Multi-Agent MCTS

TECHNISCHE UNIVERSITÄT DARMSTADT

Seminar aus maschinellem Lernen

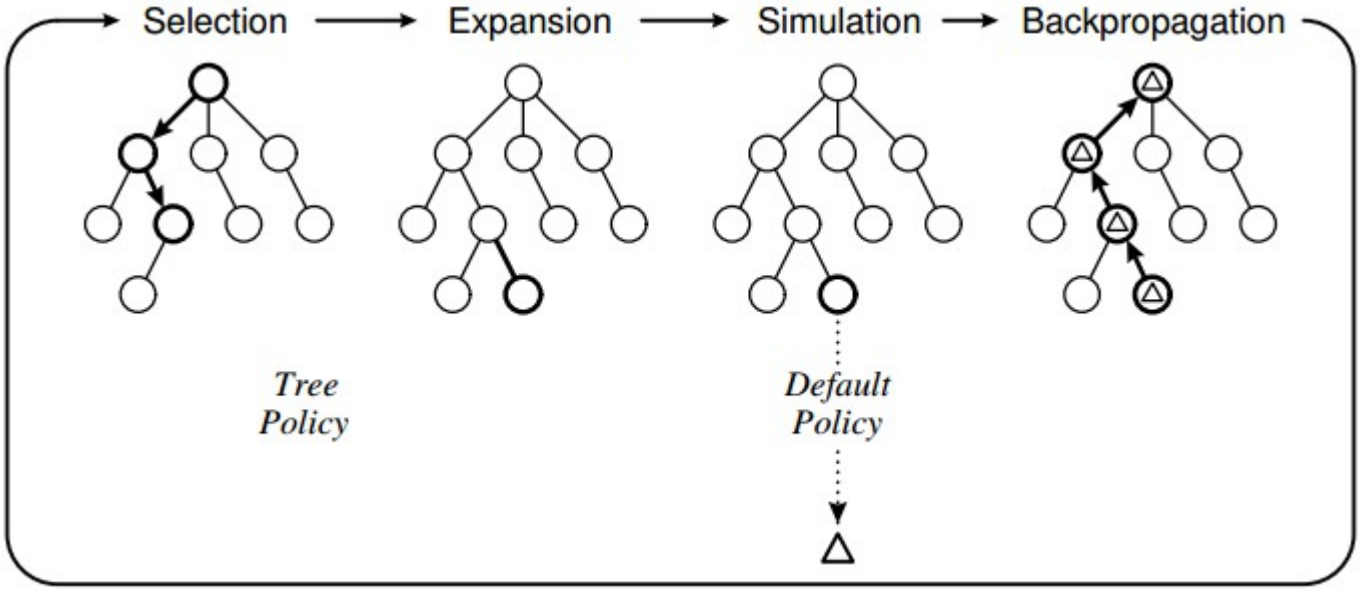
Tjark Vandommele

Agenda

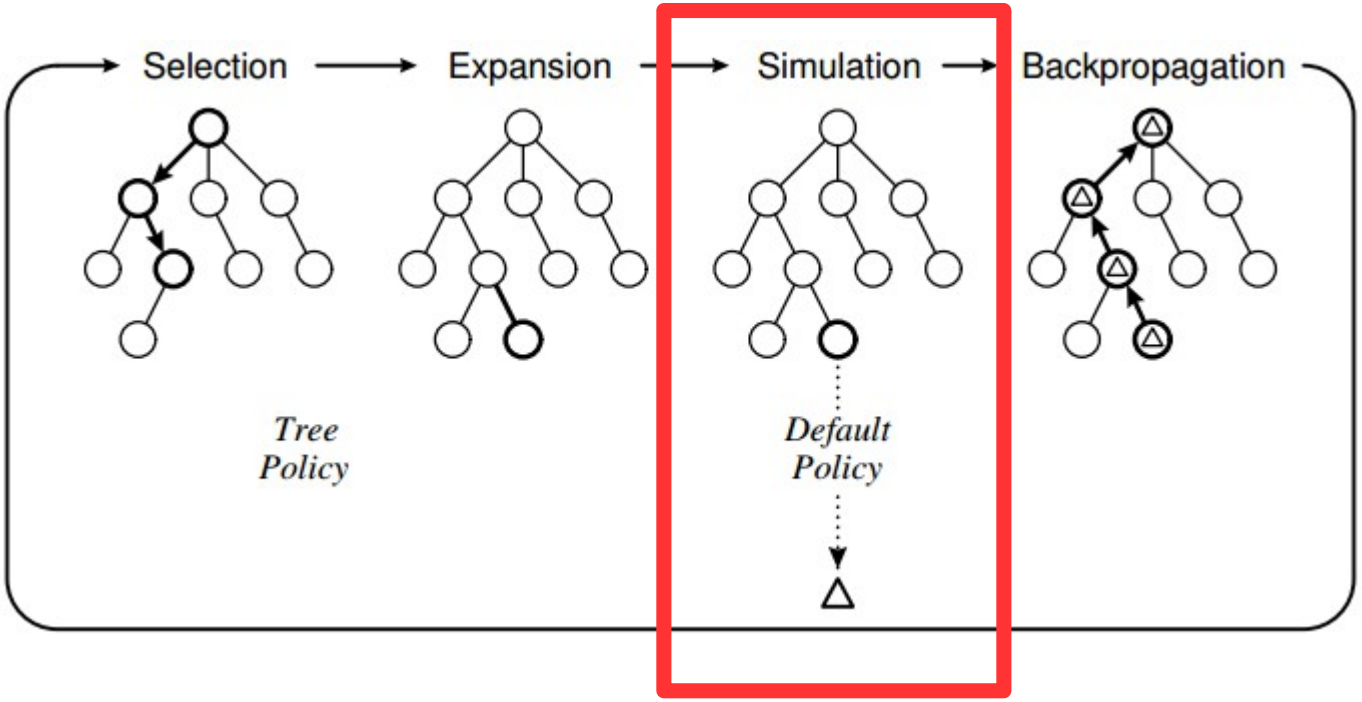
TECHNISCHE UNIVERSITÄT DARMSTADT

- Motivation
- Der Algorithmus
- Implementation
- Performance
- Fazit & Ausblick
- Quellen

Motivation


Selection
Expansion
Simulation
Backpropagation
