

UCT: Selektive Monte-Carlo-Simulation in Spielbäumen

Michael Wächter

Motivation

- wissensbasierte vs. suchintensive Ansätze
- Standard-Algorithmen (z.B. Alphabet) nicht anwendbar

Motivation

- daher: MC-Simulation
- reine MC-Simulation hat Schwächen
- daher: UCT

Überblick

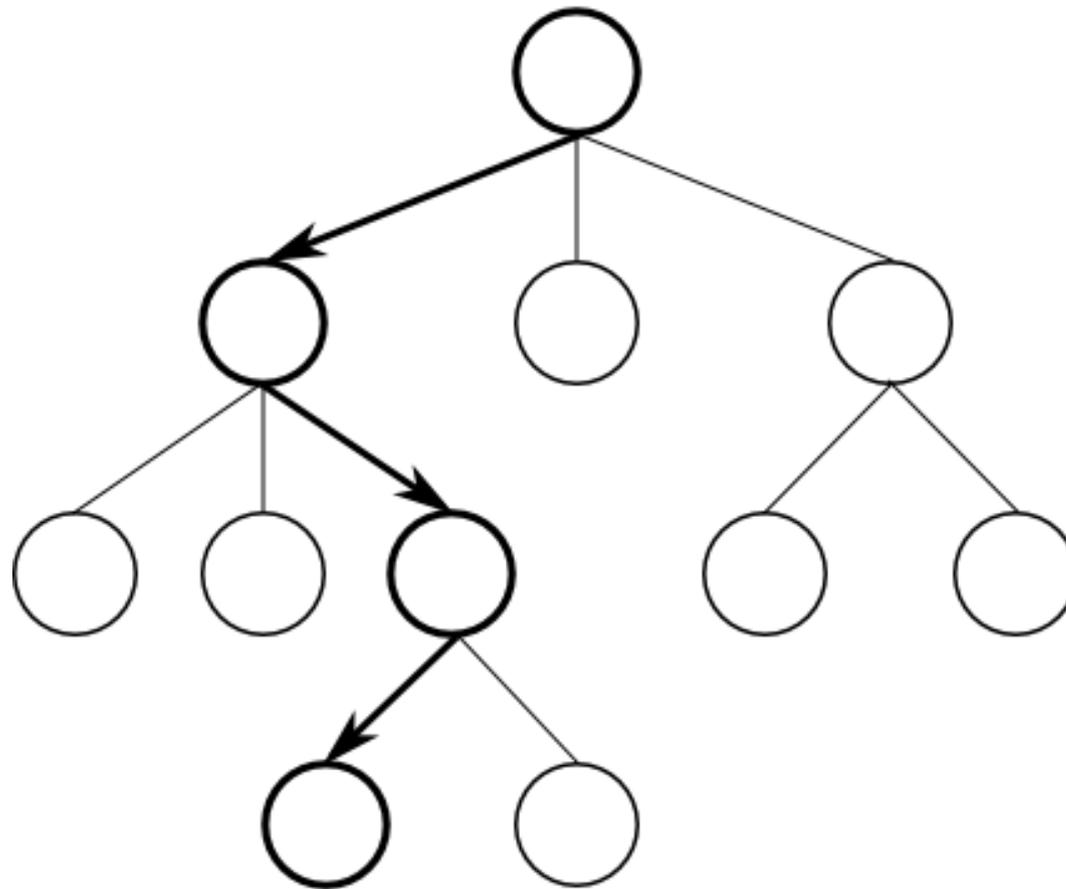
- Motivation
- Funktionsweise
- Eigenschaften
- UCT und Pokern
- Erweiterungen

