

Opponent modelling

Ausnutzen von Gegnerspezifischen Schwächen





„Knowledge is power, if you know it about the right person.“
– Erastus Flavel Beadle (1821–1894)

„Knowledge is power, if you know it about the right person.“
– Erastus Flavel Beadle (1821–1894)

Was bedeutet „**Opponent Modelling**“ (zu deutsch: „Gegner-Modellierung“)?

„Knowledge is power, if you know it about the right person.“
– Erastus Flavel Beadle (1821–1894)

Was bedeutet „**Opponent Modelling**“ (zu deutsch: „Gegner-Modellierung“)?

Erstellen eines maschinen-verständlichen Abbildes des Verhaltens eines Gegners.

„Knowledge is power, if you know it about the right person.“
– Erastus Flavel Beadle (1821–1894)

Was bedeutet „**Opponent Modelling**“ (zu deutsch: „Gegner-Modellierung“)?

Erstellen eines maschinen-verständlichen Abbildes des Verhaltens eines Gegners.

Vorteile:

- ▶ Die Stärke der gegnerischen Hand aus den Aktionen folgern.
- ▶ Die Voraussage seiner spezifischen Aktion in einer gegebenen Situation.
- ▶ Wähle die eigene Strategie so, dass der Gewinn höher als der spieltheoretische Wert ausfällt.

Einleitung

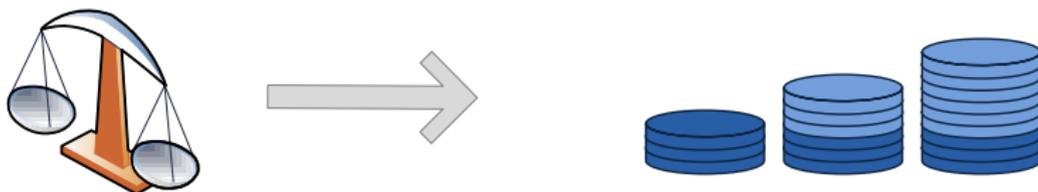
Aber warum?



- ▶ Eine optimale Strategie führt zu einem Ergebnis, das mindestens so gut ist wie jede andere Strategie, wenn man gegen einen unfehlbaren Gegner spielt.
- ▶ Genügt die optimale Strategie denn nicht?

Einleitung

Aber warum?



- ▶ Eine optimale Strategie führt zu einem Ergebnis, das mindestens so gut ist wie jede andere Strategie, wenn man gegen einen unfehlbaren Gegner spielt.
- ▶ **Genügt die optimale Strategie denn nicht?**
- ▶ unsere Gegner agieren nicht immer optimal. Wir könnten ihre Schwächen ausnutzen.

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein gegnerisches Modell kann sehr schnell sehr **komplex** werden.

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein gegnerisches Modell kann sehr schnell sehr **komplex** werden.
 - ▶ Fehlende Informationen

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein gegnerisches Modell kann sehr schnell sehr **komplex** werden.
 - ▶ Fehlende Informationen
 - ▶ (zu) viele verschiedene Einflüsse

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein gegnerisches Modell kann sehr schnell sehr **komplex** werden.
 - ▶ Fehlende Informationen
 - ▶ (zu) viele verschiedene Einflüsse
 - ▶ menschliche Intuition

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein gegnerisches Modell kann sehr schnell sehr **komplex** werden.
 - ▶ Fehlende Informationen
 - ▶ (zu) viele verschiedene Einflüsse
 - ▶ menschliche Intuition
 - ▶ Ziel der gegnerischen Strategie

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein falsches Modell kann zu schlechten Strategien führen.

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein falsches Modell kann zu schlechten Strategien führen.
 - ▶ Möglicherweise veraltetes Modell

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein falsches Modell kann zu schlechten Strategien führen.
 - ▶ Möglicherweise veraltetes Modell
- ▶ Gegner muss in relativ kurzer Zeit mit relativ wenigen Informationen analysiert werden

Einige der schwierigsten Probleme der künstlichen Intelligenz werden von Opponent Modelling abgedeckt.

