ for the hand.

HONORA The count of cards higher than 9. Range is $[0, 5]$ for a single suit and $[0, 13]$ for the hand.

ACEA The count of aces. Ranges are $[0, 1]$ and $[0, 4]$.

KINGA The count of kings. Ranges are $[0, 1]$ and $[0, 4]$.

QUEENA The count of queens. Ranges are $[0, 1]$ and $[0, 4]$.

JACKA The count of jacks. Ranges are $[0, 1]$ and $[0, 4]$.

TENA The count of tens. Ranges are $[0, 1]$ and $[0, 4]$.

RKCBS The count of aces in the whole hand the count of kings in the suit. Range is $[0, 5]$.

A.2. Attributes for Evaluating the Distribution of the Hand

Attribute Description

BAL How balanced is the hand? (0 = balanced-6 = freak)

DPTA Points for short non-trump suits. 1 point for each doubleton, 2 points for each singleton and 3 points for void. Ranges are $[0, 3]$ and $[0, 9]$.

DIF2S The difference in length between 2 suits. Range is $[-13, 13]$.

SUITS The number of suits that are longer than 3. Range is $[1, 3]$.

SUITLENS The length of the suit. $[0, 13]$.

SUITLEN2S The total length of the two suits. $[0, 13]$.

LMAJ The length of the longer major suit. $[0, 13]$.

LMIN The length of the longer minor suit. $[0, 13]$.

LONGESTS Is the suit the longest in the hand? $[0, 1]$.

LONGS Is the suit not shortest in the hand? $[0, 1]$.

FIRSTBIDS Is the suit appropriate for opening in 5M system? $[0, 1]$.

UPLINE2S Is it appropriate to bid s1 when the last bid suit is s2? $[0, 1]$.

A.3. Attributes for Evaluating the Quality of the Suits

Attribute Description

BIDDBLES The number of times we can bid the suit throughout the auction. $[0, 8]$.

INTERA The weighted count of the intermediate cards ($T = 2$, $9 = 1$, $8 = 0.5$). Ranges are $[0, 3.5]$ for a suit and $[0, 14]$ for the hand.

LOSERSTA Expected number of losers in s2 (a suit or the hand) when s1 is played as trump. Ranges: $[0, 3]$ and $[0, 12]$.

STOP-NTS Stopper quality of the suit when playing in NT. $[0 = \text{no stopper} - 4 = \text{double stopper}]$.

STOP-TRUMPTS Stopper quality of the s2 when playing with trump s1. $[0 = \text{no stopper} - 4 = \text{double stopper}]$.

QLTYS A general quality assessment of the suit $[0 = \text{bad} - 9 = \text{excellent}]$.

A.4. Attributes for Evaluating the Quality of the Hand

Attribute Description

NTLP Lindelöf points (Lindelöf, 1983) for the purpose of NT playing.

RATIO The ratio between the sum of HCP of the two longest suits and the total HCP of the whole hand. $[0, 1]$.

POINTSTA HCPTA + DPTA.

UNLOSS Estimated number of losers in the hand without partner support for the longest suit.

A.5. Dynamic Attributes

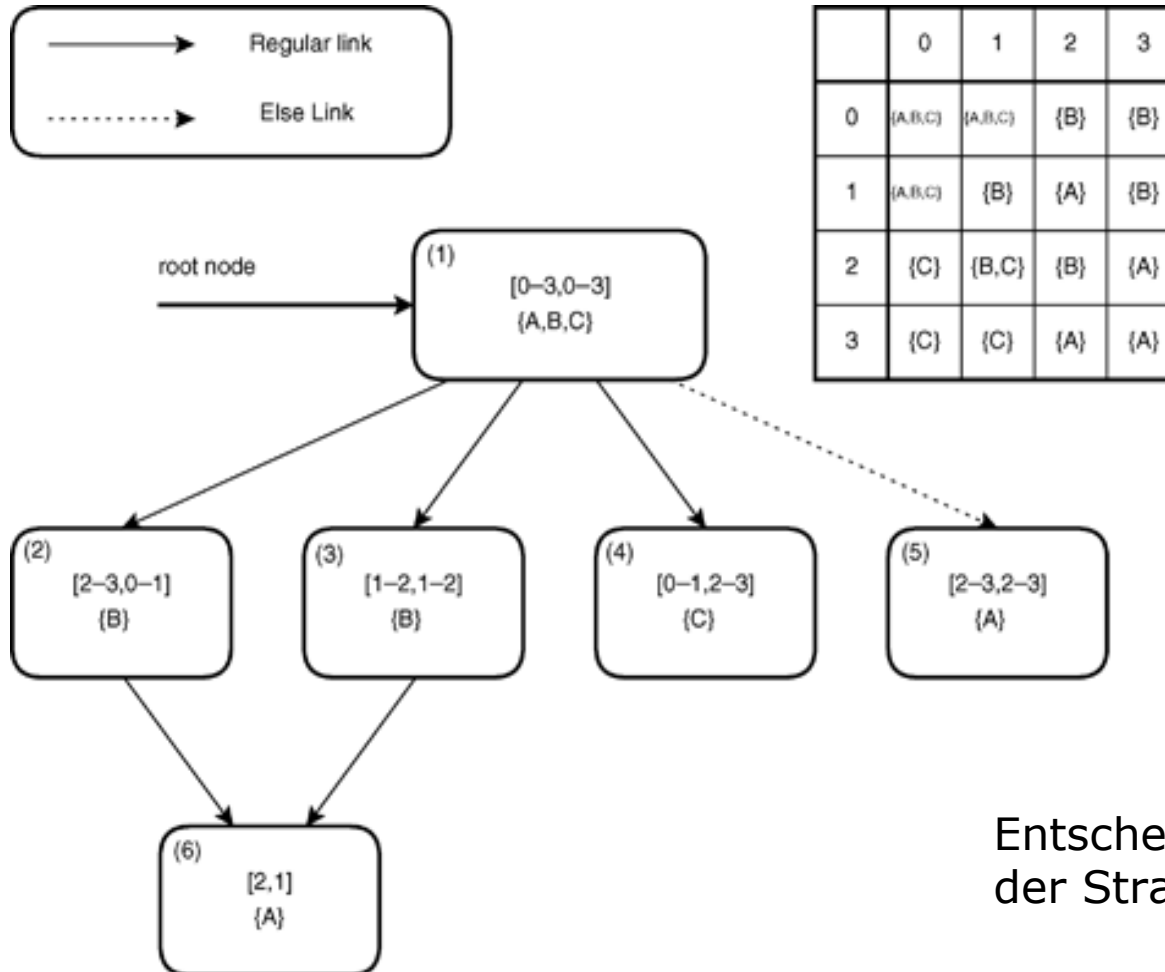
Attribute Description

COVERTA Estimated number of cover cards in s2 when s1 is played as trump. $[0-13]$.

FASTT* A Estimated number of fast losers (that can be cashed by the opponent) in s2 when s1 is played as trump.

LP Lindelöf points.

Anhang: Entscheidungsnetzwerk



Entscheidungsnetzwerk
der Strategie-Funktion