



Knowledge Engineering und Lernen in Spielen

Thema: Opening Book Learning
von: Thomas Widjaja

Gliederung

- Allgemeines
- Drei Beispielfverfahren zum Anpassen eines Opening Books
 - Deep Blue
 - Logistello (Michael Buro)
 - Crafty (Hyatt)
- Resümee

Teil I

Allgemeines

Motivation

- Menschliche Spieler können in vielen Situationen „ohne“ nachzudenken spielen.
 - Vorbereitung (Bsp. Schachtheorie)
- Gut für den Computer geeignetes Konzept
 - Computer vergessen nicht!
 - Gute Möglichkeit dem Computer Expertenwissen zugänglich zu machen.

Was ist ein „Book“?

- **Ziel:** Das Spiel des Programms verbessern, indem für wichtige Situationen „gute“ Züge **offline** berechnet werden.
(d.h. es werden entweder Zugfolgen, oder Positionen mit Vorschlägen für Züge in einer Datenbank –**dem Book**– gespeichert)
- Besonders bei:
 - **Spielanfang** / „Opening Book“
 - **Spielende** / „Endgames“

Grund z.B. Schach: die WS eine Situation in einem „Mittelgame“ wieder zu finden ist sehr gering... (Hängt von der Komplexität des Spiels ab.)

Motivation für die Verwendung eines „Opening Books“

- Wozu ein „Opening Book“
 - Strategische Schwächen der meisten Programme besonders zu Beginn des Spiels.
 - „Gute“ Züge mit wenig Rechenaufwand (besonders unter Turnierbedingung vorteilhaft)

Anforderungen an ein **Verfahren zum Anpassen** eines „Opening Books“

- **Automatisch** aus gegebenen Spielinformationen „gute“ Zugfolgen finden.
- **Automatisch** neue Strategien (Variationen) finden und bewerten.

Automatische Anpassung ist wichtig, weil die aktuellen Programme ununterbrochen auf Internetservern spielen und manuelle Anpassung zu aufwendig wäre.

Anforderung an ein **Opening Book**

- Access-time (Rekonstruktion der Informationen)
- Update-time (beim Lernprozess)
- Speicherplatz
- Dynamisch (z.B. Nicht immer auf die selbe Weise verlieren)

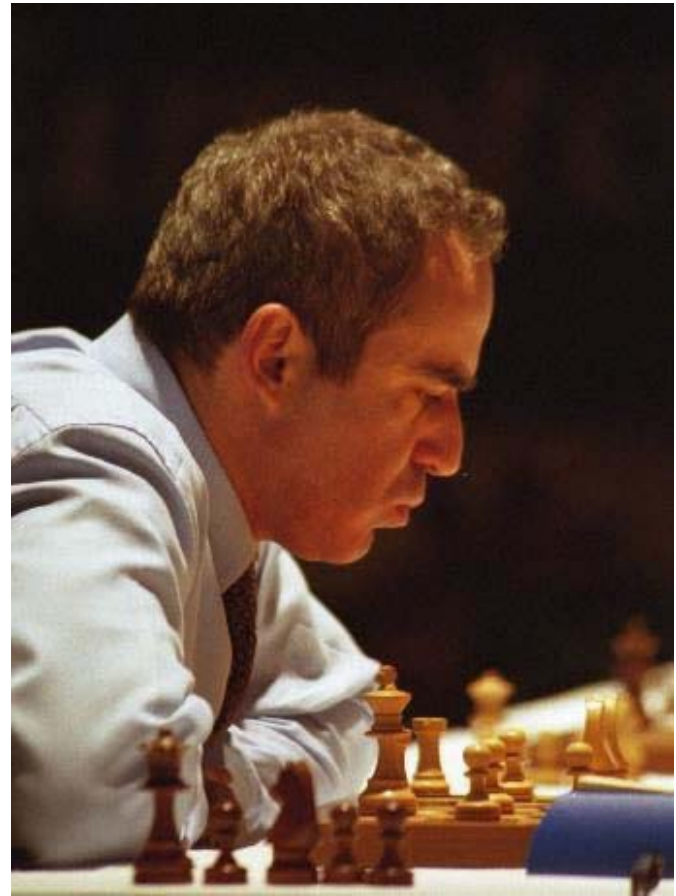
Teil II

Die Drei Strategien

- Deep Blue
- Logistello (Michal Buro)
- Crafty (Hyatt)

Exkurs: Deep Blue vs. Garry Kasparov

- Deep Blue war 1996 der erste Computer der einen amtierenden Schachweltmeister unter regulären Turnierbedingungen schlagen konnte. Das Match entschied Kasparov jedoch 4 zu 2 für sich.