Tree Policy
Default Policy

Motivation



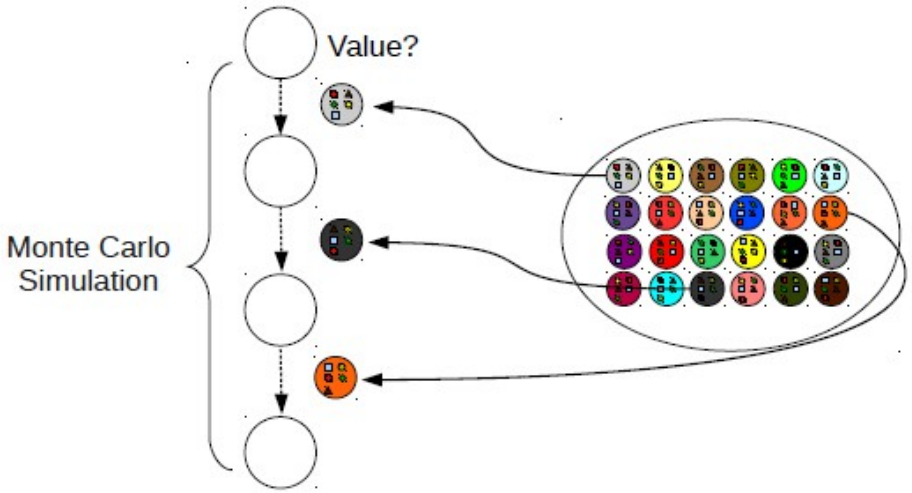
- Random policy führt zu schlechten Zügen
- Eine konkrete Heuristik ist möglicherweise nicht die Beste

→ Verschiedene Heuristiken (Agenten) nutzen

Der Algorithmus

TECHNISCHE UNIVERSITÄT DARMSTADT

Ursprüngliche Idee: Turnier unterschiedlicher Agenten gegeneinander
→ Hoher Aufwand, kaum Verbesserung