- ▶ Ein falsches Modell kann zu schlechten Strategien führen.
 - ▶ Möglicherweise veraltetes Modell
- ▶ Gegner muss in relativ kurzer Zeit mit relativ wenigen Informationen analysiert werden
 - ▶ konkreter Zustand eines Spiels ist meist nicht bekannt, lediglich dessen Klasse

- ▶ Typen von Opponent-Modeling
 - ▶ generic opponent modeling
 - ▶ specific opponent modeling

- ▶ Typen von Opponent-Modeling
 - ▶ ~~generic opponent modeling~~
 - ▶ specific opponent modeling
- ▶ reduziere das Modellierungsproblem auf ein Lernproblem
 - ▶ Schwäche-Einschätzung

- ▶ Typen von Opponent-Modeling
 - ▶ ~~generic opponent modeling~~
 - ▶ specific opponent modeling
- ▶ reduziere das Modellierungsproblem auf ein Lernproblem
 - ▶ Schwäche-Einschätzung
- ▶ Gegner Beobachtung
 - ▶ Parameter Learning

- ▶ Typen von Opponent-Modeling
 - ▶ ~~generic opponent modeling~~
 - ▶ specific opponent modeling
- ▶ reduziere das Modellierungsproblem auf ein Lernproblem
 - ▶ Schwäche-Einschätzung
- ▶ Gegner Beobachtung
 - ▶ Parameter Learning
 - ▶ Strategy Learning

- ▶ Typen von Opponent-Modeling
 - ▶ ~~generic opponent modeling~~
 - ▶ specific opponent modeling
- ▶ reduziere das Modellierungsproblem auf ein Lernproblem
 - ▶ Schwäche-Einschätzung
- ▶ Gegner Beobachtung
 - ▶ Parameter Learning
 - ▶ Strategy Learning
- ▶ den Gegner **ausbeuten**
 - ▶ Nutze die „**Schwäche**“ des Gegners, anstatt den Gegner **vollständig** zu beschreiben.

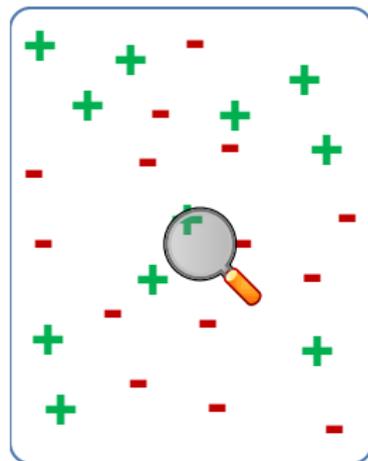
- ▶ Typen von Opponent-Modeling
 - ▶ ~~generic opponent modeling~~
 - ▶ specific opponent modeling
- ▶ reduziere das Modellierungsproblem auf ein Lernproblem
 - ▶ Schwäche-Einschätzung
- ▶ Gegner Beobachtung
 - ▶ Parameter Learning
 - ▶ Strategy Learning
- ▶ den Gegner **ausbeuten**
 - ▶ Nutze die „**Schwäche**“ des Gegeners, anstatt den Gegner **vollständig** zu beschreiben.
 - ▶ Begrenze das Risiko durch **Abwägung** unterschiedlich guter Strategien.



Idee war: Nutze die „Schwäche“ des Gegeners, anstatt den Gegner **vollständig** zu beschreiben.

Idee war: Nutze die „Schwäche“ des Gegeners, anstatt den Gegner **vollständig** zu beschreiben.

- ▶ Die Strategien der Gegner müssen anhand ihrer **Schwäche** bzw. ihrer **Ausnutzbarkeit** kategorisiert werden.
- ▶ Aber wie „misst“ man Schwäche?



Messen heißt Vergleichen

Einführung einer **Teacher-Funktion** $T(s)$ ¹:

$$\text{Weakness}_{\phi, T}(s) \Leftrightarrow \phi(s) \neq T(s)$$

Schwäche kann also im Zustand s durch Vergleich der Gegner-Strategie ϕ mit der Teacher-Strategie T gemessen werden.

¹Markovitch & Reger [3]

Schwäche-Abschätzung

Optimierung

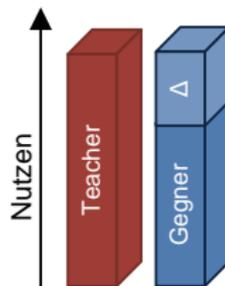
Problem: Eine von der Teacher-Strategie **verschiedene** gegnerische Strategie kann aber einen **gleichen Nutzen** besitzen.

Problem: Eine von der Teacher-Strategie **verschiedene** gegnerische Strategie kann aber einen **gleichen Nutzen** besitzen.

