

# Othello

— Entwicklung einer KI für das Spiel —

Patrick Müller, Max Zepnik

27. März 2019

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Othello</b>	<b>2</b>
2.1	Spielregeln . . . . .	3
2.2	Spielverlauf . . . . .	3
2.3	Spielstrategien . . . . .	4
2.4	Eröffnungszüge . . . . .	4
<b>3</b>	<b>Grundlagen</b>	<b>6</b>
3.1	Spieltheorie . . . . .	6
3.2	Spielstrategien . . . . .	7
3.2.1	MiniMax . . . . .	8
3.2.2	Alpha-Beta Abscheiden . . . . .	8
3.2.3	Suboptimale Echtzeitentscheidungen . . . . .	11
3.3	Monte Carlo Algorithmus . . . . .	13
3.3.1	Algorithmus . . . . .	13
3.3.2	Überlegungen zu Othello . . . . .	13
<b>4</b>	<b>Implementierung der KI</b>	<b>15</b>
4.1	Grundlegende Spiel-Elemente . . . . .	15
4.1.1	Der Spielablauf . . . . .	15
4.1.2	Die Klasse „Othello“ . . . . .	16
4.1.3	Die übrigen Komponenten . . . . .	19
4.2	Heuristiken . . . . .	19
4.2.1	Nijssen 2007 Heuristik . . . . .	20
4.2.2	Stored Monte-Carlo-Heuristik . . . . .	20
4.2.3	Cowthello Heuristik . . . . .	23
4.3	Start Tabellen . . . . .	24
4.4	Agenten . . . . .	25
4.4.1	Human Agent . . . . .	25
4.4.2	Random . . . . .	26
4.4.3	Monte Carlo . . . . .	26
4.4.4	Alpha-Beta Pruning . . . . .	29

<b>5</b>	<b>Evaluierung</b>	<b>32</b>
5.1	Random vs. Monte Carlo . . . . .	32
5.2	Monte Carlo vs Random . . . . .	33
5.3	Random vs Alpha-Beta . . . . .	33
5.4	Alpha-Beta vs Random . . . . .	33
5.5	Alpha-Beta: Nijssen 2007 Heuristik vs Random . . . . .	33
5.6	Alpha-Beta: Cowthello Heuristik vs Random . . . . .	33
5.7	Random vs Alpha-Beta: Nijssen 2007 Heuristik . . . . .	33
5.8	Random vs Alpha-Beta: Cowthello Heuristik . . . . .	33
5.9	Alpha-Beta: Nijssen 2007 Heuristik vs Monte Carlo . . . . .	33
5.10	Alpha-Beta: Cowthello Heuristik vs Monte Carlo . . . . .	33
5.11	Monte Carlo vs Alpha-Beta: Nijssen 2007 Heuristik . . . . .	33
5.12	Monte Carlo vs Alpha-Beta: Cowthello Heuristik . . . . .	33
<b>6</b>	<b>Fazit</b>	<b>34</b>

# Kapitel 1

## Einleitung

Computergegner ..

. test..

text1 ...

am Ende schreiben

auf Fazit beziehen?

# Kapitel 2

## Othello

Othello wird auf einem 8x8 Spielbrett mit zwei Spielern gespielt. Es gibt je 64 Spielsteine, welche auf einer Seite schwarz, auf der anderen weiß sind. Der Startzustand besteht aus einem leeren Spielbrett, in welchem sich in der Mitte ein 2x2 Quadrat aus abwechselnd weißen und schwarzen Steinen befindet. Anschließend beginnt der Spieler mit den schwarzen Steinen.

Die Spielfelder werden in verschiedene Kategorien eingeteilt (siehe Abbildung 2.1):

- Randfelder: äußere Felder (blaue Felder) [o.V15]
- C-Felder: Felder, welche ein Feld horizontal oder vertikal von den Ecken entfernt sind (vgl. [Ber])
- X-Felder: Felder, welche ein Feld diagonal von den Ecken entfernt sind [o.V15]
- Zentrum: innerste Felder von C3 bis F6 (grüne Felder) [o.V15]
- Zentralfelder: Felder D4 bis E5 [o.V15]
- Frontsteine: die äußersten Steine auf dem Spielbrett um das Zentrum (vgl. [Ort]).

Diese Kategorien sind für die spätere Strategie wichtig.

	A	B	C	D	E	F	G	H
1		C					C	
2	C	X					X	C
3								
4				W	S			
5				S	W			
6								
7	C	X					X	C
8		C					C	

Abbildung 2.1: Kategorien des Spielfeldes

Von Othello gibt verschiedene Varianten. Eine Variante ist Reversi. Die verschiedenen Varianten sind allerdings bis auf die Startposition gleich. Bei der Variante Reversi sind die Zentralfelder noch nicht besetzt und die Spieler setzen die vier Steine selbst, während bei Othello die Startaufstellung fest vorgegeben ist.

## 2.1 Spielregeln

Othello besitzt einfache Spielregeln, welche im Spielverlauf aber auch taktisches oder strategisches Geschick erfordern. Jeder Spieler legt abwechselnd einen Stein auf das Spielbrett. Dabei sind folgende Spielregeln zu beachten welche auch in Abbildung 2.2 abgebildet sind:

- Ein Stein darf nur in ein leeres Feld gelegt werden.
- Es dürfen nur Steine auf Felder gelegt werden, welche einen oder mehrere gegnerischen Steine mit einem bestehenden Stein umschließen würden. Dies ist im linken Spielbrett durch grüne Felder und im mittleren Spielbrett durch das gelbe Feld hervorgehoben. Das gelbe Feld (F5) umschließt mit dem Feld D5 (blau) einen gegnerischen Stein. Es können auch mehrere Steine umschlossen werden. Allerdings dürfen sich dazwischen keine leeren Felder befinden.
- Von dem neu gesetzten Stein in alle Richtungen ausgehend werden die umschlossenen gegnerischen Steine umgedreht, sodass alle Steine die eigene Farbe besitzen. In dem Beispiel ist das im dem rechten Spielbrett zu sehen. F5 umschließt dabei das Feld E5 (rot). Dieses Feld wird nun gedreht und wird schwarz.
- Ist für einen Spieler kein Zug möglich muss dieser aussetzen. Ein Spieler darf allerdings nicht freiwillig aussetzen wenn noch mindestens eine Zugmöglichkeit besteht.
- Ist für beide Spieler kein Zug mehr möglich, ist das Spiel beendet. Der Spieler mit den meisten Steinen seiner Farbe gewinnt das Spiel.
- Das Spiel endet auch wenn alle Felder des Spielbrettes besetzt sind. In diesem Fall gewinnt ebenfalls der Spieler mit den meisten Steinen seiner Farbe.

	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5								
6								
7								
8								

	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5								
6								
7								
8								

	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5								
6								
7								
8								

Abbildung 2.2: valide Zugmöglichkeiten für Schwarz und ausgeführter Zug

## 2.2 Spielverlauf

Das Spiel wird in drei Abschnitte eingeteilt [Ort]:

- Eröffnungsphase
- Mittelspiel
- Endspiel

Diese Abschnitte sind jeweils 20 Spielzüge lang. Im Eröffnungs- und Endspiel stehen zum Mittelspiel wenige Zugmöglichkeiten zur Verfügung, da entweder nur wenige Steine auf dem Spielbrett existieren oder das Spielbrett fast gefüllt ist und nur noch einzelne Lücken übrig sind. Im Mittelspiel existieren sehr viele Möglichkeiten, da sich schon mindestens 20 Steine auf dem Spielbrett befinden und diese sehr gute Anlegemöglichkeiten bieten.

## 2.3 Spielstrategien

Wie in anderen Spielen gibt es auch in Othello verschiedene Strategien. Dabei kann beispielsweise offensiv gespielt werden, indem versucht wird möglichst viele Steine in einem Zug zu drehen. Es gibt auch defensive „stille“ Züge. Ein „stiller“ Zug dreht keinen Frontstein um und dreht möglichst nur wenige innere Steine um (vgl. [Ort]).

Generell ist eine häufig genutzte Strategie die eigene Mobilität zu erhöhen und die Mobilität des Gegners zu verringern. Mit dem Begriff Mobilität sind die möglichen Zugmöglichkeiten gemeint. Durch das Einschränken der gegnerischen Mobilität hat dieser weniger Zugmöglichkeiten und muss so ggf. strategisch schlechtere Züge durchführen.

Die Position der Steine auf dem Spielbrett sollte ebenfalls nicht vernachlässigt werden. beispielsweise sollen Züge auf X-Felder vermieden werden, da der Gegner dadurch Zugang zu den Ecken bekommt. Dadurch können ggf. die beiden Ränder und die Diagonale gedreht werden und in den Besitz des Gegners gelangen.

In der Eröffnungsphase sollten die Randfelder ebenfalls vermieden werden, da diese in dieser frühen Phase des Spiels noch gedreht werden können und der taktische Vorteil in einen strategischen Nachteil umgewandelt wird.

## 2.4 Eröffnungszüge

In der nachfolgenden Tabelle 2.1 sind verschiedene Spieleröffnungen und deren Häufigkeit in Spielen aufgelistet. Spielzüge werden in Othello durch eine Angabe der Position, auf welche der Stein gesetzt wird, dargestellt. Ein vollständiges Spiel lässt sich deshalb in einer Reihe von maximal 60 Positionen darstellen.

Name	Häufigkeit	Spielzüge
Tiger	47%	F5 D6 C3 D3 C4
Rose	13%	F5 D6 C5 F4 E3 C6 D3 F6 E6 D7
Buffalo	8%	F5 F6 E6 F4 C3
Heath	6%	F5 F6 E6 F4 G5
Inoue	5%	F5 D6 C5 F4 E3 C6 E6
Shaman	3%	F5 D6 C5 F4 E3 C6 F3

Tabelle 2.1: Liste von Othelloeröffnungen [Ort]

[Ort] gibt folgende weitere Tipps für Eröffnungen:

- Versuche weniger Steinchen zu haben als dein Gegner.
- Versuche das Zentrum zu besetzen.

- Vermeide zu viele Frontsteine umzudrehen.
- Versuche eigene Steine in einem Haufen zu sammeln statt diese zu verstreuen.
- Vermeide vor dem Mittelspiel auf die Kantfelder zu setzen.

Viele dieser Tipps können auch im späteren Spielverlauf verwendet werden.



# Kapitel 3

## Grundlagen

In diesem Kapitel werden die theoretischen Grundlagen für eine Künstliche Intelligenz für das Spiel Othello erläutert.