1. Deep Blues Strategie in der Eröffnungsphase

■ „Dualstrategie“

1. Gewöhnliche Datenbank von Zugfolgen

Zugfolgen aus Expertenwissen die (manuell) an die Spielweise des Systems angepasst wurden. (relativ klein)

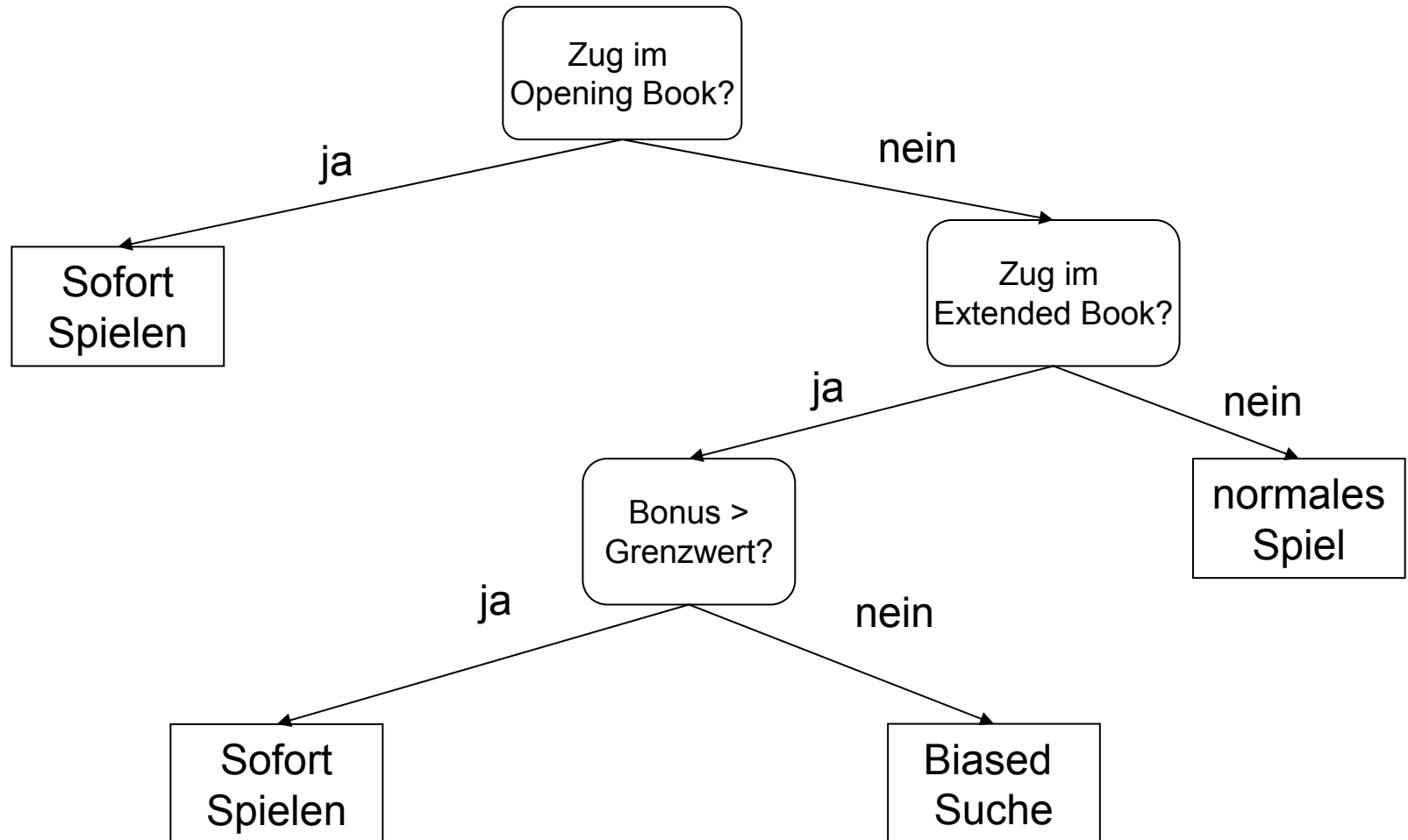
2. Extended Book

Datenbank mit **Bewertungen** der ersten 30 Züge von ca. 700.000 GM-Spielen

Bewertung der Züge (im Extended Book) berücksichtigt folgende Faktoren:

- Wie oft wurde der entsprechende Zug in der Datenbank gefunden?
 - Relative Häufigkeit der gefundenen Zügmöglichkeiten
 - Ergebnis des Zuges
 - Wie lange liegt das Spiel zurück
 - Spielstärke der Spieler
 - Kommentare zu dem Zug (stark/schwach)
 - Gab es (von Kommentatoren) vorgeschlagene Zugalternativen?
- Kann die Bewertung der Züge um den $\frac{1}{2}$ Wert eines Bauern erhöhen oder reduzieren. (klein genug um Raum für die eigene Suche zu lassen)

Schema:



Beispiel zur Funktionsweise dieser Dualstrategie

- Deep Blue vs. Garry Kasparov / 1997 (2. Spiel)
 - Zug 1 bis 9: „Opening Book“
 - Zug 10 bis 17: „Automatic Extended-Book-Move“
 - Zug 18 bis 19: Biased Search durch Boni aus dem Extended Book
 - Zug >19: „Normales“ Spiel

Anmerkung: Die ersten 17 Züge wurden (fast) ohne Rechenaufwand gespielt!

Fazit

- Das Zusammengefasste Expertenwissen wird durch die Boni dazu verwendet die Suche von Deep Blue zu „leiten“.
 - ⇒ Gute Ausnutzung der Eröffnungstheorie, aber trotzdem noch flexibel genug neue Züge zu finden.

2. Die Strategie von Logistello (M. Buro)

- Wir in dem Othello Programm „Logistello“ verwendet.
 - Inkrementelles Aufbauen eines Opening Books:
 - Jede Gespielte Situation kommt in einen Baum. (dieser Baum kann auch aus externen Quellen erstellt werden.)
 - Für jede Situation kommt nicht nur der gespielte Zug, sondern auch der Zug mit der 2. besten (heuristischen) Bewertung in den Baum.
- => Es gibt immer zwei Möglichkeiten: Entweder den bereits gespielten Zug zu wiederholen oder den noch nicht gespielten Zug zu wählen.

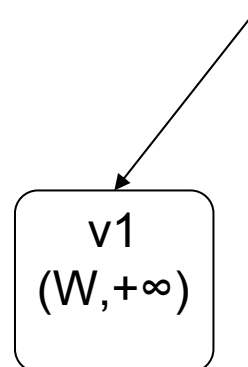
Legende

- Jedes Blatt hat eine Bewertung!
Entweder Heuristik, oder den Ausgang des Spiels.
- Während des Spiels findet das Programm die aktuelle Position im Baum und bewertet diese mit einer Negamax Suche. (Knoten)