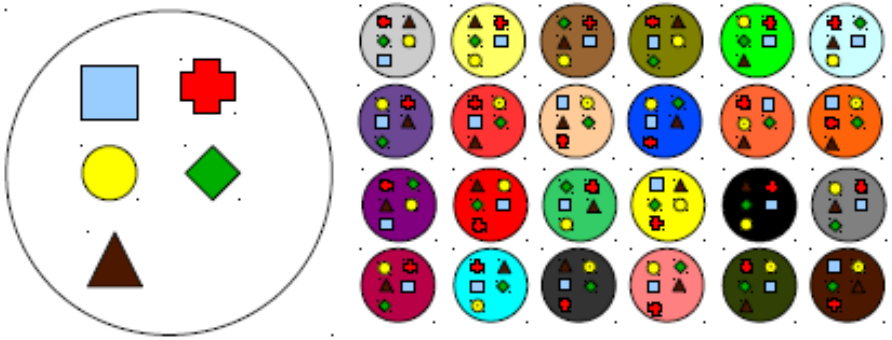


Neue Idee: Für jeden Simulations-Schritt einen anderen Agenten wählen
→ Mehr Exploration
→ Weniger schwache Züge

Implementation
Agenten erzeugen

TECHNISCHE UNIVERSITÄT DARMSTADT

- Basis ist die hierarchische Heuristik von *Fuego*
  - Nakade
  - Atari Capture
  - Atari Defend
  - Lowlib
  - Pattern
- Neuordnung der 5 Teil-Heuristiken führt zu 120 hierarchischen Heuristiken



$$5! = 120$$

Implementation
Agenten auswählen

TECHNISCHE UNIVERSITÄT DARMSTADT

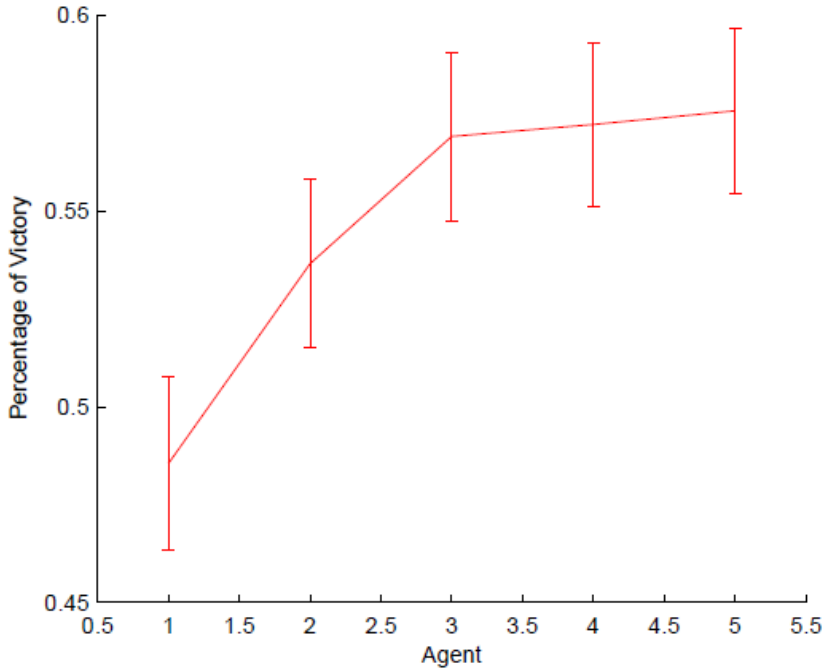
- Erster Versuch mit 120 Agenten vs. Fuego nur 41,2% (± 2,1%) Erfolg
  → Scheinbar schlechte Agenten verantwortlich

- Einfacher Lernalgorithmus, um die besten Agenten zu finden
  1. Beginne mit dem original Fuego Agenten
  2. Füge einen zufälligen Agenten hinzu
  3. Überprüfe ob neue Agentenkombination besser ist als bisherige
  4. Wenn ja behalte den neusten Agenten; wenn nicht entferne ihn
  5. Wiederhole bis alle 120 Agenten getestet wurden

- Greedy hill climbing Algorithmus
- Führt nicht zwangsläufig zur optimalen Agentenkombination
- Potential für Verbesserungen

