

# Othello

— Entwicklung einer KI für das Spiel —

Patrick Müller, Max Zepnik

27. November 2018

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>2</b>
<b>2</b>	<b>Grundlagen</b>	<b>3</b>
2.1	Spieltheorie . . . . .	3
2.2	Spielstrategien . . . . .	3
2.2.1	Min-Max . . . . .	4
2.2.2	Alpha-Beta Pruning . . . . .	4
2.2.3	Suboptimale Echtzeitentscheidungen . . . . .	6
<b>3</b>	<b>Othello</b>	<b>9</b>
<b>4</b>	<b>Implementierung der KI</b>	<b>10</b>
<b>5</b>	<b>Evaluierung</b>	<b>11</b>
<b>6</b>	<b>Fazit</b>	<b>12</b>

# Kapitel 1

## Einleitung

Computergegner ..

. test..

text1 ...

am Ende schreiben

auf Fazit beziehen?

# Kapitel 2

## Grundlagen

### 2.1 Spieltheorie

Einleitung fehlt

**Definition 1 (Spiel (Game))** Ein **Game** besteht aus einem Tupel der Form

$$\mathcal{G} = \langle S_0, \text{player}, \text{actions}, \text{result}, \text{terminalTest}, \text{utility} \rangle$$

$S_0$  beschreibt den Startzustand des Spiels.

PLAYER ist auf der Menge der Spieler definiert und gibt den aktuellen Spieler zurück.

ACTIONS gibt die validen Folgezustände eines gegebenen Zustands zurück.

RESULT definiert das Resultat einer durchgeführten Aktion  $a$  und in einem Zustand  $s$ .

TERMINALTEST prüft ob ein Zustand  $s$  ein Terminalzustand, also Endzustand, darstellt.

UTILITY gibt einen Zahlenwert aus den Eingabewerten  $s$  (Terminalzustand) und  $p$  (Spieler) zurück.

Positive Werte stellen einen Gewinn, negative Werte einen Verlust dar.

quelle S.162

Eine spezielle Art von Spielen sind **Nullsummenspiele**.

Definition  
States davor

**Definition 2 (Nullsummenspiele)** In einem **Nullsummenspiel** ist die Summe der utility Funktion eines Zustands über alle Spieler 0. Dies bedeutet, dass wenn ein Spieler gewinnt mindestens ein Gegenspieler verliert.

Durch den Startzustand  $S_0$  und der Funktion ACTION wird ein **Spielbaum (Game Tree)** aufgespannt.

**Definition 3 (Spielbaum (Game Tree))** Ein **Spielbaum** besteht aus einer einzigen Wurzel, welche einen bestimmten Zustand (meistens  $S_0$ ) darstellt. Die Kindknoten der Wurzel stellen die durch ACTIONS erzeugten Zustände dar. Die Kanten zwischen der Wurzel und den Kindknoten stellen jeweils die durchgeführte Aktion dar, die ausgeführt wurde um vom State  $s$  zum Kindknoten zu gelangen.

**Definition 4 (Suchbaum (Search Tree))** Ein **Suchbaum** ist ein Teil des Spielbaums.

[?]

Überleitung  
einfügen

### 2.2 Spielstrategien

Es gibt verschiedene Spielstrategien. Im Folgenden werden diese kurz erläutert und anschließend verglichen.

### 2.2.1 Min-Max

Der erste hier erläuterte Strategie ist der Min-Max Algorithmus. Dieser ist folgendermaßen definiert:

Patrick

$$MinMax(s) = \begin{cases} Utility(s); & \text{wenn TerminalTest}(s) == \text{true} \\ \max(\{a_e \text{ Actions}(s) MinMax(Result(s, a))\}); & \text{wenn Spieler am Zug} \\ \min(\{a_e \text{ Actions}(s) MinMax(Result(s, a))\}); & \text{wenn Gegner am Zug} \end{cases}$$

Der Spieler sucht den bestmöglichen Zug aus ACTIONS, der ihm einen für eine Züge einen Vorteil schafft aber gleichzeitig nur „schlechte“ Zugmöglichkeiten für den Gegner generiert. Der Gegner kann dadurch aus allen ehemals möglichen Zügen nicht den optimalen Zug spielen, da dieser in den aktuell enthaltenen Zügen nicht vorhanden ist. Er wählt aus den verfügbaren ACTIONS nach den gleichen Vorgaben seinen besten Zug aus.