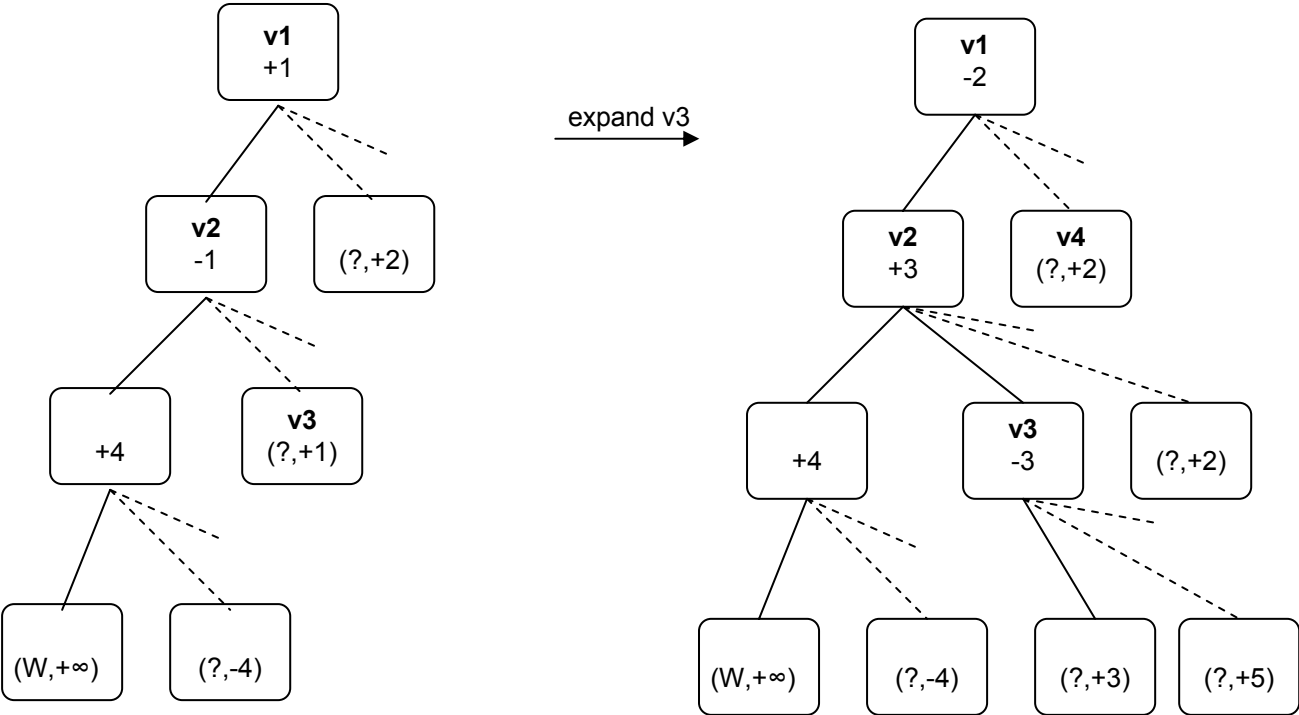
Beispiel für ein Blatt:

Erste Komponente kann „W“ (Win) „L“ (Lost) „D“ (Draw) oder auch „?“ sein.

Die **Zweite** Komponente beschreibt die heuristische Evaluierung



Book Expansion



Details des Algorithmus

- Abwechselnd „Draws“ als Sieg für Spieler A und Spieler B werten.
- Der Algorithmus kann zwischen „privaten“ und „öffentlichen“ Draws unterscheiden.

Entweder schnelle Negamax-Suche, oder die Werte der inneren Knoten müssen im Voraus berechnet werden.

(Vorteil der online Negamax-Suche: Anpassung während des Spiels & Zufallskomponente werden möglich)

Zufallskomponente

Der Zug wird erst nach zwei Durchläufen
Ausgewählt:

1. Der Negamax Wert, v wird bestimmt (analog der eben beschriebenen Vorgehensweise)
2. Es werden die Zugfolgen bestimmt die einen Negamaxwert im Intervall $[v-\epsilon, v]$ haben.
3. Einer der Ermittelten Zugfolgen wird zufällig ausgewählt.

Vor- und Nachteile des Algorithmus

■ Vorteile

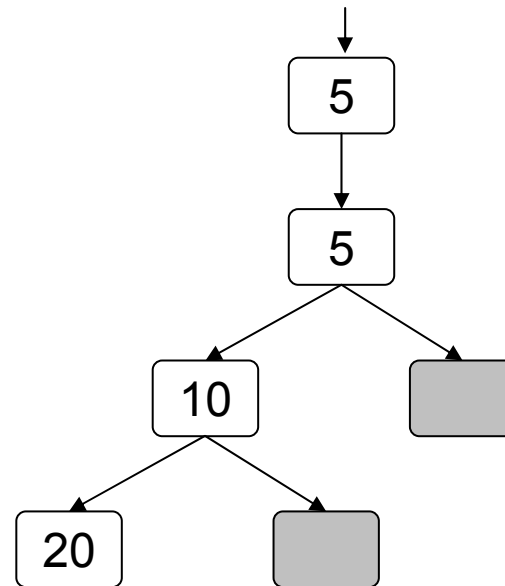
- Verliert nicht immer auf die gleiche Weise
- Bei hinreichend guter Evaluierungsfunktion können neue Zugfolgen entdeckt werden. (Automatische Erweiterung)
- Kann Parallel Berechnet werden (Subtrees auf Prozessoren verteilen)

3. Die Strategie aus „Crafty“ / Hyatt (1999)

- Geht von einem **gegeben** (ggf. manuell erstelltem) Book aus.
- Es werden Heuristiken der ersten 10 Züge nach Verlassen des Books als Bewertung für die „Book“ Zugfolge verwendet. (**Gambits**)
- Im Grunde "Reinforcement learning" Ansatz!