Performance
Reiner Zufall

TECHNISCHE UNIVERSITÄT DARMSTADT

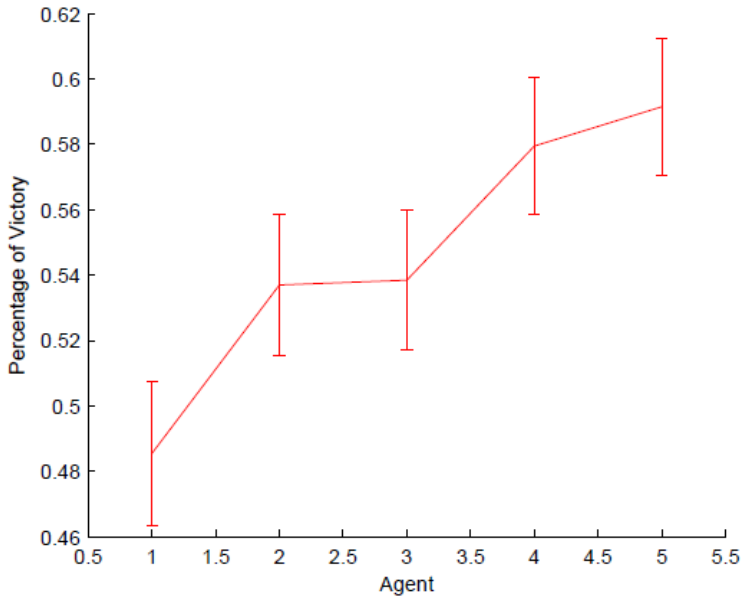


- Alle Experimente auf 9x9 Go Brett
- Je 500 Spiele als Weiß und Schwarz
- Gegner immer Standard Fuego
  → Eine Simulation pro Blatt

- 5 Agenten gefunden
- 57,55% (± 2,1%) Erfolg
- Verbesserung um 9%

Agent Number	Percentage of Victory
0	$48.50\% \pm 2.20\%$
5 (α)	$52.85\% \pm 2.15\%$
6 (β)	$53.60\% \pm 2.15\%$
64	$57.30\% \pm 2.15\%$
70	$29.60\% \pm 1.90\%$

Performance
Händisch angepasst



- 15 „gute“ Agenten ausgewählt
- Teils intuitiv, teils via trial and error
- „Gute“ Agenten werden zuerst überprüft

- 5 Agenten gefunden
- 59,15% (± 2,1%) Erfolg
- Weitere Verbesserung um 1,6%
- Teils identische Agenten wie zuvor

Agent Number	Percentage of Victory
0	$48.55\% \pm 2.20\%$
1 (α)	$54.40\% \pm 2.15\%$
2 (β)	$54.55\% \pm 2.15\%$
3	$57.05\% \pm 2.15\%$
42	$50.90\% \pm 2.20\%$

Fazit & Ausblick

TECHNISCHE UNIVERSITÄT DARMSTADT

Fazit:

- Multi-Agent Ansatz kann Single-Agent Fuego schlagen
- Ein Team aus Agenten kann stärker sein als seine Individuen
- Auch „schlechte“ Agenten können das Team stärker machen
- Erstaunlicherweise sehr kleine Teams aus fünf Agenten

Ausblick:

- Algorithmus um Agenten auszuwählen sollte verbessert werden
- Gruppen von Agenten sollten betrachtet werden
- Andere Heuristiken bzw. Heuristik-Bausteine denkbar
- Performance gegen andere Ansätze muss untersucht werden

Quellen

TECHNISCHE UNIVERSITÄT DARMSTADT

- L. S. Marcolino and H. Matsubara, **“Multi-Agent Monte Carlo Go**”
  - Proc. Int. Conf. Auton. Agents Multi. Sys., Taipei, Taiwan,2011, pp. 21–28.

- M. Enzenberger, M. Müller, B. Arneson, and R. B. Segal, “**Fuego - An Open-Source Framework for Board Games and Go Engine Based on Monte Carlo Tree Search**”
  - IEEE Trans. Comp. Intell. AI Games, vol. 2, no. 4, pp. 259–270, 2010.