Die Strategie ist einer Tiefensuche und erkundet jeden Knoten zuerst bis zu den einzelnen Blättern bevor ein Nachbarknoten ausgewählt wird. Dies setzt das mindestens einmalige Durchlaufen des gesamten Search Trees voraus. Bei einem durchschnittlichen Verzweigungsfaktor von  $f$  bei einer Tiefe von  $d$  resultiert daraus eine Komplexität von  $O(d^f)$ . Bei einem einmaligen Erkunden der Knoten können die Werte aus den Blättern rekursiv von den Blättern zu den Knoten aktualisiert werden. Dadurch muss im nächsten Zug nur das Minimum aus ACTIONS ermittelt werden, da alle Kindknoten schon evaluiert wurden. Für übliche Spiele kann die Min-Max-Strategie allerdings nicht verwendet werden, da die Komplexität zu hoch für eine akzeptable Antwortzeit ist und der benötigte Speicherplatz für die berechneten Zustände sehr schnell wächst.

### 2.2.2 Alpha-Beta Pruning

Der Min-Max Algorithmus berechnet nach dem Prinzip „depth-first“ stets den kompletten Game Tree. Bei der Betrachtung des Entscheidungsverhaltens des Algorithmus fällt jedoch schnell auf, dass ein nicht unerheblicher Teil aller möglichen Züge gar nicht erst in Betracht gezogen wird. Dies geschieht aufgrund der Tatsache, dass diese Züge in einem schlechteren Ergebnis resultieren würden als die letztendlich ausgewählten.

Max

Dem Alpha-Beta Pruning Algorithmus liegt der Gedanke zugrunde, dass die Zustände, die in einem realen Spiel nie auftreten würden auch nicht berechnet werden müssen. Damit steht die dafür regulär erforderliche Rechenzeit und der entsprechende Speicher dafür zur Verfügung andere, vielversprechendere Zweige zu verfolgen.

#### Demonstration an einem Beispiel

Um den Algorithmus zu verdeutlichen betrachten wir das, an XXX angelehnte, folgende Beispiel. Das Dargestellte Spiel besteht aus lediglich zwei Zügen, die abwechselnd durch die Spieler gewählt werden. An den Knoten der untersten Ebene des Game Tree werden die Werte der Zustände gemäß der UTILITY Funktion angegeben. Die Werte  $\alpha$  und  $\beta$  geben den schlechtmöglichen bzw. den bestmöglichen Spielausgang für einen Zweig, immer aus der Sicht des beginnenden Spielers, an. Die ausgegrauten Knoten wurden noch nicht betrachtet.

Quelle Norvig  
Hinzufügen,  
Verweis auf  
Abbildung  
einfügen

Betrachten wir nun den linken Baum in der zweiten Zeile: Der Algorithmus beginnt damit alle möglichen Folgezustände bei der Wahl von B als Folgezustand zu evaluieren. Dabei wird zunächst der Knoten E betrachtet und damit der Wert 5 ermittelt. Dies ist der Bisher beste Wert. Er wird als  $\beta$  gespeichert. Eine Aussage über den schlechtesten Wert kann noch nicht getroffen werden.

Warum

Im nachfolgenden Game Tree wird der nächste Schritt verdeutlicht. Es wird der Knoten F betrachtet. Dieser hat einen Wert von 13. Am Zuge ist jedoch der zweite Spieler. Dieser wird, geht man davon aus, dass er ideal spielt, jedoch keinen Zug wählen der ein besseres Ergebnis für den Gegner bringt als unbedingt nötig. Der bestmögliche Wert für den ersten Spieler bleibt damit 5.

