



Samuel's Checkers Program

Seminar: Knowledge Engineering und Lernen in Spielen
29.06.2004

Ge Hyun Nam

Überblick



- Einleitung
- Basis Dame-Programm
- Maschinelles Lernen
 - Auswendiglernen
 - Verallgemeinerndes Lernen
 - Vergleich

Einleitung



- Maschinelles Lernen angewandt auf Spiele
- Wieso Dame?

- Einfachheit der Regeln ermöglicht Konzentration auf Lerntechniken
- Erfüllt Grundeigenschaften einer intellektuellen Aktivität
 - Aktivität darf nicht deterministisch sein
 - Bestimmtes Ziel muss existieren
 - Regeln der Aktivität müssen exakt sein
 - Hintergrundwissen bzgl. Aktivität, gegen diese der Lernprozess getestet werden kann
 - Aktivität sollte Gruppe von Menschen bekannt sein

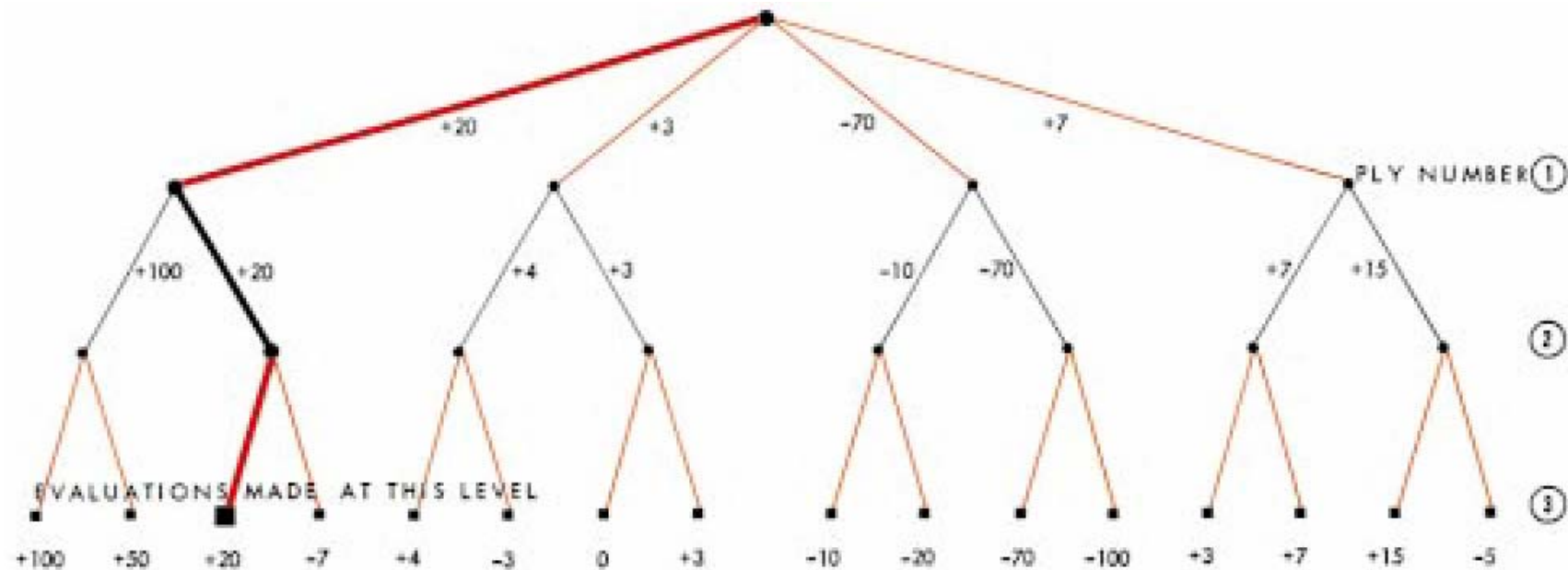


Basis Dame-Programm

- Zuerst muss Computer programmiert werden gültiges Dame zu spielen
 - Regeln des Spiels in Maschinensprache ausdrücken
 - Gegnerischen Zug akzeptieren
 - Zug des Computers melden
- Basis Dame-Programm
- Computer spielt indem er ein paar Züge voraus schaut und resultierende Spielpositionen bewertet