Die Bewertungen werden rückwärts durch den Baum propagiert.

- Ab dem Zug, mit dem das Book verlassen wurde, wird die Bewertung bestimmt.
- Der „Baum“ im Book wird rückwärts durchlaufen und die Bewertung wird jeweils durch die möglichen Alternativen geteilt und zugewiesen.



Graue Felder, repräsentieren Züge die nicht gespielt wurden.

Bewertungsfunktion

Die Bewertungsfunktion beachtet folgende Faktoren:

- Die ermittelte Bewertung (aus den ersten 10 Zügen nach dem Book)
- Stärke des Gegners
- Negative Bewertungen werden stärker gewichtet als positive. (Grund: Blunders)
- Tiefe der Suche für die Züge
Art des Spiels (one minute vs. Open end)

- Die Häufigkeit des Zuges im Book

Distributed Learning

- Crafty ist ein Schachprogramm das jährlich ca. 20.000 Partien auf **verschiedenen** Servern spielt
 - PGN Dateien werden generiert (**P**ortable **G**ame **N**otation)
 - Im Moment: Austausch der Daten per E-Mail, eine Erweiterung wäre eine Client/Server Architektur
- Auch möglich die PGN-Dateien manuell zu importieren. => Anpassung an Gegner

Probleme

- Lernen ist Permanent -> „aging learning“ für schlechte Züge
- Das Gelernte ist Gegnerspezifisch
- Crafty erweitert sein Book nicht automatisch

Teil III

Resümee

Vergleich der Opening Book Strategien

	Deep Blue	Logistello	Crafty
Aus gegebenen Spielinformationen „gute“ Zugfolgen finden	Ja	Ja	Ja
Generierung von neuen Varianten	Nein	Ja	Nein
Dynamische Strategie (z.B. Nicht auf die gleiche Weise verlieren)	Nein (nur sehr langsam)	Ja	Ja

Andere Domänen

- Die hier gezeigten Strategien können auf andere Domänen übertragen werden, wo Expertenwissen zur Verfügung steht.

Beispiele:

- Andere Spiele
- Medizinische Diagnose

Ausblick

- Distributed Learning
- Private & Public Informationen
- Strategische Planung

Quellen

- **Michael Buro.** Toward opening book learning. In H. J. van den Herik and H. Iida, editors, *Games in AI Research*, pages 47-54. Universiteit Maastricht, 2000.
- **Robert M. Hyatt.** Book learning --- a methodology to tune an opening book automatically. *International Computer Chess Association Journal*, 22(1):3-12, March 1999.
- **Murray Campbell:** Knowledge Discovery in Deep Blue. *Communications of the ACM* 42(11): 65-67, 1999.
- **Johannes Fürnkranz.** Machine Learning in Game Playing: A Survey. In J. Fürnkranz and M. Kubat (eds.), *Machines that Learn to Play Games*, pp.11-59, Nova Science Publishers, 2001.

Exkurs: Wie entwickeln Schachgroßmeister ihre Eröffnungsstrategie?

- Typischerweise werden wenige Züge intensiv durchdacht. (Sehr Tiefe Suche)
Auswahl dieser Züge mit:
 - Intuition
 - Erfahrungen in vergangenen Spielen
 - Fachliteratur

Exkurs (Crafty): Verschiedene Verläufe der Heuristik in den ersten 10 Zügen nach dem Opening Book

- Konstanter Abstieg/Anstieg der Evaluierung => letzter (schlechtester/bester) Heuristikwert wird als Bewertung für den Book-Zug gespeichert.
- Erst fallender dann stetig steigender Wert der Evaluierung (nach einem gespielten Gambit) =>