Nach der Auswertung des Knotens G steht fest, dass es keinen besseren und keinen schlechteren Wert aus Sicht des ersten Spielers gibt. Daraufhin wird die 5 auch als schlechtester Wert in  $\alpha$  gespeichert. Ausgehend von A ist der schlechteste Wert damit 5 ggf. kann jedoch noch ein besseres Ergebnis herbeigeführt werden.  $\alpha$  wird entsprechend gesetzt und  $\beta$  verbleibt undefiniert.

Nun werden die Kindknoten von C betrachtet. Mit einem Wert von 3 wäre der Knoten H das bisher beste Ergebnis für die Wahl von C. Der Wert wird entsprechend gespeichert. Würde C gewählt gäbe man dem Gegenspieler die Chance ein im Vergleich zu der Wahl des Knotens B schlechteres Ergebnis herbeizuführen. Da Ziel des Spielers jedoch ist die eigenen Punkte zu maximieren gilt es diese Chance gar nicht erst zu gewähren. Entsprechend werden die Auswertung der weiteren Knoten abgebrochen.

Der Kindknoten K des Knotens D ist mit einem Wert von 42 vielversprechend und wird in  $\beta$  gespeichert. Da dieser Wert größer ist als die gespeicherten 5 wird auch der entsprechende Wert von A aktualisiert. Der anschließend ausgewertete Knoten L ermöglicht nun ein schlechteres Ergebnis von 6  $\beta$  muss also aktualisiert werden. Der Knoten M liefert schließlich den schlechtesten Wert von 1. Da der Gegenspieler im Zweifel diesen Wert wählen würde bleibt der bisher beste Wert das Ergebnis in E. In A wird der Spieler daher B auswählen

Dieses einfache Beispiel zeigt bereits recht gut wie die Auswertung von weiteren Zweigen vermieden werden kann. In der Praktischen Anwendung befinden sich die wegfallenden Zustände häufig nicht nur in den Blättern des Baumes sondern auch auf höheren Ebenen. Der eingesparte Aufwand wird dadurch häufig noch größer.

## Implementierung

Nachfolgend wird eine Pseudoimplementierung des um Alpha-Beta Pruning erweiterten MinMax Algorithmus angegeben.:

```

global Suchtiefe
int minMax(Spiel AktuellerZustand, int Spieler, int Tiefe, int alpha, int beta) {
    if (Tiefe == 0) {
        return Utility(AktuellerZustand, Spieler);
    }
    int bisherigerMaximalWert = alpha
    Zuege = mengeDerFolgezuege(aktuellerZustand);
    for (Zug z in Zuege) {
        Spiel NeuerZustand = äwähleZug(Aktueller Zustand, z)
        wert = -minMax(NeuerZustand, anderer(Spieler), Tiefe-1, -beta, -bisherigerMaximalWert)
        if (wert > bisherigerMaximalWert) {
            bisherigerMaximalWert = wert
            if (bisherigerMaximalWert >= beta) {
                break;
            }
            if (Tiefe = Suchtiefe) {
                speichereZug(z)
            }
        }
    }
    return bisherigerMaximalwert;
}

```

Es handelt sich um eine rekursive Implementierung. Im Basisfall ist der Gametree bereits bis in die angegebene Suchtiefe erforscht. In diesem Fall wird der Wert der Utility Funktion für den aktuellen Spieler bei dem aktuellen

Zustand zurückgegeben.

Handelt es sich nicht um einen solchen Fall werden alle möglichen Folgezüge betrachtet. Um den Wert des Zuges zu bestimmen wird rekursiv die minMax-Methode erneut aufgerufen. Dabei wird entsprechend der Neue Zustand, der andere Spieler und eine um die um eins verringerte Tiefe übergeben. Der beste Wert für den anderen Spieler ist der schlechteste wert für den ersten Spieler. Daher wird der bisherige Wert von beta als alpha übergeben. Der bisher beste Wert ist aus sicht des anderen Spielers der schlechteste daher wird dieser als neues beta Übergeben. Da die Utility Funktion so implementiert ist, dass die Summe der Wertigkeiten eines Zustandes Null ergibt muss noch das Vorzeichen geändert werden.