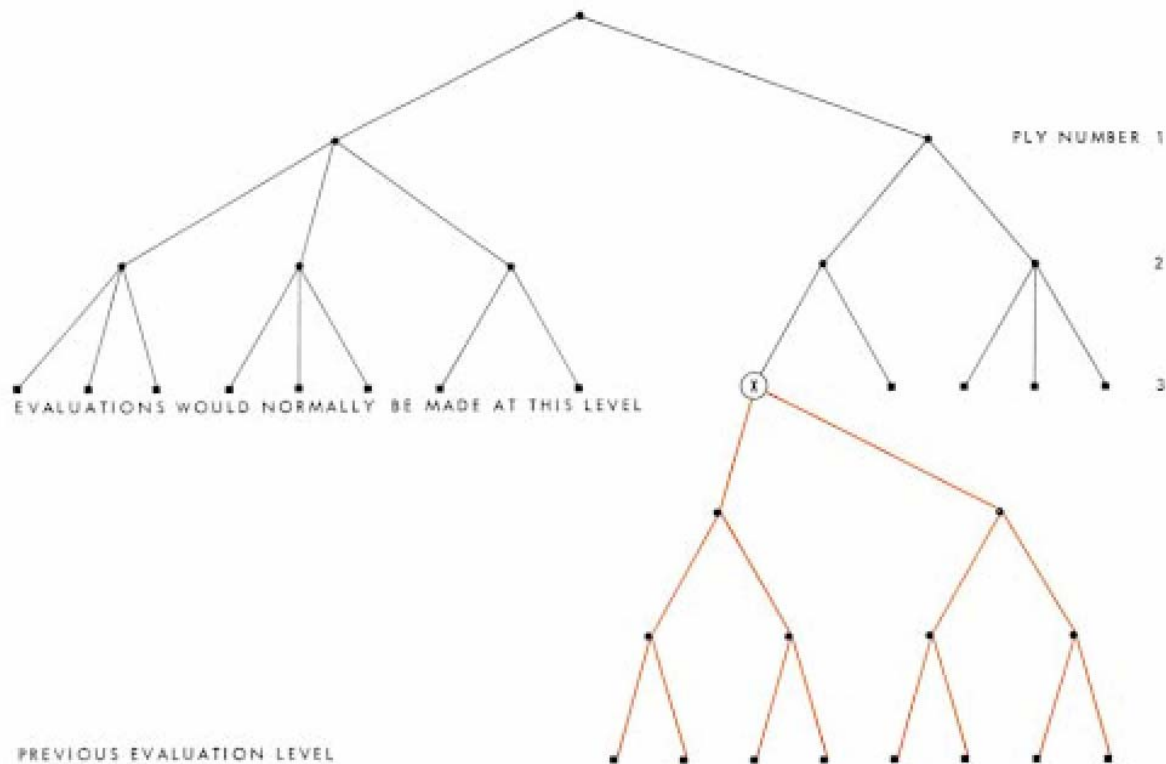
Basis Dame-Programm (2)

- Parameter der Evaluationsfunktion werden entweder festgesetzt oder das Programm kann diese auswählen
- Minimax-Suche (in späteren Versionen auch alpha-beta pruning)



Auswendiglernen

- Einfachste Art von Lernen
- Alle während des Spiels auftauchende Positionen werden mit ihrer Bewertung gespeichert
- Einsparen von Berechnungszeit



Auswendiglernen (2)



- Weitere Eigenschaften:
- ‚Sense of direction‘
 - Sieg erzwingen
 - Tiefere Spielzüge
- Katalogisierung von Spielpositionen
 - Standardisierung: Weiss ist am Zug
 - bei Damespielen Spielfeld spiegeln
 - Gruppierung in Datensätze
- Redundanz vermeiden
- Verwerfen von Spielzügen
 - Häufigkeit der Benutzung mittels Alter-Attribut
 - Auffrischen: bei Verweis Alter halbieren
 - Vergessen: bei Erreichen eines Maximum-Wertes löschen
- Bewertungsfunktion: (1) piece advantage, (2) denial of occupancy, (3) Mobilität, (4) Hybrid Term: Zentrumskontrolle + piece advantage