Zunächst werden dazu einige Begriffe der Spieltheorie eingeführt. Diese werden dann zur Beschreibung einiger deterministischer Algorithmen zum treffen einer Spielentscheidung verwendet. Den Abschluss des Kapitels bildet schließlich die Beschreibung einer stochastische Methode für diesen Zweck.

### 3.1 Spieltheorie

In dem folgenden Unterkapitel werden grundlegende Definitionen eingeführt. Diese sind an [RN16] angelehnt.

**Definition 1 (Spiel (Game))**(vgl. [RN16] S. 162)) Ein Spiel besteht aus einem Tupel der Form

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

Definiere  $Q$  als die Menge aller Zustände im Spiel. Ein Spielzustand besteht aus dem aktuellen Spielbrett, der Zugnummer, dem aktiven Spieler und weiteren Parametern, welche zur genauen Darstellung eines Spielzustands führen.

- $Q$  : Menge aller Spielzustände (*states*)
- $S_0 \in Q$  beschreibt den Startzustand des Spiels.
- $\text{Players}$  : Menge aller Spieler:  $\text{Players} = \{S, W\}$
- $\text{player} : Q \rightarrow \text{Players}$  ist auf der Menge der Spielzustände definiert und gibt den aktuellen Spieler  $p$  zurück.
- $\text{actions} : Q \rightarrow 2^{\text{moves}}$  gibt die Menge der validen Zugmöglichkeiten eines gegebenen Zustands zurück.  
 $\text{moves}$  ist die Menge aller Zugmöglichkeiten.  
 $\text{moves} = \{ \langle X, Y \rangle \mid X \in \{0..7\} \wedge Y \in \{0..7\} \}$   
Diese geben die Koordinate der Zugposition des neuen Steines auf dem Spielbrett an.
- $\text{result} : Q \times \text{actions} \rightarrow Q$  definiert das Resultat einer durchgeführten Aktion  $a$  und in einem Zustand  $s$ .

- `terminalTest` :  $Q \rightarrow \mathbb{B}$  prüft ob ein Zustand  $s$  ein Terminalzustand ( $s \in \text{terminalStates}$ ), also Endzustand, darstellt.

$$\text{terminalTest}(s) = \begin{cases} \text{True} & | s \in \text{terminalStates} \\ \text{False} & | s \notin \text{terminalStates} \end{cases}$$

- `utility` :  $\text{terminalStates} \times \text{Players} \rightarrow \{-1, 0, 1\}$  gibt einen Zahlenwert aus den Eingabenwerten  $s$  (Terminalzustand) und  $p$  (Spieler) zurück.  
Positive Werte stellen einen Gewinn, negative Werte einen Verlust dar. „0“ stellt ein Unentschieden dar.

Eine spezielle Art von Spielen sind [Nullsummenspiele](#).

**Definition 2 (Nullsummenspiele (vgl. [RN16] S. 161))** In einem [Nullsummenspiel](#) ist die Summe der utility Funktion eines Endzustandes (`terminalState`) über alle Spieler 0. Es gilt also:

$$\forall s \in \text{terminalStates} : \sum_{p \in \text{Players}} \text{utility}(s, p) = 0$$

In Othello spielen zwei Spieler gegeneinander. Es gibt also nur die drei Möglichkeiten:

- Weiß gewinnt, Schwarz verliert
- Schwarz gewinnt, Weiß verliert
- Unentschieden

Durch den Startzustand  $S_0$  und der Funktion `actions` und `result` wird ein [Spielbaum \(Game Tree\)](#) aufgespannt.

**Definition 3 (Spielbaum (Game Tree)(vgl. [RN16] S. 162))** Ein [Spielbaum](#) besteht aus einer Wurzel, welche einen bestimmten Zustand (Startzustand  $S_0$ ) 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 Zustand  $s$  zum Kindknoten  $s'$  zu gelangen. Diese Kindknoten können wiederum weitere Knoten enthalten oder ein Endpunkt des Baumes (Blatt = Spielende) darstellen. Die mathematische Definition wird rekursiv durchgeführt:

`Spielbaum` :  $Q \times \text{List}(\text{actions}) \times \text{List}(\text{Spielbaum})$

`Node(S, M, T)`  $\in$  `Spielbaum` g.d.w:

- $S \in Q \cup \{\Omega\}$
- $M = [m_1, \dots, m_n] \in \text{List}(\text{actions})$ , wobei gilt:  $\forall i \in \{1..n\} : m_i \in \text{actions}(S)$
- $T = [t_1, \dots, t_n] \in \text{List}(\text{Spielbaum})$ , wobei gilt:  $\forall i \in \{1..n\} : t_i = \text{result}(S, m_i)$

**Definition 4 (Suchbaum (Search Tree)(vgl. [RN16] S. 163))** Ein [Suchbaum](#) ist ein Teil des Spielbaums. Die Wurzel des Spielbaums besteht aus dem aktuellen Spielzustand. Alle Kindknoten und Blätter entsprechen dem Spielbaum beginnend ab dem aktuellen Zustand. Es gilt also:

$$B \in A \wedge B = A[s] \text{ mit } A \in \text{Spielbaum}, B \in \text{Suchbaum}(s), s \in Q$$

umformulieren ?

Überleitung  
einfügen

## 3.2 Spielstrategien

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

### 3.2.1 MiniMax

Die erste hier erläuterte Strategie ist der MiniMax Algorithmus. Dieser ist folgendermaßen definiert (siehe [RN16] S.164):

$$MiniMax(s, p) = \begin{cases} Utility(s, p); & \text{wenn TerminalTest}(s) \\ \max(\{MiniMax(result(s, a) | a \in actions(s)\}, (p + 1) \% 2); & \text{wenn Spieler am Zug} \\ \min(\{MiniMax(result(s, a) | a \in actions(s)\}, (p + 1) \% 2); & \text{wenn Gegner am Zug} \end{cases}$$

$getBestMove : Q \times \{-1, 0, 1\} \rightarrow \mathbf{actions}$

$getBestMove(s, score, p) = \{a | a \in \mathbf{actions} \wedge MiniMax(result(s, a), (p + 1) \% 2) == score\}.any()$

Die Funktion `getBestMove` ermittelt eine Menge aller Züge, welche den von `MiniMax` zurückgegeben Wert besitzen und wählt aus dieser Menge einen Zug aus.

Der Spieler sucht durch diese Funktion den bestmöglichen Zug aus den verfügbaren Zügen (`actions`), der ihm einen für seine 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 eine 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  $\mathcal{O}(d^f)$ . 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 MiniMax-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.

### 3.2.2 Alpha-Beta Abscheiden

Der MiniMax Algorithmus berechnet nach dem Prinzip „depth-first“ stets den kompletten Spielbaum.

Bei der Betrachtung des Entscheidungsverhaltens des Algorithmus fällt jedoch schnell auf, dass ein nicht unerheblicher Teil aller möglichen Züge durch einen menschlichen Spieler gar nicht erst in Betracht gezogen wird. Dies geschieht aufgrund der Tatsache, dass diese Züge in einem schlechteren Ergebnis resultieren würden als ein anderer, letztendlich ausgewählter, Zug.

Dem Alpha-Beta Abscheiden (Alpha-Beta Pruning) Algorithmus liegt der Gedanke zugrunde, dass die Zustände, die in einem realen Spiel nie ausgewählt 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 [RN16] 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 Spielbaum werden die Werte der Zustände gemäß der `utility` Funktion angegeben. Die Werte  $\alpha$  und  $\beta$  geben den schlecht möglichsten bzw. den bestmöglichen Spielausgang für einen Zweig, immer aus der Sicht des beginnenden Spielers, an. Die ausgegrauten Knoten wurden noch nicht betrachtet.

Abbildung 3.1: Beispielhafter Spielbaum



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.

Im nachfolgenden Spielbaum 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

Warum

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 einer Invariante des Alpha-Beta Abschneiden Algorithmus angegeben (siehe Listing 3.1). Es handelt sich um eine angepasste Version von [Str19]

Listing 3.1: Pseudoimplementierung von Alpha-Beta Abschneiden

```

1  alphaBeta(State, player, alpha = -1, beta = 1) {
2      if (finished(State)) {
3          return utility(State, player)
4      }
5      val := alpha
6      for (ns in nextStates(State, player)) {
7          val = max({ val, -alphaBeta(ns, other(player), -beta, -alpha) })
8          if (val >= beta) {
9              return val
10         }
11         alpha = max({ val, alpha })
12     }
13     return val
14 }
```

Bei der angegebenen Implementierung handelt es sich um eine rekursive Umsetzung. Nachfolgend sei das Programm erleutert:

1. Im Basisfall wurde bereits ein Blatt des Spielbaumes erreicht. Damit ist das Spiel bereits beendet. In diesem Fall kann mit *utility* der Wert des Zustands *State* für den entsprechenden Spieler *player* zurückgegeben werden.
2. In der Variable *val* wird der maximale Wert aller von *State* erreichbaren Zustände, sofern *player* einen Zug ausführt, gespeichert.

Da der Algorithmus per Definition alle Wertigkeiten kleiner  $\alpha$  ausschließen soll kann die Variable mit  $\alpha$  initialisiert werden.

3. Nun wird über alle Folgezustände  $ns$  aus der Menge  $nextStates(State, player)$  iteriert.
4. Nun wird rekursiv jeder Zustand  $ns$  ausgewertet. An dieser Stelle ist jedoch der andere Spieler an der Reihe. Entsprechend erfolgt dies für den anderen Spieler. Da es sich um ein Nullsummenspiel handelt ist der Wert eines Zustandes aus Sicht des Gegners von  $player$  genau der negative Wert der Wertigkeit für  $player$ . Aus diesem Grund müssen die Rollen von  $\alpha$  und  $\beta$  vertauscht und außerdem die Vorzeichen invertiert werden.
5. Da laut der Spezifikation des Algorithmus nur die Wertigkeit von Zuständen berechnet werden sollen in denen diese kleiner oder gleich  $\beta$  ist, wird die Auswertung aller Folgezustände mit einem  $val$  der größer oder gleich  $\beta$  ist abgebrochen. In diesem Fall wird  $val$  zurückgegeben.
6. Sobald ein Folgezustand mit einem größeren Wert als  $\alpha$  gefunden wurde kann  $\alpha$  auf den entsprechenden Wert erhöht werden. Sobald klar ist, dass der Wert  $val$  erreicht werden kann, so sind Werte kleiner als  $val$  nicht mehr relevant.