Funktionsweise

1. Selektion
2. Expansion
3. Simulation
4. Backpropagation

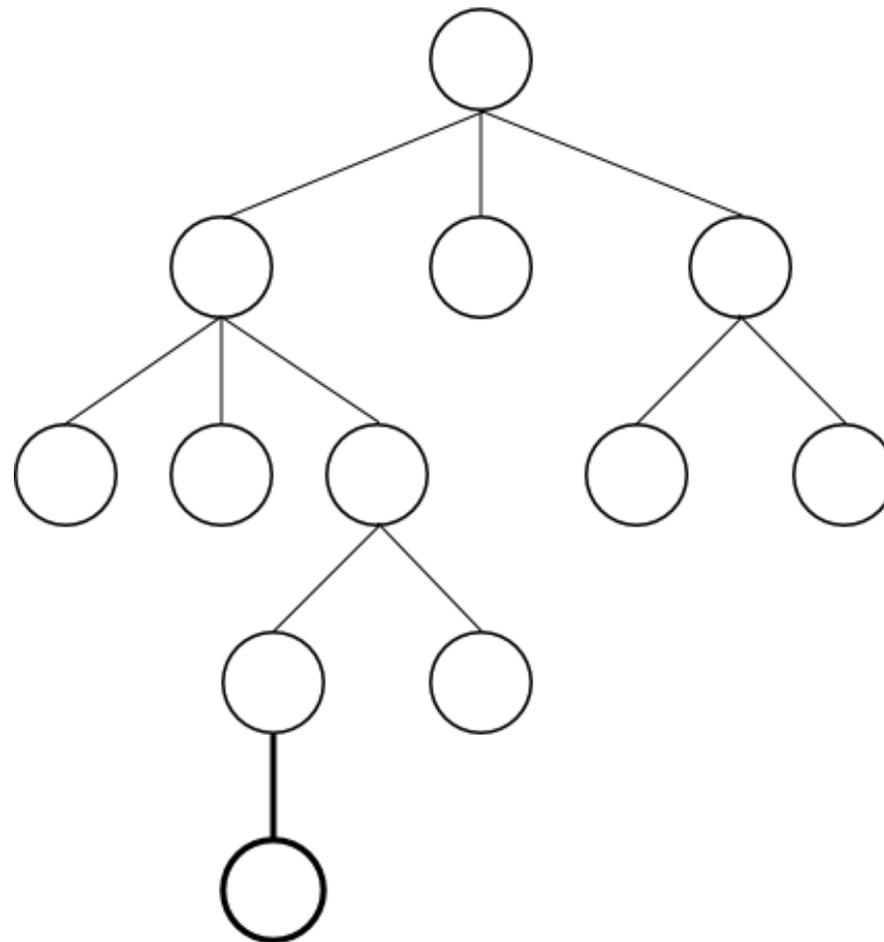
Funktionsweise

Selektion



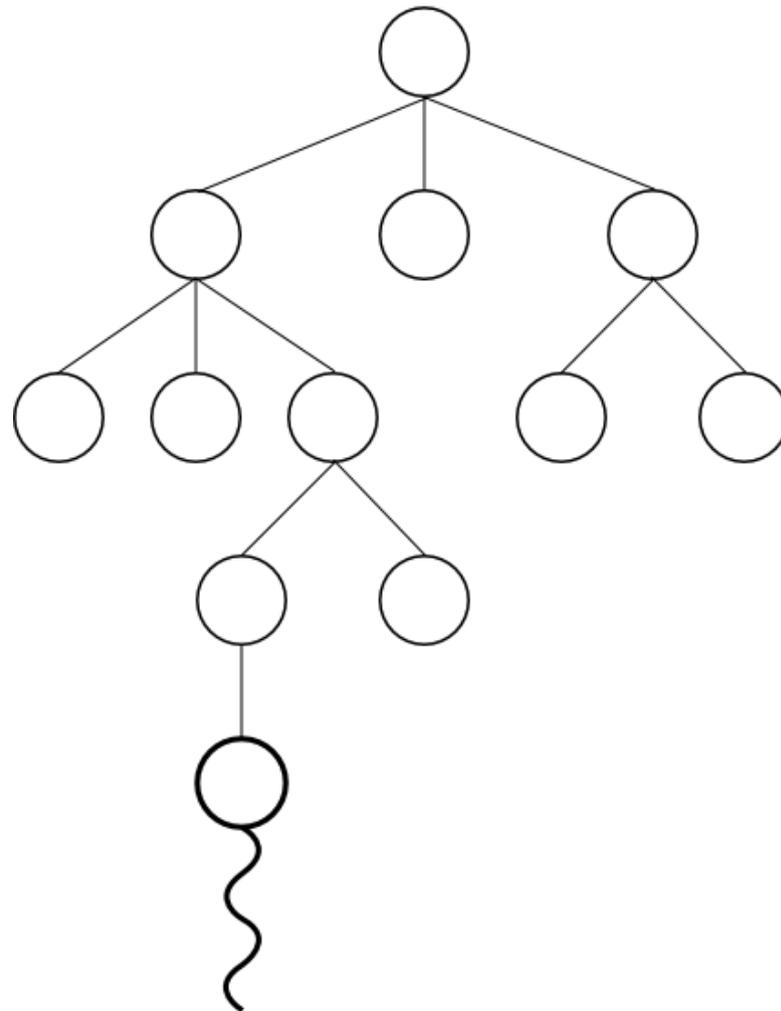
Funktionsweise

Expansion

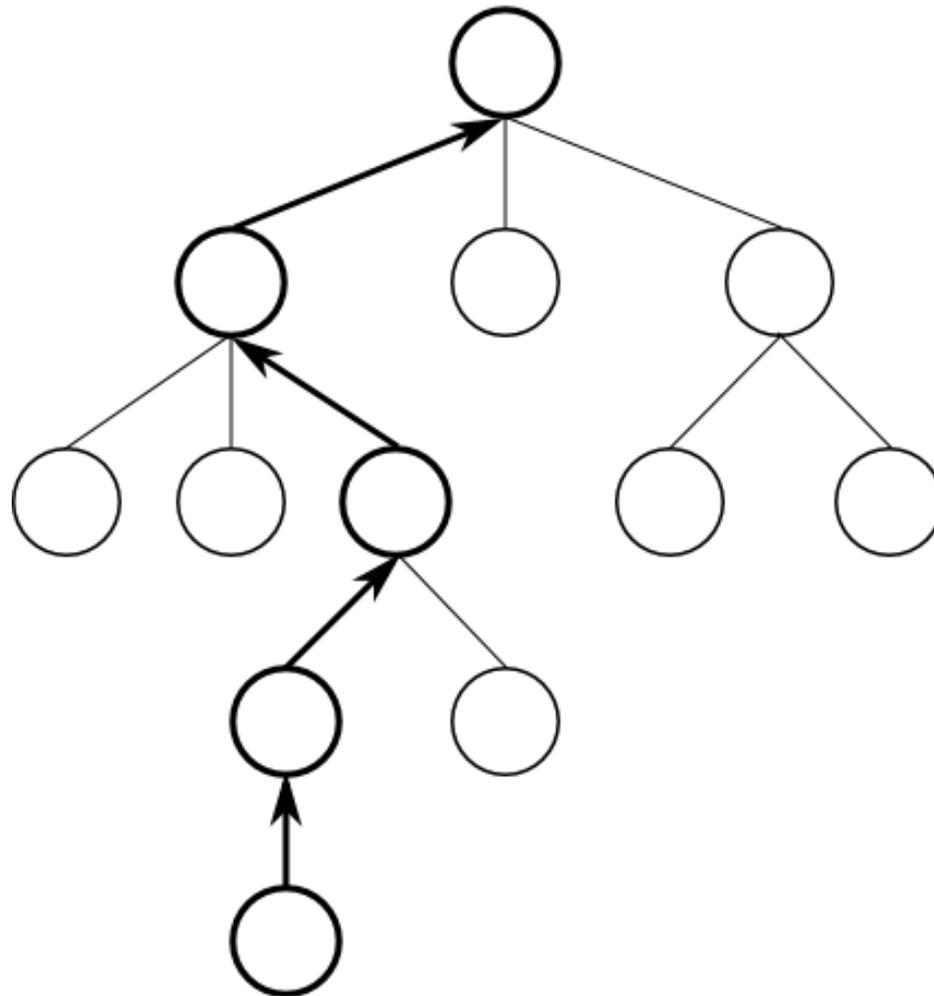


Funktionsweise

Simulation



Funktionsweise Backpropagation



Eigenschaften

- Konvergenz
- kein Expertenwissen (oder nur zur Leistungsverbesserung)
- „Anytime“-Eigenschaft
- Strategie vs. Taktik

Eigenschaften

- Ergebnisse in realen Systemen:
 - Trendwende in Go-Programmen
 - in Pokern wissenschaftlichen Ansätzen überlegen (Billings et al., 1999)

UCT und Pokern

- Nötige Adaptionen:
 - Nichtdeterminismus
 - unvollständiges Wissen
 - mehr als zwei Spieler

Erweiterungen

- Simulationsstrategien
 - Patterns, handkodierte Expertenwissen und KNNs
- Spielerprofile
- Parallelisierung
- Progressive Bias
- Erweiterungen aus Schach
 - Transposition Tables und Pondering

Vielen Dank für Ihre Aufmerksamkeit!