Auswendiglernen (3)



- Tests

- Spiel gegen sich selbst
- Spiel gegen verschiedene Menschen (teilweise Damemeister)
- Book Games

- Resultat

- Programm lernt sehr gutes Eröffnungs- und Endspiel
- Keine grosse Verbesserung im Mittelspiel
- Besser als überdurchschnittlicher Anfänger



Verallgemeinerndes Lernen

- Auswendiglernen ist durch Speicherkapazität beschränkt
- Effektivere Lernmethode: gemachte Erfahrungen verallgemeinern und nur diese Verallgemeinerungen speichern
- Beinhaltet folgende Abstraktion:
 - Programm ist in der Lage die Terme des Bewertungspolynoms selbst auszuwählen und das Vorzeichen und die Größe der Koeffizienten anhand des Spielerfolgs zu bestimmen

Verallgemeinerndes Lernen (2)

- Bewertungspolynom:
 - $V(\text{board}) = c_1 * F_1 + c_2 * F_2 + c_3 * F_3 + \dots + c_n * F_n$
- Programm übernimmt Rolle beider Spieler
 - Alpha verallgemeinert seine Erfahrung nach jeden Zug, indem die Koeffizienten angepasst werden
 - Beta benutzt das ganze Spiel lang das gleiche Bewertungspolynom
 - Alpha im Spiel gegen Menschen, Alpha vs. Beta im Selbstspiel
- Bei Sieg von Alpha, erhält Beta sein Bewertungssystem
- Liegt hingegen Beta in Führung -> ‚black mark‘
- Bei drei ‚black marks‘ -> auf den falscher Weg; führender Koeffizient wird auf Null gesetzt

Verallgemeinerndes Lernen (3)

- Modifikation der Termkoeffizienten anhand des Spielerfolges beinhaltet folgende Schwierigkeit:
 - Nicht immer sicher, welcher Spielzug zum Gewinn oder Verlust geführt hat
 - Also nicht sicher, welcher Term neu gewichtet werden sollte
- Lösung:
- Während des Spiels nach jedem Zug Bewertungspolynom bewerten
- Vergleich der Bewertungen für die Stellungen, die sich ergeben, einmal mit dem Polynom und zum anderen nach der Minimax Methode

Verallgemeinerndes Lernen (4)

- Differenz ist Delta, Maß für die Güte des Bewertungspolynoms
- Wenn Delta positiv, Polynom zu pessimistisch
 - Alle positiven Terme stärker bewerten
- Wenn Delta negativ, Polynom zu optimistisch
 - Alle negativen Terme stärker bewerten
- Nach Beendigung des Spiels werden die für jeden Zug ermittelten Delta-Werte mit dem Vorzeichen der verschiedenen Terme während eines ganzen Spiels korreliert
- Koeffizienten werden je nach berechneter Korrelation für das nächste Spiel verändert



Auswendiglernen vs. Verallgemeinerndes Lernen

- Auswendiglernen führte sehr schnell zu einer guten Eröffnung
- Aber schlechtes Mittelspiel
- Verallgemeinerndes Lernen hingegen nicht so gutes Eröffnungsspiel
- Aber drastische Verbesserung des Mittelspiels
- Einmal in Steinvorteil, hatte der Gegner meist keine Chance mehr
- Meisterliches Spiel bei Kombination beider Lernmodi

Quellenangaben



- A. L. Samuel: Some Studies in Machine Learning Using the Game of Checkers
- A. L. Samuel: Some Studies in Machine Learning Using the Game of Checkers. II – Recent Progress
- S. Russel, P. Norvig: Artificial Intelligence A Modern Approach
- R. Sutton, A. Barto: Reinforcement Learning: An Introduction