### 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.

### 3.2.3 Suboptimale Echtzeitentscheidungen

Selbst die gezeigten Verbesserung des MiniMax-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 die Tatsache, dass viele Spiele unter Verwendung von derzeit erhältlicher Hardware (noch) nicht lösbar sind.

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 versucht den Wert eines Spielzustandes anhand einzelner Eigenschaften des Zustandes anzunähern. Wie in Kapitel 2.3 erläutert, hat ein Spieler der eine Ecke des Feldes besetzt in der Regel einen Vorteil. Ein solcher Zustand würde durch die Heuristik entsprechend besser bewertet werden.

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. Jedoch wird die Berechnung der

Heuristik mit steigender Genauigkeit meist komplizierter und somit auch rechnen- und damit zeitintensiver. Aus diesem Grund muss immer eine Abwägung aus Genauigkeit und Geschwindigkeit vorgenommen werden.

### 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.

### Vorwärtsabschneiden

Vorwärtsabschneiden (Forward Pruning) durchsucht nicht den kompletten Spielbaum, sondern durchsucht nur einen Teil. Eine Möglichkeit ist eine Strahlensuche, welche nur die „besten“ Züge durchsucht (vgl. [RN16] S. 175). Die Züge mit einer geringen Erfolgswahrscheinlichkeit werden abgeschnitten und nicht bis zum Blattknoten evaluiert. Durch die Wahl des jeweiligen Zuges mit der höchsten Gewinnwahrscheinlichkeit können aber auch sehr gute bzw. schlechte Züge nicht berücksichtigt werden, wenn diese eine geringe Wahrscheinlichkeit besitzen. Durch das Abschneiden von Teilen des Spielbaum wird die Suchgeschwindigkeit deutlich erhöht. Der in dem Othello-Programm „Logistello“ verwendete „Probcut“ erzielt außerdem eine Gewinnwahrscheinlichkeit von 64% gegenüber der ursprünglichen Version ohne Vorwärtsabschneiden (vgl. [RN16] S. 175).

### Suche gegen Nachschlagen

Viele Spiele kann man in 3 Haupt-Spielabschnitte einteilen:

- Eröffnungsphase
- Mittelspiel
- Endphase

In der Eröffnungsphase und in der Endphase gibt es im Vergleich zum Mittelspiel wenige Zugmöglichkeiten. Dadurch sinkt der Verzweigungsfaktor und die generelle Anzahl der Folgezustände. In diesen Phasen können die optimalen Spielzüge einfacher berechnet werden. Eine weitere Möglichkeit besteht aus dem Nachschlagen des Spielzustands aus einer Lookup-Tabelle.

Dies ist sinnvoll, da gewöhnlicherweise sehr viel Literatur über die Spieleröffnung des jeweiligen Spiels existiert. Das Mittelspiel jedoch hat zu viele Zugmöglichkeiten, um eine Tabelle der möglichen Spielzüge bis zum Spielende aufstellen zu können. In dem Kapitel 2.4 werden die bekanntesten Eröffnungsstrategien aufgelistet.

Viele Spielstrategien wie beispielsweise die MiniMax-Strategie setzen den kompletten oder wenigstens einen großen Teil des Spielbaums voraus. Dieser kann entweder berechnet werden oder aus einer Lookup-Tabelle gelesen werden. Je nach Verzweigungsfaktor der einzelnen Spielzüge kann diese allerdings sehr groß sein. Selbst im späten Spielverlauf gibt es verschiedene Spiele, welche einen großen Spielbaum besitzen.

Beispielsweise existieren für das Endspiel in Schach mit einem König, Läufer und Springer gegen einen König 3.494.568 mögliche Positionen (vgl. [RN16] S.176).

Dies sind zu viele Möglichkeiten um alle speichern zu können, da noch sehr viel mehr Endspiel-Kombinationen als diese existieren.

Anstatt die Spielzustände also zu speichern, können auch die verbleibenden Spielzustände berechnet werden.

Othello besitzt gegenüber Schach den Vorteil, dass die Anzahl der Spielzüge auf 60 bzw. 64 Züge begrenzt sind. Dadurch kann in der Endphase des Spiel ggf. der komplette verbleibende Spielbaum berechnet werden, da die Anzahl der möglichen Zugmöglichkeiten eingeschränkt wird.

Bei der Berechnung der Spielzüge sind die Suchtiefe und der Verzweigungsfaktor entscheidend für die Berechnungsdauer. Aus diesem Grund können im Mittelspiel keine MiniMax-Algorithmen bis zu den Blattknoten des Spielbaumes ausgeführt werden, da die Menge des benötigten Speicherplatzes außerhalb jeglicher Grenzen eines Arbeits- oder Gamingscomputers liegen.

## 3.3 Monte Carlo Algorithmus

Im Gegensatz zu den bisher gezeigten Algorithmen verwendet der Monte Carlo Algorithmus einen Stochastischen Ansatz um einen Zug auszuwählen. Im nachfolgenden Abschnitt wird die Funktionsweise des Monte Carlo Algorithmus erklärt. Daran angeschlossen folgen Möglichkeiten Strategische Überlegungen zum Spiel Othello einzubringen. Dabei sind die nachfolgenden Ausführungen stark angelehnt an jene von [Nij07].

### 3.3.1 Algorithmus

Als Ausgangspunkt legt der Monte Carlo Algorithmus die Menge der Züge zugrunde, die ein Spieler unter Wahrung der Spielregeln wählen kann. Diese Züge seien nachfolgend Zug-Kandidaten genannt. Enthält die Menge keine Züge, so bleibt dem Spieler nichts anderes übrig als auszusetzen. Enthält die Menge nur einen Zug, so muss der Spieler diesen ausführen. Per Definition ist dies dann der bestmögliche Zug. Enthält die Menge hingegen mindestens zwei mögliche Züge, so gilt es den besten unter ihnen auszuwählen. Um den besten Zug zu ermitteln, wird das Spiel mehrfach, die Anzahl sei  $N_P$ , bis zum Ende simuliert. Während der Simulation wird jeder mögliche Zug gleich häufig gewählt. Der Rest des simulierten Spiels wird dann zufällig zu Ende gespielt. Sobald alle Durchgänge erfolgt sind, wird das Durchschnittliche Ergebnis für jeden möglichen Zug berechnet. Dabei gibt es zwei Möglichkeiten dieses Ergebnis zu berechnen:

Wahlweise kann die Durchschnittliche Punktzahl eines Zuges oder die durchschnittliche Anzahl an gewonnenen Spielen herangezogen werden. Jener Zug, der nun das beste Ergebnis verspricht wird gespielt.

### 3.3.2 Überlegungen zu Othello

Bisher spielt der Algorithmus auf gut Glück ohne sich jeglicher Informationen des Spiels zu bedienen. In der Hoffnung das Spiel des Algorithmus zu verbessern, werden nun weitere Informationen zu Othello herangezogen. Hier sein zwei Möglichkeiten beschrieben um dies zu erreichen:

**Vorverarbeitung** Wie in entsprechenden Kapitel gezeigt, gibt es strategisch gute und strategisch eher schlechte Züge. In seiner Reinform betrachtet der Monte Carlo Algorithmus jedoch beide Arten von Zügen gleich stark. Die Idee der Methode der Vorverarbeitung besteht darin, schlechte Züge in einem Vorverarbeitungsschritt auszuschließen um diese in den Simulationen gar nicht erst zu spielen. Um zu entscheiden, welche Züge ausgeschlossen werden, werden die einzelnen Spielzustände nach Ausführung des Zuges bewertet. Dazu werden die im entsprechenden Kapitel beschriebenen Kategorien von Spielsteinen herangezogen und mit einem entsprechenden Punktwert belegt. Für die Entscheidung an sich kann nun zwischen zwei Strategien gewählt werden:

Entweder kann eine feste Anzahl, diese sei  $N_S$ , an best bewertesten Züge ausgewählt werden oder alternativ eine variable Anzahl. Dies geschieht in dem der Durchschnitt aller Bewertungen bestimmt wird und nur jene



Züge ausgewählt werden die eine gewisse Bewertung relativ zum Durchschnittswert haben. Dies wird in einer prozentualen Erfüllung des Durchschnittswertes, diese sei  $p_s$ , angegeben.

**Pseudozufällige Zugauswahl** In der Standardversion des Monte Carlo-Algorithmus werden die simulierten Spiele nach der Wahl des ersten Zuges zufällig zu Ende gespielt. Dem hier beschriebenen Ansatz liegt die Idee zu Grunde auch in dieser Phase der Simulation einige Züge anderen gegenüber zu bevorzugen. Dies geschieht nach dem gleichen Prinzip wie im Abschnitt zur Vorverarbeitung beschrieben. Da es jedoch sehr zeitaufwändig ist, die Bewertung jedes einzelnen Spielzustandes innerhalb der Simulationen vorzunehmen, erfolgt dies nur bis zu einer bestimmten Tiefe. Diese sei  $N_d$ .

# Kapitel 4

## Implementierung der KI

In den folgenden Unterkapiteln werden verschiedene Spielalgorithmen vorgestellt und implementiert. Anschließend werden diese verbessert und auch unter Berücksichtigung des Laufzeitverhaltens analysiert.

Zunächst wird aber die grundsätzliche Programmstruktur erläutert und das Spielgerüst implementiert, damit unterschiedliche Spieler es ausführen können.

### 4.1 Grundlegende Spiel-Elemente

Die Python Implementierung befindet sich im Verzeichnis „python“ des zu diesem Projekt gehörenden Git Repository. Um das Spiel zur Ausführung zu bringen werden die Pakete „numpy“ sowie „pandas“ benötigt. Sind diese Abhängigkeiten vorhanden, kann das Spiel durch das Ausführen des Kommandos „python main-game“ gestartet werden.

Die einzelnen Komponenten wurden unter thematischen Gesichtspunkten in verschiedenen Dateien organisiert. Nachfolgend wird auf die einzelnen Dateien und deren Funktion kurz eingegangen.

#### 4.1.1 Der Spielablauf

In „main-game.py“ wird das Rahmenprogramm gestartet. In diesem werden zunächst die Spieler festgelegt. Anschließend wird ein Spiel erstellt, initialisiert und das Spielbrett ausgegeben ( siehe Listing 4.1 Z. 2-4). Die Methode „game\_is\_over“ gibt bei Spielende „True“, ansonsten „False“ zurück. In der Schleife von Zeile 5 bis Zeile 11 wird der Agent des aktuellen Spielers ermittelt (Z. 6f.). Für diesen wird die Funktion „get\_move“ aufgerufen. Diese gibt den nach der Strategie des jeweiligen Agenten besten Spielzug zurück. Je nach Spielagent werden unterschiedliche Algorithmen zur Ermittlung dieses Zuges verwendet. Dieser Zug wird gespielt und auf das Spielbrett angewendet (Z. 9). Abschließend wird das aktualisierte Spielbrett und der zuletzt durchgeführte Zug ausgegeben (Z. 10f.). Die Schleife wird bis zum Spielende wiederholt. Danach wird der Gewinner und die gesamte Spieldauer ermittelt und ausgegeben (Z. 12-15).