Ist der neue Wert größer als der bisherige Maximalwert so wird dieser aktualisiert. Da der zweite Spieler versucht die Punktzahl des Gegners zu maximieren bricht dieser die Auswertung aller Zweige ab bei der ein Ergebnis, welches besser ist als das bisher schlechteste Ergebnis, möglich wird. Abschließend wird der ausgewertete Zug gespeichert um ihn später ausführen zu können.

### Ordnung der Züge

Wie in obigen Beispiel an den Zweigen unter dem Knoten C zu sehen war kann, je nach der Reihenfolge in der die Folgezüge untersucht werden, die Auswertung eines Folgezustandes früher oder später abgebrochen werden. Optimalerweise werden die besten Züge, also jene Züge die einen möglichst frühen Abbruch der Betrachtung eines Knotens herbeiführen zuerst betrachtet. Um dies Abschätzen zu können bedient man sich in der Praxis einer Heuristik die Aussagen über die Güte eines Zuges im Vergleich zu den übrigen Zügen zulässt. Anhand dieser Heuristik kann dann die Reihenfolge der Auswertung einzelner Folgezustände dynamisch angepasst werden.

### 2.2.3 Suboptimale Echtzeitentscheidungen

Selbst die gezeigten Verbesserung des MinMax-Algorithmus besitzt noch einen wesentlichen Nachteil. Da es sich um einen "depth-first-Algorithmus handelt muss jeder Pfad bis zu einem Endzustand betrachtet werden um eine Aussage über den Wert des Zuges treffen zu können. Dem steht jedoch die Tatsache entgegen, dass in der Praxis eine Entscheidung möglichst schnell, idealer Weise innerhalb weniger Minuten, getroffen werden soll. Hinzu kommt, dass je nach der verwendeten Datenstruktur für ein Spiel bei entsprechend hohem Verzweigungsfaktor und einer großen Anzahl von Zügen der Hauptspeicher eines handelsüblichen Computers nicht mehr ausreicht um diese zu fassen

Es gilt also eine Möglichkeit zu finden, die Auswertung des kompletten Baumes zu vermeiden.

#### Heuristiken

Dieses Problem lösen sogenannte Heuristiken. Dabei handelt es sich um eine Funktion die den Wert eines Spielzustandes annähert.

Die Nutzung der Heuristik wird vereinfacht, wenn Sie so definiert ist, dass sie, sofern es sich um einen Endzustand handelt den Wert der Utility Funktion zurückgibt. Der Vorteil dieses Verhaltens wird im nächsten Abschnitt betrachtet.

Kommt eine Heuristik zur Anwendung so ist die Genauigkeit mit der diese den tatsächlichen Wert approximiert der wesentliche Aspekt der die Qualität des Spiel-Algorithmus ausmacht. Um zu verhindern, das versehentlich die besten Züge nicht betrachtet werden, ist es essentiell, dass eine Heuristik den tatsächlichen Wert eines Zustandes

Max

max

consistent/admiss

nie überschätzt. Das unterschätzen des Wertes hingegen ist möglich darf im Sinne der Genauigkeit der Heuristik jedoch nicht allzu ungleichmäßig Auftreten.

### Abschnittskriterium der Suche

Gibt die Heuristik im Falle eines Endzustandes den Wert der Utility Funktion zurück so kann die oben gezeigte Implementierung so angepasst werden, dass statt der Utility Funktion einfach die Heuristik ausgewertet wird. Dadurch muss nicht mehr der Vollständige Zweig durchsucht werden und das Abbrechen nach einer gewissen Suchtiefe wird möglich.