Fragen

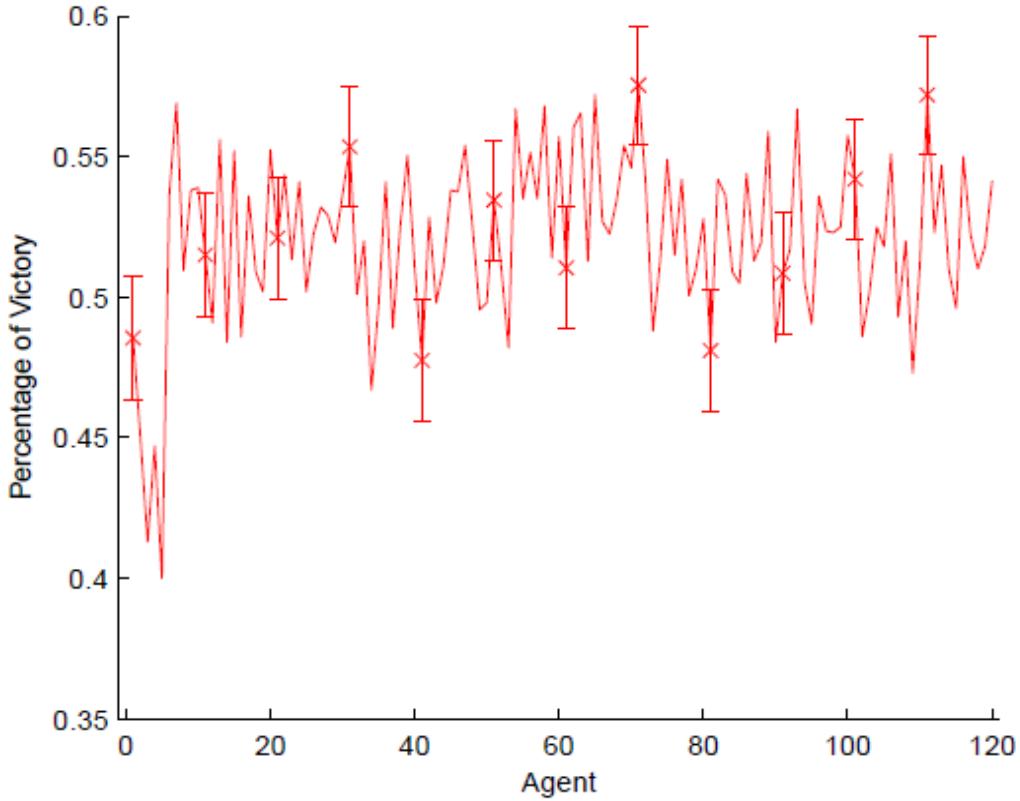
TECHNISCHE UNIVERSITÄT DARMSTADT

Danke für die Aufmerksamkeit!
Gibt es noch Fragen?