Listing 4.1: Spielablauf in „main-game.py“

```
1  players = {PLAYER_ONE: player_one, PLAYER_TWO: player_two}
2  game = Othello()
3  game.init_game()
4  game.print_board()
5  while not game.game_is_over():
```

```

6     current_player = game.get_current_player()
7     player_object = players[current_player]
8     move = player_object.get_move(game)
9     game.play_position(move)
10    game.print_board()
11    print(f"Played position: ({COLUMN_NAMES[move[1]]}{move[0] + 1})")
12    duration = time.time() - start
13    print("Game is over")
14    print(f"Total duration: {duration} seconds")
15    print(f"Winner is {PRINT_SYMBOLS[game.get_winner()]}")

```

### 4.1.2 Die Klasse „Othello“

Die Klasse „Othello“ modelliert einen Spielzustand und enthält die Grundlegende Spiellogik wie bspw. die Berechnung erlaubter Züge. Nachfolgend wird auf die Art der Speicherung eines Spielzustandes und auf die wichtigsten Funktionen dieser Klasse eingegangen.

**Klassenvariablen** In Listing 4.2 sind alle Klassenvariablen, sowie deren Initialisierungswerte angegeben.

1. „**\_board**“: Bei „\_board“ handelt es sich um eine Liste von Listen, zur Modellierung der zweidimensionalen Struktur des Spielbretts. Initialisiert wird das Spielbrett in seinem Leerzustand. Daher wird zu Beginn jedes Feld auf „0“ zur Repräsentation des leeren Feldes gesetzt.
2. „**\_current\_player**“: Speichert den Spieler, der im modellierten Spielzustand an der Reihe ist.
3. „**\_last\_turn\_passed**“ wird verwendet um zu speichern ob der vorherige Spieler passen musste. Dadurch kann das Spiel sobald zwei Spieler unmittelbar nacheinander passen müssen beendet werden.
4. „**\_game\_is\_over**“ wird auf „True“ gesetzt sobald das Spiel beendet ist
5. „**\_fringe**“ In „\_fringe“ werden alle Felder des Spielfeldes gespeichert die in eine Richtung unmittelbar neben einem bereits besetzten Feld liegen. Durch die Mitführung dieser Information muss zur Berechnung der erlaubten Züge nicht jedes mal über das Spielfeld iteriert werden um zunächst die infrage kommenden Felder zu ermitteln.
6. „**\_turning\_stones**“: Enthält als Schlüssel alle erlaubten Züge und als Wert jeweils eine Liste jener Spielsteine die durch ausführen des Zuges umgedreht werden. Da zur Ermittlung der erlaubten Züge diese Information bereits berechnet werden muss wird sie in Form des Dictionarys vorgehalten um diese an anderer Stelle nicht erneut berechnen zu müssen.
7. „**\_taken\_moves**“: Speichert alle ausgeführten Züge die erforderlich waren um den modellierten Spielzustand zu erreichen, da einige Algorithmen diese Information benötigen.
8. „**\_turn\_nr**“: Speichert die Nummer des aktuellen Spielzuges, da die verwendete Strategie bei einigen Algorithmen davon abhängt, wie weit das Spiel schon fortgeschritten ist.

Listing 4.2: Klassenvariablen der Klasse „Othello“

```

1     _board = [[0 for _ in range(8)] for _ in range(8)]
2     _current_player = None

```

```
3
4     _last_turn_passed = False
5     _game_is_over = False
6
7     _fringe = set()
8     _turning_stones = dict()
9
10    _taken_moves = dict()
11    _turn_nr = 0
```

Die Funktion „`compute_available_moves`“ ist in Listing 4.3 angegeben und wird verwendet um die Inhalte des Dictionaries „`turning_stones`“ zu berechnern.

Da beiden Spielern in der Regel nicht die gleichen Züge zur Verfügung stehen muss zunächst der vorherige Inhalt von „`turning_stones`“ gelöscht werden. Dies geschieht in Zeile 2 indem die Datenstruktur neu initialisiert wird.

In Zeile 3 wird eine lokale Referenz des zur Darstellung eines durch den aktuellen Spieler besetzten Feldes verwendeten Symbols erzeugt.

Mit der in Zeile 4 beginnenden Schleife wird über alle in Frage kommenden Züge in der Menge „`fringe`“ iteriert um zu ermitteln, ob dieser Zug erlaubt wäre.

Dazu wird zunächst eine lokale Menge initialisiert um die durch spielen dieser Position gedrehten Steine zu speichern (Zeile 5).

Nun muss ausgehend von der derzeit betrachteten Position ermittelt werden ob in irgendeine Richtung Spielsteine gedreht werden würden. Dies erfolgt durch die in Zeile 6 beginnende Schleife.

Dazu wird zunächst das nächste Feld in diese Richtung unter Verwendung einer Hilfsfunktion ermittelt (Zeile 7) und anschließend eine weitere temporäre Menge der in dieser Richtung gedrehten Steine initialisiert (Zeile 8)

Die entsprechende Richtung muss nun solange weiter verfolgt werden wie ein weiterer Nachbar in diese Richtung vorhanden ist. Dies geschieht durch die in Zeile 9 beginnende Schleife.

Da in den nachfolgenden Schritten ermittelt werden muss welcher Spieler das derzeit betrachtete Feld besetzt hat, werden die Indizes der derzeitigen ausgepackt (Zeile 10) und dann verwendet um zu ermitteln welchen Wert das Feld derzeit hat (Zeile 11).

Nun gibt es drei mögliche Fälle:

1. Es befindet sich kein Stein auf dem derzeit betrachteten Feld. In diesem Fall wird die Abfrage in Zeile 12 positiv ausgewertet und diese Richtung muss nicht weiter verfolgt werden. Entsprechend wird die while-Schleife in Zeile 13 abgebrochen.
2. Das derzeit betrachtete Feld wird durch den anderen Spieler besetzt. In diesem Fall ist die Abfrage in Zeile 14 positiv. Da der Stein ggf. umgedreht werden würde, wird die aktuelle Position gespeichert (Zeile 15)
3. Das derzeit betrachtete Feld wird durch den Spieler selbst besetzt. In diesem Fall wird die Abfrage in Zeile 16 positiv ausgewertet. Nun werden alle Steine zwischen der Ausgangsposition und der derzeitigen gedreht. Daher wird die Menge der durch diesen Zug gedrehten Steine mit der in diese Richtung befindlichen Steine vereinigt (Zeile 17) und die Schleife verlassen (Zeile 18).

In Zeile 19 wird das nächste Feld in diese Richtung berechnet.

Gemäß der Regeln muss durch jeden Zug mindestens ein Stein gedreht werden. Aus diesem Grund wird nun ermittelt ob dies bei diesem Zug gegeben wäre (Zeile 20). Ist dies der Fall so werden die gedrehten Steine für diesen Zug in „`stones_to_turn`“ gespeichert (Zeile 21).

Hat ein Spieler nun keine Möglichkeiten einen Zug durchzuführen, so sind in „`stones_to_turn`“ keine Züge enthalten. Dieser Fall muss besonders behandelt werden. Tritt er ein, so wird die Abfrage in Zeile 22 positiv ausgewertet.

In diesem Fall muss nochmal unterschieden werden, ob der vorherige Spieler ebenfalls keinen Zug zur Auswahl hatte (Zeile 23).

Falls ja so ist das Spiel zu ende. Dies wird durch setzen von „`game_is_over`“ gespeichert (Zeile 24).

Falls nein (Zeile 25), so wird gespeichert, dass der Spieler passen musste (Zeile 26) und der nächste Zug vorbereitet (Zeile 27) Hat der Spieler hingegen eine Zugmöglichkeit (Zeile 28) so hat er aus Sicht des folgenden Zuges nicht passen müssen. Entsprechend wird „`last_turn_passed`“ wieder auf „`False`“ gesetzt.

Listing 4.3: Die Funktion „`_compute_available_moves`“

```

1  def _compute_available_moves(self):
2      self._turning_stones = dict()
3      own_symbol = self._current_player
4      for current_position in self._fringe:
5          position_turns = set()
6          for direction in DIRECTIONS:
7              next_step = Othello._next_step(current_position, direction)
8              this_direction = set()
9              while next_step is not None:
10                 (current_x, current_y) = next_step
11                 current_value = self._board[current_x][current_y]
12                 if current_value == EMPTY_CELL:
13                     break
14                 elif current_value != own_symbol:
15                     this_direction.add(next_step)
16                 elif current_value == own_symbol:
17                     position_turns = position_turns | this_direction
18                     break
19                 next_step = Othello._next_step(next_step, direction)
20             if len(position_turns) > 0:
21                 self._turning_stones[current_position] = position_turns
22         if len(self._turning_stones) == 0:
23             if self._last_turn_passed:
24                 self._game_is_over = True
25             else:
26                 self._last_turn_passed = True
27                 self._prepare_next_turn()
28         else:
29             self._last_turn_passed = False

```

**Weitere Funktionen der Klasse „Othello“** Die Klasse „Othello“ enthält weitere Funktionen auf die hier jedoch nicht im Detail eingegangen werden soll. Dennoch sei hier jeweils kurz deren Verwendungszweck der wichtigsten Funktionen genannt:

1. „`play_position`“ verändert den Spielzustand dahingehend, dass der übergebene Zug, sofern er erlaubt ist,

ausgeführt wird, die entsprechenden Steine des Gegners gedreht und dessen Zug vorbereitet wird. Dabei wird auch die „`fringe`“ entsprechend aktualisiert.

2. „`set_available_moves`“ verändert „`stones_to_turn`“ dahingehend, dass nur noch übergebene Positionen enthalten sind. Kann damit zum Filtern der erlaubten Züge verwendet werden.
3. „`get_available_moves`“: Gibt die erlaubten Züge zur Verwendung in den Spielerimplementierungen zurück
4. „`other_player`“: Gibt das Symbol des anderen Spielers zurück
5. „`utility`“: Gibt gemäß der Definition eines Spiels 0, 1 oder -1 zurück.
6. „`get_winner`“: Ermittelt den Gewinner des Spiels und gibt ihn zurück.
7. „`get_statistics`“: Gibt die Anzahl der Felder pro Spieler zurück.
8. „`get_current_player`“: Gibt den derzeitigen Spieler zurück.
9. „`game_is_over`“: Gibt zurück ob das Spiel bereits zu Ende ist.
10. „`init_game`“: Bereitet den Start eines Spiels vor indem die initial besetzten Felder entsprechen gesetzt werden. Der beginnende Spieler festgelegt und die „`fringe`“ vorbereitet, sowie „`stones_to_turn`“ für den ersten Zug berechnet.