max

### Forward pruning

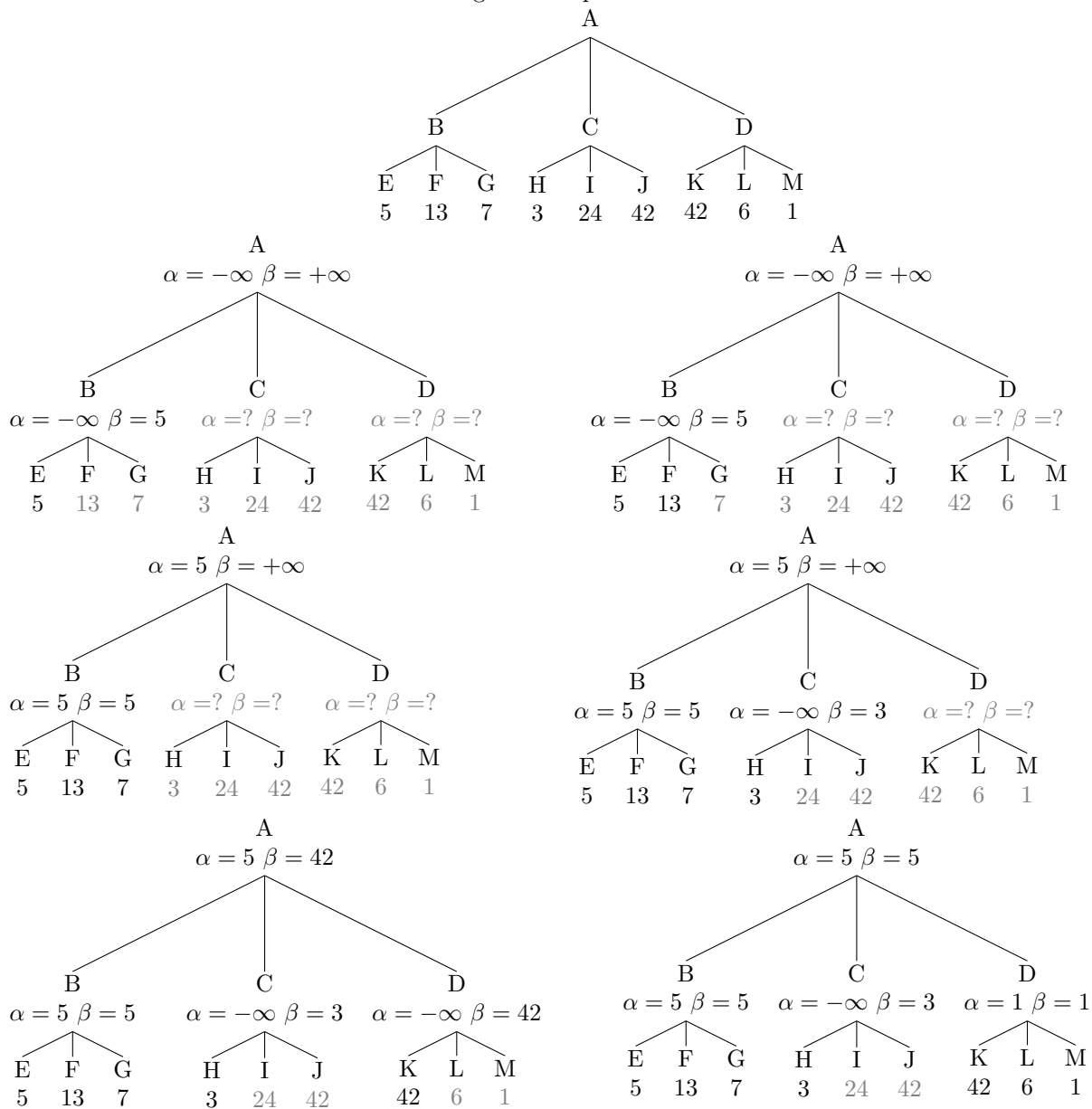
patrick

### Search versus lookup

patrick



Abbildung 2.1: Beispielhafter Game Tree



## Kapitel 3

# Othello

---

Othello erklären

---

Algo erklären

## Kapitel 4

# Implementierung der KI

## Kapitel 5

# Evaluierung

## Kapitel 6

## Fazit

# Notes

am Ende schreiben	2
auf Fazit beziehen?	2
Einleitung fehlt	3
quelle S.162	3
Definition States davor	3
Überleitung einfügen	3
Patrick	4
Max	4
Quelle Norvig Hinzufügen, Verweis auf Abbildung einfügen	4
Warum	4
Max	6
max	6
consistent/admissible	6
max	7
patrick	7
patrick	7
Othello erklären	9
Algo erklären	9

# Literaturverzeichnis