Backup

TECHNISCHE UNIVERSITÄT DARMSTADT

Zufällige Agentenreihenfolge

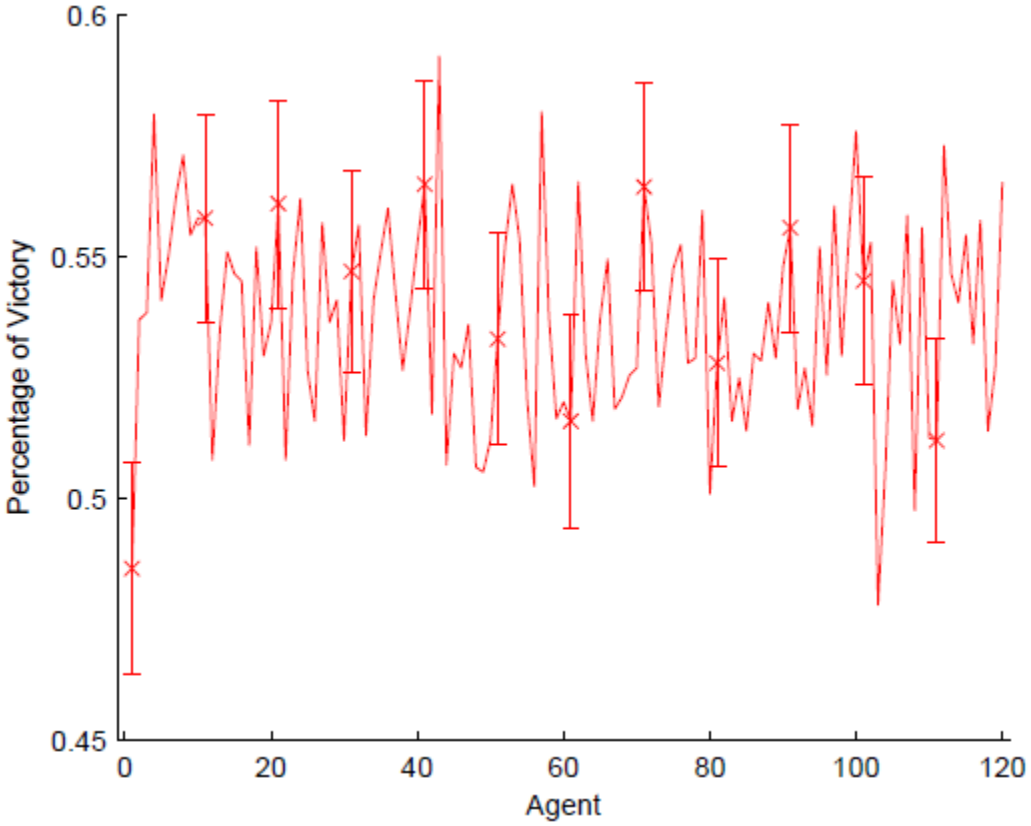


N	AC	AD	L	P
AD	N	AC	P	L
AD	N	P	AC	L
AD	AC	P	N	L
N	AC	P	L	AD

Backup

TECHNISCHE UNIVERSITÄT DARMSTADT

Angepasste Agentenreihenfolge

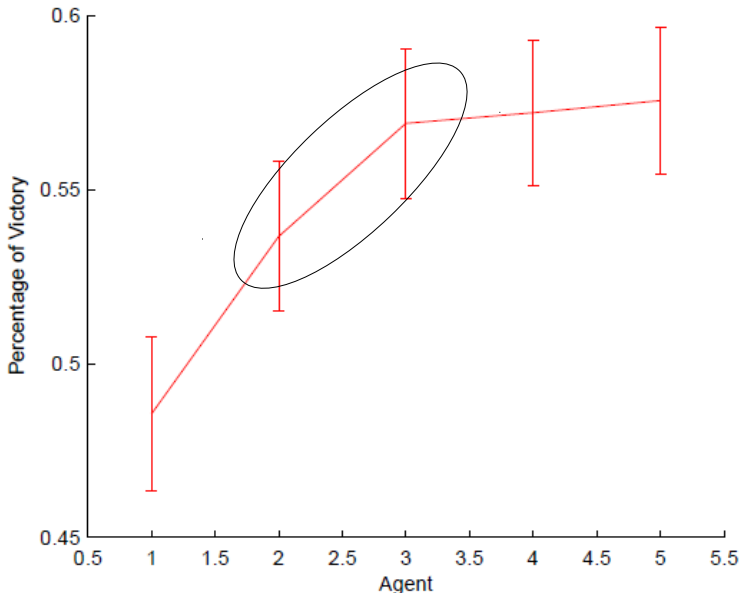


N	AC	AD	L	P
AD	N	AC	P	L
AD	N	P	AC	L
AD	P	L	N	AC
AC	N	AD	L	P

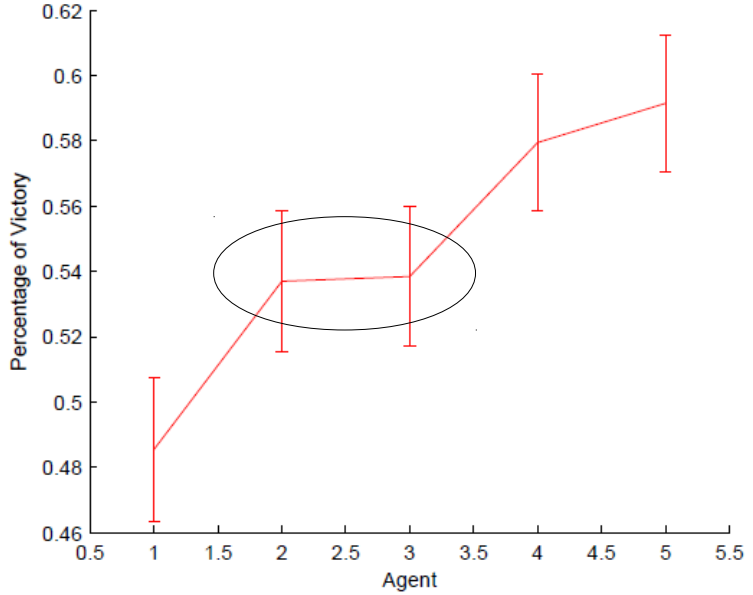
Backup

TECHNISCHE UNIVERSITÄT DARMSTADT

Performance Vergleich



Agent Number	Percentage of Victory
0	$48.50\% \pm 2.20\%$
5 (α)	$52.85\% \pm 2.15\%$
6 (β)	$53.60\% \pm 2.15\%$
64	$57.30\% \pm 2.15\%$
70	$29.60\% \pm 1.90\%$



Agent Number	Percentage of Victory
0	$48.55\% \pm 2.20\%$
1 (α)	$54.40\% \pm 2.15\%$
2 (β)	$54.55\% \pm 2.15\%$
3	$57.05\% \pm 2.15\%$
42	$50.90\% \pm 2.20\%$