### 4.1.3 Die übrigen Komponenten

Neben den zuvor detailliert besprochenen finden in der Implementierung noch die folgenden Komponenten Anwendung:

1. Konstanten in der Datei „`constants.py`“. Durch die Verwendung von Konstanten bspw. zur Symbolisierung von welchen Spieler ein Feld besetzt ist, wird einerseits von konkreten Werten abstrahiert und diese sind, sofern erforderlich einfach austauschbar. Andererseits wird eine gewisse Konsistenz, bspw. für Mapping-Funktionen zur Ausgabe, über das ganze Programm hinweg erreicht.
2. Hilfsfunktionen in der Datei „`util.py`“ werden dazu genutzt Werte vom Benutzer abzufragen.

## 4.2 Heuristiken

Wie im Abschnitt 3.2.3 besprochen werden für einige Algorithmen Funktionen zur Approximation des Wertes eines Spielzustandes benötigt. Im Rahmen dieser Arbeit wurden die folgenden Heuristiken implementiert:

1. Nijssen 2007 Heuristik
2. Stored Monte-Carlo-Heuristik
3. Cowthello Heuristik

Die Implementierung aller Heuristiken befinden sich in der Datei „`heuristics.py`“

### 4.2.1 Nijssen 2007 Heuristik

Die „Nijssen 2007 Heuristik“ wurde aus [Nij07] übernommen. Ihr liegt die im Kapitel 2 erläuterte Idee zugrunde, die einzelnen Spielfelder in Kategorien einzuteilen und jeder Kategorie einen speziellen Wert zuzuweisen. Die Bewertung des Spielzustandes  $s$  aus Sicht eines Spielers **player** ergibt sich dann nach der folgenden Formel:  $\text{heuristic}(\text{player}, S) = \sum_{f \in \text{domain}(F(s))} w_f * b_f$  wobei  $F(s)$  eine Funktion ist, die für das Spielfeld des Zustands  $s$  eine Relation zurückgibt die für jedes Feld angibt durch welchen Spieler es besetzt ist oder ob es leer ist,  $w_f$

$$\text{für das dem Feld zugeordnete Gewicht und } b_f = \begin{cases} 1, & \text{wenn } F(s)[f] = \text{player} \\ -1, & \text{wenn } F(s)[f] = \text{other}(\text{player}) \\ 0, & \text{sonst} \end{cases}$$

Für die Gewichte  $b_f$  der im Kapitel 2 eingeführten Kategorien vergibt Nijssen die folgenden Gewichte:

1. „Eck-Felder“: +5
2. „X-Felder“: -2
3. „C-Felder“: -1
4. „Zentral-Felder“: +2
5. „Andere-Felder“: +1

### 4.2.2 Stored Monte-Carlo-Heuristik

Die „Stored Monte-Carlo-Heuristik“ verwendet eine Datenbank. In dieser Datenbank werden die Anzahl der gespielten und gewonnenen Spielzüge gespeichert. Diese wird in den Unterabschnitten [Datenbank](#) und [Befüllen der Datenbank](#) erklärt. Darauf aufbauend wird die Funktionsweise der Heuristik in dem Abschnitt [Heuristik](#) erläutert.

**Datenbank** Die Datenbank speichert zu jeder Zugnummer die Gewinnwahrscheinlichkeiten der Spieler bei der Verwendung der Feldkategorien zu einer bestimmten Zugnummer. Sie ist in der Datei „`database_moves.csv`“ im CSV-Format gespeichert. Es existieren zehn Feldkategorien. Das Spielfeld ist symmetrisch zum Mittelpunkt aufgebaut. Dies bedeutet, dass Züge, welche symmetrisch zu anderen Zügen sind, als ein Zug angesehen werden können. In Abbildung 4.1 sind die symmetrischen Felder des Othello-Spielbrettes farblich hervorgehoben. Die unterschiedlichen Farben stellen die unterschiedlichen Feldkategorien dar. Diese sind von Null bis Acht durchnummeriert. Das Zentrum ist mit „X“ markiert. Da das Zentrum bei Spielstart schon besetzt ist, wird diese Feldkategorie nicht in der Datenbank gespeichert. Die restlichen neun Kategorien sind als Spalten in der Datenbank abgebildet.

	a	b	c	d	e	f	g	h
1	0	1	2	3	3	2	1	0
2	1	4	5	6	6	5	4	1
3	2	5	7	8	8	7	5	2
4	3	6	8	X	X	8	6	3
5	3	6	8	X	X	8	6	3
6	2	5	7	8	8	7	5	2
7	1	4	5	6	6	5	4	1
8	0	1	2	3	3	2	1	0

Abbildung 4.1: Symmetrie des Spielfeldes

Die Datenbank besteht aus 60 Zeilen und neun Spalten. Die neun Spalten stellen o.g. Feldkategorien („0“ bis „8“) dar. Die  $c$ -te Spalte in der  $n$ -ten Zeile stellt die Gewinnwahrscheinlichkeiten des Spieles dar, wenn im  $n$ -ten Spielzug ein Feld der  $c$ -ten Feldkategorie gespielt wird.

Die mathematische Formel dazu lautet:  $database_{n,c} = P(\text{win} | \text{move}(n) \in c)$ , wobei gilt:

- $\text{move}(n)$  = Feld des  $n$ -ten Spielzuges
- $c \in \{\text{Feldkategorie} - 'X'\}$  :  $c$  ist eine Feldkategorie außer das Zentrum („X“)

Jede Zelle in dieser Tabelle enthält ein Tripel der Form

`(won_games_player_1, won_games_player_2, total_played_games)`.

Die erste Komponente stellt die Anzahl der gewonnenen Spiele des ersten Spielers dar, die zweite Komponente speichert die Anzahl der Spiele, welche der zweite Spieler gewonnen hat und die dritte Komponente gibt die Gesamtanzahl der gespielten Spiele dieser Spielkategorie an. Die genaue Funktionsweise wird in Kapitel [Befüllen der Datenbank](#) erklärt.

Zwischen den drei Komponenten gibt es folgenden mathematischen Zusammenhang:

$$\text{won\_games\_player\_1} + \text{won\_games\_player\_2} \leq \text{total\_played\_games}$$

Durch unentschiedene Spiele kann die Gesamtanzahl der Spiele größer als die Summe der gewonnenen Spiele der zwei Spieler sein. In Abbildung 4.2 ist der Initialzustand der Datenbank dargestellt (vgl. auch Abbildung 4.1 bzgl. der Feldkategorien). Die Startwerte der Tupel sind jeweils „(0, 0, 0)“, da initial noch keine Spiele gespeichert sind.



Zugnummer	Feldkategorie								
	0	1	2	3	4	5	6	7	8
0	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
1	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
2	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
3	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
4	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
5	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)

Abbildung 4.2: Ausschnitt des Initialzustandes der Datenbank

**Befüllen der Datenbank** Nachdem die Datenbank im vorherigen Abschnitt initialisiert erstellt wurde, wird diese nun mit Spieldaten aus zufällig gespielten Spielen befüllt. In Listing 4.4 ist das Spielen eines zufälligen Spieles und den Aufruf der Funktion „`update_fields_stats_for_single_game`“ abgebildet. Zunächst wird ein Othello-Spiel initialisiert (Z. 1f.). Durch die Nutzung des „Random“-Agenten wird ein komplettes Spiel durchgeführt (Z. 3f.). Anschließend wird der Gewinner ermittelt (Z. 5) und die Zugreihenfolge ermittelt (Z. 6). Die Funktion „`update_fields_stats_for_single_game`“ verwendet diese beiden Parameter um die Datenbank zu aktualisieren. Die Funktion wird in Listing 4.5 abgebildet.

Listing 4.4: Befüllen der Datenbank 1

```

1 g = Othello()
2 g.init_game()
3 while not g.game_is_over():
4     g.play_position(Random.get_move(g))
5 winner = g.get_winner()
6 moves = g.get_taken_mv()
7 self.update_fields_stats_for_single_game(moves, winner)

```

In der Funktion „`update_fields_stats_for_single_game`“ (Listing 4.5 Z. 9-12) wird die Liste der Züge in einzelne Züge aufgeteilt (Z. 10), welche anschließend jeweils eine Spalte in einer Datenbankzeile aktualisieren. Dazu werden die Züge, beispielsweise „a1“, in Zeile 11 in eine Feldkategorie umgerechnet (im Beispiel „0“) und ebenfalls der Methode „`update_field_stat`“ übergeben.

Diese Methode liest die Werte der durch Zugnummer und Feldkategorie festgelegte Zelle aus (Z.2). Je nachdem, welcher Spieler dieses Spiel gewonnen hat, wird die Zahl der gewonnenen Spiele des ersten, zweiten oder keinem Spieler um eins erhöht (Z. 3-6). Abschließend wird die Zahl der insgesamt durchgeführten Spiele inkrementiert (Z.7 dritte Komponente) und in die Datenbank zurückgeschrieben (Z.7).

Listing 4.5: Befüllen der Datenbank 2

```

1 def update_field_stat(self, turn_nr, field_type, winner):
2     (won_games_pl1, won_games_pl2, total_games_played) = self._data[turn_nr][field_type]
3     if winner == PLAYER_ONE:
4         won_games_pl1 += 1
5     elif winner == PLAYER_TWO:
6         won_games_pl2 += 1
7     self._data[turn_nr][field_type] = (won_games_pl1, won_games_pl2, total_games_played + 1)
8
9 def update_fields_stats_for_single_game(self, moves, winner):
10     for turn_nr in range(len(moves)):
11         position = self.translate_position_to_database(moves[turn_nr])

```

```
12 self.update_field_stat(turn_nr, position, winner)
```

Die Datenbank wurde mit 140.000 Spielen trainiert, um statistische Abweichungen zu minimieren. Hierbei wirkt das Gesetz der großen Zahlen, das besagt, dass sich die Wahrscheinlichkeiten eines Ereignisses bei sehr vielen Wiederholungen dem Erwartungswert annähert.

Zur Auswertung der Datenbank wurde die Klasse „Analyse“ in der Datei „analyse\_database.py“ erstellt. Diese liest die Datenbank aus und liefert eine farbige Ansicht der Gewinnwahrscheinlichkeiten eines Spielers jeweils pro Zug auf dem Terminal. Neben dem dargestellten Spielbrett wird auch das Minimum, das Maximum, der Durchschnittswert und die Standardabweichung des dargestellten Spielbrettes berechnet.

**Heuristik** Mithilfe der o.g. Datenbank wird eine Heuristik implementiert. Die Methode „heuristic“ ist in Listing 4.6 abgebildet. Diese ermittelt die verfügbaren Zugmöglichkeiten und die aktuelle Zugnummer (Z. 2f.). Für jede Zugmöglichkeit wird die Gewinnwahrscheinlichkeit für den übergebenen Spieler in ein Dictionary gespeichert (Z. 6f.). Aus diesem Dictionary wird die höchste Wahrscheinlichkeit ermittelt und zurückgegeben (Z. 9f.).

Listing 4.6: Stored Monte-Carlo-Heuristik Funktion

```
1 def heuristic(current_player, game_state: Othello):
2     moves = game_state.get_available_moves()
3     turn_nr = game_state.get_turn_nr()
4     move_probability = dict()
5
6     for move in moves:
7         move_probability[move] = database.db.get_likelihood(move, turn_nr, current_player)
8
9     selected_move = max(move_probability.items(), key=operator.itemgetter(1))[0]
10    return move_probability[selected_move]
```

### 4.2.3 Cowthello Heuristik

Die „Cowthello Heuristik“ wurde aus [o.VoJ] übernommen. Ihr liegt ebenfalls der in Kapitel 2 erläuterte Idee zugrunde, die einzelnen Spielfelder in Kategorien einzuteilen und jeder Kategorie einen speziellen Wert zuzuweisen. Die Heuristik unterscheidet sich von der „Nijssen 2007 Heuristik“ nur in der genaueren Definition der Feldkategorien und unterschiedliche Gewichtungen dieser Felder. Die Gewichte der einzelnen Felder sind in Abbildung dargestellt. Die Unterschiede zur „Nijssen 2007 Heuristik“ bestehen in der detaillierteren Gewichtung beispielsweise der Zentralfelder („1“, „5“ oder „50“) im Gegensatz zu „2“. Die restliche Implementierung ist identisch zur „Nijssen 2007 Heuristik“.

	a	b	c	d	e	f	g	h
1	100	-25	25	10	10	25	-25	100
2	-25	-50	1	1	1	1	-50	-25
3	25	1	50	5	5	50	1	25
4	10	1	5	1	1	5	1	10
5	10	1	5	1	1	5	1	10
6	25	1	50	5	5	50	1	25
7	-25	-50	1	1	1	1	-50	-25
8	100	-25	25	10	10	25	-25	100

Abbildung 4.3: Gewichtung der einzelnen Spielfelder

## 4.3 Start Tabellen

In dem Grundlagenkapitel 3.2.3 wurden die Vor- und Nachteile von Suche vs. Nachschlagen beschrieben. In diesem Kapitel wird die Implementierung der Starttabellen mit Eröffnungszügen erklärt. Im Folgenden werden Spieltabellen und Startdatenbank synonym verwendet, da diese Tabellen als Matrix in einer Datenbank gespeichert werden.

Für viele Spiele, beispielsweise für Schach, existieren Starttabellen, welche die „besten“ Eröffnungszüge speichern. Für Othello existieren zwar mehrere Spieltabellen, aber nur wenige Eröffnungsspiele. In der Implementierung werden die 77 Eröffnungszüge von Robert Gatliff [o.V05] verwendet. Diese bestehen jeweils aus einer Zugreihenfolge, z.B. „c4, c3, d3, c5, b2“.

Diese Züge werden im Spiel mit dem aktuellen Spielzustand verglichen. Wenn eine oder mehrere gespeicherte Züge mit dem aktuellen Spielzustand übereinstimmen, wird ein Zug aus den verfügbaren Zügen ausgewählt und gespielt. Die Funktion „get\_available\_moves\_of\_start\_tables“ ist in Listing 4.7 abgebildet.

Listing 4.7: Befüllen der Datenbank 2

```

1 def get_available_moves_of_start_tables(self, game: Othello):
2     if len(self._start_tables) == 0:
3         self._init_start_tables()
4     turn_nr = game.get_turn_nr()
5     available_moves = []
6     taken_mv = game.get_taken_mv_text()
7     for game in self._start_tables:
8         turn = 0
9         for move in game:

```

```

10     if turn < turn_nr:
11         if taken_mv[turn] != move:
12             break
13     else: # turn == turn_nr
14         if move != "i8" or move != "nan": # invalid field
15             available_moves.append(move)
16         break
17     turn += 1
18     available_moves = list(dict.fromkeys(available_moves))
19     if "nan" in available_moves:
20         available_moves.remove("nan")
21     return available_moves

```

Beim ersten Aufrufen der Starttabellen, wird die Datenbank initialisiert (Z. 2f.). Ab Zeile sieben werden alle Zeilen der Datenbank mit der aktuellen Zugreihenfolge des Spiels verglichen (Z. 9ff.). Wenn ein Zug in der bestehenden Zugreihenfolge abweichend von dem aktuellen Eintrag der Startdatenbank ist (Z. 11f.), wird der weitere Vergleich mit diesem Eintrag abgebrochen und mit dem folgenden Eintrag fortgefahren.

Da das Spielbrett symmetrisch zum Mittelpunkt ist, wurden die Eröffnungszüge gespiegelt. Dadurch wurde die Datenbank auf 308 Züge erweitert. Diese ist in der Datei „`start_moves.csv`“ als CSV gespeichert.

Die Agenten „Monte Carlo“ und „Alpha-Beta Pruning“ verwenden in der Standardeinstellung Starttabellen. Wenn beide Agenten gegeneinander spielen, werden die ersten Spielzüge beider Spieler sehr stark beschleunigt. Die Agenten müssen keine Züge berechnen, sondern können, bei verfügbaren Eröffnungszügen, durch das Nachschlagen in der Datenbank und die Auswahl eines Spielzuges sehr viel Spielzeit einsparen. Erst wenn die Datenbank keine passende Zugmöglichkeit mehr enthält, starten die Agenten die Berechnung des besten Zuges.

## 4.4 Agenten

Beim Start eines Partie stehen dem Nutzer mehrere Agenten zur Auswahl die die Rolle eines Spielers übernehmen können. Die Implementierung zu diesen Agenten befindet sich in Unterverzeichnis „`Agents`“.

Die folgende Agenten stehen dabei zur Auswahl:

1. Human
2. Random Player
3. Monte Carlo
4. Alpha-Beta Pruning

### 4.4.1 Human Agent

Der „Human Agent“ bzw. menschliche Agent stellt eine Schnittstelle die es einem menschlichen Spieler ermöglicht eine Spielentscheidung zu treffen. Die Implementierung findet sich in der Datei „`human.py`“ Neben dem Spielfeld bekommt der Nutzer dabei eine Liste aller für Ihn möglichen Züge dargestellt. Durch die Eingabe eines Zuges wird dieser im Spielmodell ausgeführt und der nächste Agent wird aufgerufen. Da das Ziel dieser Arbeit darin besteht eine künstliche Intelligenz zur Wahl der Züge zu entwickeln, soll an dieser Stelle nicht weiter auf diesen Agenten eingegangen sein.

### 4.4.2 Random

Die Implementierung des Agenten „Random“ befindet sich in der Datei „`random.py`“. Dieser Agent ist die einfachste Form der im Rahmen dieser Arbeit eingesetzten Strategien einer künstlichen Intelligenz. Ihr liegt die Idee zugrunde, dass der Agent einen zufälligen Zug aus der Menge der erlaubten Züge auswählt.

In der mit der Arbeit entwickelten Implementierung ist dies derartig umgesetzt, dass der Agent die Liste der möglichen Züge aus dem Spielzustand abrufen und einen zufälligen Index der Liste auswählt um dann den dort angegebenen Zug zu spielen.

### 4.4.3 Monte Carlo

Hinter diesem Agenten steht das im Kapitel 3.3 erläuterte Prinzip, der zufällig gespielten Spiele zur Ermittlung des Spielzuges mit der höchsten Gewinnwahrscheinlichkeit. Dabei werden ausgehend von einem derzeitigen Spielzustand  $N$  Spiele simuliert in dem für beide Spieler der Agent „Random“ verwendet wird. Dabei wird für den jeweils ersten simulierten Zug gespeichert wie oft dieser durchgeführt wurde und wie häufig dies in einer gewonnenen Simulation resultierte. In dem nicht simulierten Spiel wird dann jener Spielzug ausgeführt, bei dem der Quotient aus der Anzahl der nach spielen dieses Zuges gewonnenen Spiele und der Anzahl der simulierten Spiele die mit diesem Zug begonnen wurden, am größten ist.

Nachfolgend wird dieses Prinzip anhand des Quellcodes nochmals erläutert.

Die Spielerklasse „MonteCarlo“ ist in der Datei „`python/Agents/monteCarlo.py`“ zu finden.

#### Parameter

Das genaue Verhalten des Agenten kann durch die folgenden Parameter beeinflusst werden:

- „`big_n`“: Die Anzahl  $N$  der zufällig gespielten Spiele je Zug
- „`use_start_libs`“: Bool'scher Wert ob Startbibliotheken verwendet werden sollen
- „`preprocessor`“: Der verwendete Vorverarbeiter bzw. Präprozessor. Dabei stehen die folgenden bereits im Kapitel 3.3 besprochenen Varianten zur Auswahl:
  - Der feste Selektivität Präprozessor
  - Der variable Selektivität Präprozessor
  - Kein Präprozessor
- „`preprocessor_parameter`“: Parameter des verwendeten Präprozessors. Je nach Auswahl des Präprozessors steht dieser Parameter für:
  - Die Anzahl der Züge die durch den Präprozessor gelangen
  - Die Prozentuale Abweichung vom Mittelwert der Wertigkeiten der Züge
- „`heuristic`“: Eine der in 4.2 beschriebenen. Wird im Präprozessor zur Bewertung von Spielzuständen herangezogen.
- „`use_multiprocessing`“: Bool'scher Wert der angibt ob die  $N$  simulierten Spiele unter Verwendung mehrerer Prozesse durchgeführt werden sollen. Bei Systemen mit mehr als einem Prozessor können damit durch Parallelisierung Geschwindigkeitsvorteile erreicht werden.

## Die Präprozessoren

Nachfolgend wird kurz auf die Implementierung der in Kapitel 3.3 beschriebenen Präprozessoren eingegangen.

**Der feste Selektivität Präprozessor** Die in 4.8 abgedruckte Funktion gibt die Implementierung des Präprozessors mit fester Selektivität an.

Die Funktion erhält einen Spielzustand „game\_state“, den Parameter „n\_s“ der Anzahl  $N_s$  der Züge die den Präprozessor passieren und eine Heuristik „heuristic“ die zur Bewertung der Spielzüge herangezogen wird. In der sich über die Zeilen 3 bis 5 erstreckenden Anweisungen wird eine nach der Bewertung des einzelnen Zuges gemäß der Heuristik sortierte Liste von Zügen erstellt. In Zeile 6 werden dann nur die ersten „n\_s“ Züge im Spielzustand „game\_state“ gesetzt.

Listing 4.8: Die Funktion „preprocess\_fixed\_selectivity“

```

1  @staticmethod
2  def preprocess_fixed_selectivity(game_state: Othello, n_s, heuristic):
3      heuristic_values = sorted(
4          MonteCarlo.preprocess_get_heuristic_value(game_state, heuristic=heuristic).items(),
5          key=operator.itemgetter(1))
6      game_state.set_available_moves(heuristic_values[:n_s][0])

```

**Der variable Selektivität Präprozessor** In 4.9 ist die Implementierung des Präprozessors mit variabler Selektivität angegeben.

Wie der Präprozessor mit fester Selektivität erhält auch diese Funktion einen Spielzustand „game\_state“ und eine Heuristik „heuristic“. Anders als bei dem oben beschriebenen Präprozessor wird hier jedoch ein Wert „p\_s“ zur Beschreibung der maximalen prozentualen Abweichung  $p_s$  vom Mittelwert der Wertigkeit der Züge übergeben.

Für die Berechnung wird zunächst der Wert der Heuristik für alle Spielzüge berechnet (Zeile 3). Daraufhin wird in der Zeile 5 der Durchschnittliche Wert berechnet und mit der Anweisung in den zeilen 6 und 7 nur jene Züge im Spielzustand gesetzt, die einen Wert größer als  $p_s$  multipliziert mit dem Durchschnittlichen Wert haben.

Listing 4.9: Die Funktion „preprocess\_variable\_selectivity“

```

1  @staticmethod
2  def preprocess_variable_selectivity(game_state: Othello, p_s, heuristic):
3      heuristic_value_dict = MonteCarlo.preprocess_get_heuristic_value(game_state, heuristic=heuristic)
4      heuristic_values = [v for _, v in heuristic_value_dict.items()]
5      average_heuristic_value = sum(heuristic_values) / len(heuristic_values)
6      game_state.set_available_moves(
7          [m for m, v in heuristic_value_dict.items() if v >= p_s * average_heuristic_value])

```

## Die Funktion „get\_move“

Eine Funktion „get\_move“ wird von allen Agenten bereitgestellt und in der „main-game.py“ zur Auswahl eines Zuges gerufen. Beschrieben wird der Zug durch ein Paar, welches die Koordinaten auf dem Spielbrett darstellt. Nachfolgend wird die in Listing 4.10 abgedruckte Funktion diskutiert: Als Parameter erhält die Funktion den

Spielzustand „gameSpielzustand“, zu dem die Entscheidung über den nächsten Zug getroffen werden soll. In

Zeile 2 wird nun überprüft, ob die Verwendung der Starttabellen aktiviert ist und zusätzlich in dem aktuellen Spiel weniger als 21 Züge gespielt wurden. Die zweite Komponente der „if“-Abfrage erfolgt, da die Startbibliothek maximal Strategien bis zum 20-ten Zug enthält.

Ist dies der Fall so werden in Zeile 3 alle gemäß der Startbibliothek in Frage kommenden Züge ermittelt. Steht mindestens ein Zug zur Auswahl (Z. 4) so wird dieser gespielt (Z. 5). Nun werden die im vorherigen Zug ermittelten Wahrscheinlichkeiten gelöscht (Z. 6) und das Symbol, welches zur Repräsentation des eigenen Spielers verwendet wird zwischengespeichert (Z.6)

Ist die Option zur Verwendung eines Präprozessors aktiviert wird die Abfrage in Zeile 8 positiv ausgewertet. In diesem Fall wird der selektierte Präprozessor aufgerufen (Z. 9). Danach wird ermittelt ob die Option zur Verwendung mehrerer Prozesse aktiviert ist (Z. 10). Ist dies nicht der Fall, so werden „big\_n“ zufällige Spiele gespielt (Z. 11). Die dazu aufgerufene Hilfsfunktion gibt ein Object des Typs „Dictionary“ zurück. Dieses enthält für jeden möglichen Zug die Anzahl der gewonnenen Spiele wenn dieser Zug zuerst gespielt wurde und die Gesamtanzahl von Spielen.

Ist die option zur Verwendung mehrere Prozesse aktiviert, so wird der „else“-Zweig in den Zeilen 12 bis 20 ausgeführt:

Dazu wird zunächst die Anzahl der Verfügbaren Prozessorkerne ermittelt (Z. 13). Im weiteren Verlauf werden entsprechend viele Prozesse verwendet. So können alle verfügbaren Prozessorkerne genutzt werden, ohne einen zu großen Verwaltungsaufwand für die Prozesse zu generieren. In Zeile 14 wird dann ein entsprechender Pool von Prozessen angelegt und durch die Anweisung der Zeilen 15 und 16 spezifiziert, dass mittels dieser Prozesse so häufig wie Prozesse vorhanden sind asynchron jeweils „big\_n“ geteilt durch die Anzahl der Prozesse zufällige Spiele simuliert werden. Da die entsprechende Funktion eine Ganzzahl erwartet, kommt an dieser Stelle die Ganzzahldivision „//“ zum Einsatz. Dadurch ergibt die Summe aller durchgeführten Simulationen ggf. nicht „big\_n“; bei entsprechend großen Werten kann dies jedoch vernachlässigt werden.

Da die Spiele asynchron simuliert werden muss die Berechnung zunächst durch Aufruf der Funktion „get“ an jedem Listenelement angestoßen werden. Für das erste Element geschieht dies in Zeile 17. Für die weiteren Elemente erfolgt dies in der Schleife zur Zusammenführung der Ergebnisse. In Zeile 20 werden dann alle Prozesse wieder geschlossen.

Nun muss anhand der Daten zu gewonnenen und insgesamt gespielten Spiele für jeden Zug die Gewinnwahrscheinlichkeit berechnet werden. Dies erfolgt in den Zeilen 21 bis 23 indem der Quotient aus den beiden Werten gebildet wird (Z. 23)

Nun wird jener Zug mit der größten Gewinnwahrscheinlichkeit ermittelt (Z. 24) und an die aufrufende Funktion zurückgegeben (Z. 25).

Listing 4.10: get\_move Funktion des Monte-Carlo Agenten

```

1  def get_move(self, game_state: Othello):
2      if self.use_start_lib and game_state.get_turn_nr() < 21:
3          moves = self.start_tables.get_available_moves_of_start_tables(game_state)
4          if len(moves) > 0:
5              return UtilMethods.translate_move_to_pair(moves[random.randrange(len(moves))])
6      self.move_probability.clear()
7      own_symbol = game_state.get_current_player()
8      if self.preprocessor is not None:
9          self.preprocessor(game_state, self.preprocessor_parameter, self.heuristic)
10     if not self.use_multiprocessing:
11         winning_statistics = MonteCarlo.play_n_random_games(own_symbol, game_state, self.big_n)
12     else:

```

```

13         number_of_processes = mp.cpu_count()
14         pool = mp.Pool(processes=number_of_processes)
15         list_of_stats = [pool.apply_async(MonteCarlo.play_n_random_games, args=(own_symbol,
16             game_state.deepcopy(), self.big_n // number_of_processes)) for _ in range(number_of_processes)]
17         winning_statistics = list_of_stats[0].get()
18         for single_list in list_of_stats[1:]:
19             MonteCarlo.combine_statistic_dicts(winning_statistics, single_list.get())
20         pool.close()
21         for single_move in winning_statistics:
22             (games_won, times_played) = winning_statistics[single_move]
23             self.move_probability[single_move] = games_won / times_played
24         selected_move = max(self.move_probability.items(), key=operator.itemgetter(1))[0]
25         return selected_move

```

#### 4.4.4 Alpha-Beta Pruning

Der Agent „Alpha-Beta Pruning“ basiert auf dem in Kapitel 3.2.2 beschriebenen Prinzip des Alpha-Beta Abschneidens. Wie bereits erläutert handelt es sich dabei um eine Tiefensuche, bei der weniger vielversprechende Zweige nicht bis zum Ende ausgewertet werden.

Wie in 3.2.3 erläutert wird dabei ab einer gewissen Suchtiefe eine Heuristik zur Bewertung eines Spielzustandes herangezogen. Dadurch, muss nicht der komplette Baum bis zum Ende ausgewertet werden.

Nachfolgend wird die in der Datei „alphaBetaPruning.py“ bedindliche Implementierung besprochen.

##### Parameter

Auch das detaillierte Verhalten des Agenten „Alpha-Beta Pruning“ kann durch verschiedene Parameter beeinflusst werden. Sie lauten:

- „`_use_start_libs`“: Bool’scher Wert der angibt ob Startbibliotheken verwendet werden sollen.
- „`_search_depth`“: Gibt die Tiefe an, ab der einzelne Zweige nicht weiter verfolgt werden und stattdessen die Heuristik zur Zustandsbewertung verwendet wird.
- „`_heuristic`“: Eine der in Kapitel 3.2.3 beschriebenen Heuristiken.
- „`_use_monte_carlo`“: Bool’scher Wert der angibt ob in der Tiefe „`_search_depth`“ + 1 statt der Heuristik das Prinzip des Monte Carlo Agentens zur Bewertung der Züge verwendet werden soll.
- „`_mc_count`“: Anzahl der simuliert Spiele wenn das Monte Carlo Prinzip zur Bewertung des Zustandes verwendet wird.

##### Die Funktion „`value`“

Die in Listing 4.11 abgedruckte Funktion „`value`“ entspricht im wesentlichen der in Listing 3.1 abgedruckten und in 3.2.2 besprochenen „`value`“-Funktion des Alpha-Beta Abschneiden Algorithmus.

Hier soll einzig auf den sich unterscheidenden Basisfall (Zeile 3-6) eingegangen werden. Ist ein Endzustand erreicht, so ist das Spiel beendet und die Abfrage in Zeile 3 wird positiv ausgewertet. In diesem Fall wird der Wert der „`utility`“-Funktion für den aktuellen Spieler zurück gegeben (Z. 4). Um das Gewinnen bzw. Verlieren jedoch deutlich stärker zu gewichten als die Bewertungen der Heuristik wird dies zuvor mit 1000 multipliziert.



Ist die maximale Suchtiefe erreicht, so wird die Abfrage in Zeile 5 positiv ausgewertet. In diesem Fall wird die Heuristik zur Zustandsbewertung herangezogen (Z. 6)

Listing 4.11: „value“ Funktion des Alpha-Beta Spielers

```

1  @staticmethod
2  def value(game_state: Othello, depth, heuristic, alpha=-1, beta=1):
3      if game_state.game_is_over():
4          return game_state.utility(game_state.get_current_player()) * 1000
5      if depth == 0:
6          return heuristic(game_state.get_current_player(), game_state)
7      val = alpha
8      for move in game_state.get_available_moves():
9          next_state = game_state.deepcopy()
10         next_state.play_position(move)
11         val = max({val, -1 * AlphaBetaPruning.value(next_state, depth - 1, heuristic, -beta, -alpha)})
12         if val >= beta:
13             return val
14         alpha = max({val, alpha})
15     return val

```

**Änderungen an „value\_monte\_carlo“** Wird statt der Heuristik zur Bewertung der „Monte-Carlo Agent“ herangezogen so wird statt der Funktion „value“ die angepasste Funktion „value\_monte\_carlo“ aufgerufen. Dabei wird in dem in Listing 4.11 in Zeile 6 beschriebenen „if“-Zweig eine Bewertung aller möglichen Züge vorgenommen und die Gewinnwahrscheinlichkeit des besten Zuges zurückgegeben.

### Die Funktion „get\_move“

Eine Funktion „get\_move“ wird von allen Agenten bereitgestellt und in der „main-game.py“ zur Auswahl eines Zuges gerufen. Beschrieben wird der Zug durch ein Paar, welches die Koordinaten auf dem Spielbrett darstellt. Nachfolgend wird die in Listing 4.12 abgedruckte Funktion diskutiert:

Die Übergabeparameter, sowie die Zeilen bis einschließlich 5 sind identisch zu der Implmentierung der im Kapitel 4.4.3 diskutierten „get\_move“-Funktion des „Monte-Carlo“-Agenten.

In Zeile 6 wird ein Wörterbuch zur Speicherung der Bewertungen der Züge gemäß der „value“-Funktion angelegt. In diesem wird später die Bewertung als Schlüssel und als dazugehörigen Wert die Liste aller Züge die diese Bewertung erhalten haben gespeichert werden. Diese Datenstruktur wird durch die Schleife in den Zeilen 7 bis einschließlich 17 aufgebaut in dem über alle legalen Züge iteriert wird:

In den Zeilen 8 bzw. 9 wird der Ausgangszustand kopiert und der in der aktuellen Schleifeniteration betrachtete Zug „move“ ausgeführt. Nun wird die Bewertung mittels der „value“-Funktion vorgenommen. Dabei wird unterschieden ob gemäß des Parameters „use\_monte\_carlo“ in der Tiefe „search\_depth“ mittels der Funktion „value\_monte\_carlo“ das Monte-Carlo Prinzip zur Bewertung herangezogen werden soll (Z. 10f) oder ob mittels der Funktion „value“ die Heuristik verwendet wird (Z. 13f).

Nun wird überprüft ob bereits Spielzüge mit dieser Bewertung in „best\_moves“ gespeichert sind (Z. 15) Ist dies nicht der Fall, so wird die Bewertung mit der leeren Liste initialisiert (Z. 16). Anschließend wird der gerade betrachtete Zug „move“ an die entsprechende Liste angehängt.

Sobald alle legalen Züge betrachtet wurden, wird die maximale Bewertung ermittelt (Z. 18) und ein zufälliger Zug mit dieser Bewertung zurückgegeben (Z. 19)

Listing 4.12: get\_move Funktion des Alpha-Beta Spielers

```
1  def get_move(self, game_state: Othello):
2      if self._use_start_lib and game_state.get_turn_nr() < 21:
3          moves = self._start_tables.get_available_moves_of_start_tables(game_state)
4          if len(moves) > 0:
5              return UtilMethods.translate_move_to_pair(moves[random.randrange(len(moves))])
6      best_moves = dict()
7      for move in game_state.get_available_moves():
8          next_state = game_state.deepcopy()
9          next_state.play_position(move)
10         if self._use_monte_carlo:
11             result = -AlphaBetaPruning.value_monte_carlo(next_state, self._search_depth - 1, self._heuristic,
12                                                         mc_count=self._mc_count)
13         else:
14             result = -AlphaBetaPruning.value(next_state, self._search_depth - 1, self._heuristic)
15         if result not in best_moves.keys():
16             best_moves[result] = []
17         best_moves[result].append(move)
18     best_move = max(best_moves.keys())
19     return best_moves[best_move][random.randrange(len(best_moves[best_move]))]
```

# Kapitel 5

## Evaluierung

Die in Kapitel 4 beschriebenen Agenten werden nun verglichen. Sie werden zunächst in den Standardeinstellungen verglichen und in einem weiteren Schritt mit verschiedenen Variationen dieser Einstellungen. Ein Beispiel hierfür wäre die Verwendung einer anderen Heuristik. Die Standardeinstellungen der Agenten wurden empirisch an eine Laufzeit von fünf Minuten je Spiel und Agent angenähert. Dies war eine Vorgabe, dass jedem Agent etwa fünf Minuten Berechnungszeit zur Verfügung steht. In der folgenden Tabelle sind die in diesem Kapitel evaluierten Agenten aufgelistet.

Vergleich	Agent 1	Agent 2
1	Random	Monte Carlo
2	Monte Carlo	Random
3	Random	Alpha-Beta
4	Alpha-Beta	Random
5	Alpha-Beta: Nijssen 2007 Heuristik	Random
6	Alpha-Beta: Cowthello Heuristik	Random
7	Random	Alpha-Beta: Nijssen 2007 Heuristik
8	Random	Alpha-Beta: Cowthello Heuristik
9	Alpha-Beta: Nijssen 2007 Heuristik	Monte Carlo
10	Alpha-Beta: Cowthello Heuristik	Monte Carlo
11	Monte Carlo	Alpha-Beta: Nijssen 2007 Heuristik
12	Monte Carlo	Alpha-Beta: Cowthello Heuristik

Tabelle 5.1: Liste der Äquivalenzklassen

Die in dieser Tabelle aufgeführten Vergleiche werden in den folgenden Unterkapiteln durchgeführt und die Ergebnisse evaluiert. Um ein aussagekräftiges Ergebnis ermitteln zu können, werden jeweils 100 Spiele pro Vergleich durchgeführt. Da ein Unterschied existiert, welcher Agent den ersten Zug durchführt, werden beide Kombinationen verglichen.

### 5.1 Random vs. Monte Carlo

In diesem Vergleich werden die Agenten „Random“ und „Monte Carlo“ in Standardkonfiguration verglichen. Dies bedeutet, dass „Monte Carlo“ 1400 zufällige Spiele je Zug durchführt und dadurch den „besten“ Zug ermittelt. Die Auswertung ergab folgende Ergebnisse:

- Ein durchschnittliches Spiel dauert Minuten.
- Der Agent „Random“ gewann Spiele.
- Der Agent „Monte Carlo“ gewann Spiele.

Aus diese Ergebnissen können unentschiedene Spiele abgeleitet werden. Der Vergleich mit dem „Random“ Agenten soll die grundsätzliche Leistungsfähigkeit des Agenten darstellen. Statistisch ist die Gewinnwahrscheinlichkeit des „Random“ Agenten 50 Prozent. Je höher die Gewinnwahrscheinlichkeit des „Monte Carlo“ Agenten ist, dh. je mehr sie von den 50 Prozent abweicht, desto besser ist der Agent. Ist die Wahrscheinlichkeit nur minimal besser als 50 Prozent, ist der Agent nicht sehr gut, da durch simples Raten eine Gewinnwahrscheinlichkeit von 50 Prozent erreicht wird.

## 5.2 Monte Carlo vs Random

## 5.3 Random vs Alpha-Beta

## 5.4 Alpha-Beta vs Random

## 5.5 Alpha-Beta: Nijssen 2007 Heuristik vs Random

## 5.6 Alpha-Beta: Cowthello Heuristik vs Random

## 5.7 Random vs Alpha-Beta: Nijssen 2007 Heuristik

## 5.8 Random vs Alpha-Beta: Cowthello Heuristik

## 5.9 Alpha-Beta: Nijssen 2007 Heuristik vs Monte Carlo

## 5.10 Alpha-Beta: Cowthello Heuristik vs Monte Carlo






## 5.11 Monte Carlo vs Alpha-Beta: Nijssen 2007 Heuristik

## 5.12 Monte Carlo vs Alpha-Beta: Cowthello Heuristik

## Kapitel 6

## Fazit

# Notes

 am Ende schreiben . . . . .	1
 auf Fazit beziehen? . . . . .	1
 umformulieren ? . . . . .	7
 Überleitung einfügen . . . . .	7
 Warum . . . . .	9

# Literaturverzeichnis

- [Ber] Berg, Matthias. Strategieführer. <http://berg.earthlingz.de/ocd/strategy2.php>. [Online; accessed 20-January-2019].
- [Nij07] J. A. M. Nijssen. Playing othello using monte carlo, Jun 2007.
- [Ort] Ortiz, George and Berg, Matthias. Eröffnungsstrategie. <http://berg.earthlingz.de/ocd/strategy3.php>. [Online; accessed 20-January-2019].
- [o.V05] o.V. Reversi openings. <http://samsoft.org.uk/reversi/openings.htm>, 2005. [Online; accessed 24-03-2019].
- [o.V15] o.V. Spiele: Othello. [https://de.wikibooks.org/wiki/Spiele:\\_Othello](https://de.wikibooks.org/wiki/Spiele:_Othello), 2015. [Online; accessed 20-January-2019].
- [o.VoJ] o.V. Cowthello. <http://www.aurochs.org/games/cowthello/cowthello.js>, o.J. [Online; accessed 25-03-2019].
- [RN16] Stuart J. Russell and Peter Norvig. *Artificial intelligence: A modern approach*. Always learning. Pearson, 2016.
- [Str19] Prof. Dr. Karl Stroetmann. An introduction to artificial intelligence - lecture notes in progress. 2019.