Idee: Anstatt eine Strategie-Entscheidung zu bewerten, schauen wir uns den **Nutzen** an.

$$\text{Weakness}_{\phi, T, U_s}(s) := U_T(a_{T,s}(s)) - U_T(a_{\phi,s}(s))$$

Dazu verwenden wir die **Teacher-Nutzenfunktion** U_T , welche den Nutzen des aus der Aktion a resultierenden Zustands bewertet.



$$\text{Weakness}_{\phi, T, U_s}(s) := U_T(a_{T,s}(s)) - U_T(a_{\phi,s}(s))$$

- ▶ Die **Differenz** der Nutzwerte der Teacher-Aktion $a_{T,s}(s)$ und der Gegner-Aktion $a_{\phi,s}(s)$ stellt dann die Schwäche des Gegners als reelle Zahl dar.

$$\text{Weakness}_{\phi, T, U_s}(s) := U_T(a_{T,s}(s)) - U_T(a_{\phi,s}(s))$$

- ▶ Die **Differenz** der Nutzwerte der Teacher-Aktion $a_{T,s}(s)$ und der Gegner-Aktion $a_{\phi,s}(s)$ stellt dann die Schwäche des Gegners als reelle Zahl dar.
- ▶ Dies kann auch als **Prädikat** ausgedrückt werden:

$$\text{Weakness}_{\phi, T, U_s}(s) \Leftrightarrow U_T(a_{T,s}(s)) - U_T(a_{\phi,s}(s)) \geq b_\epsilon(S)$$

Wobei $b_\epsilon(S)$ eine **untere Schranke** der Schwäche-Definition darstellt.

Schwäche-Abschätzung

Teacher-Nutzenfunktion



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- ▶ Damit die Schwäche-Abschätzung zuverlässig funktioniert, muss die **Teacher-Nutzenfunktion** U_T sehr genau sein.

Schwäche-Abschätzung

Teacher-Nutzenfunktion

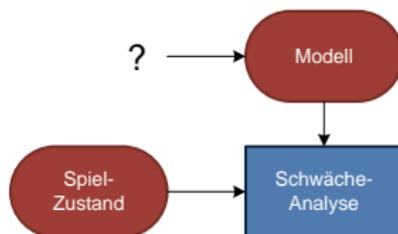


- ▶ Damit die Schwäche-Abschätzung zuverlässig funktioniert, muss die **Teacher-Nutzenfunktion** U_T sehr genau sein.
- ▶ Sie basiert daher auf einem **offline gelernten Modell**, das aus der Analyse von sehr vielen vorherigen Spielen gewonnen werden kann.



Das Modellierungs-Problem ist nun ein **Lernproblem**.

Das Modellierungs-Problem ist nun ein **Lernproblem**.



- ▶ Wie bekommt man die gegnerische Strategie?
- ▶ Wie erforscht man den Gegner?

- ▶ Problem: Der Lernprozess muss **schnell** erfolgen und **gute** Ergebnisse liefern. Beobachtete Objekte sind **nur** die gegnerischen Aktionen.



- ▶ Problem: Der Lernprozess muss **schnell** erfolgen und **gute** Ergebnisse liefern. Beobachtete Objekte sind **nur** die gegnerischen Aktionen.
- ▶ menschlicher Ansatz: Intuition einer Theorie und ausprobieren der Theorie.



- ▶ Problem: Der Lernprozess muss **schnell** erfolgen und **gute** Ergebnisse liefern. Beobachtete Objekte sind **nur** die gegnerischen Aktionen.
- ▶ menschlicher Ansatz: Intuition einer Theorie und ausprobieren der Theorie.
- ▶ maschineller Ansatz: **Parameter Learning** oder **Strategy Learning**



- ▶ Ansatz: **Beobachte** und **bewerte** die Parameter der gegnerischen Strategie².
- ▶ **Beobachten:**

- ▶ **Bewerten:**

²Hoen, Southey, Holte & Bulitko [1]

- ▶ Ansatz: **Beobachte** und **bewerte** die Parameter der gegnerischen Strategie².
- ▶ **Beobachten**: Die Parameter beschreiben eine gemischte Strategie mit unterschiedlichen Wahrscheinlichkeiten. Einführung eines Counters pro Parameter.
- ▶ **Bewerten**:

²Hoen, Southey, Holte & Bulitko [1]

- ▶ Ansatz: **Beobachte** und **bewerte** die Parameter der gegnerischen Strategie².
- ▶ **Beobachten**: Die Parameter beschreiben eine gemischte Strategie mit unterschiedlichen Wahrscheinlichkeiten. Einführung eines Counters pro Parameter.
- ▶ **Bewerten**: Sie werden von einer *Maximum a Priori* (MAP) Funktion ausgewertet und von einer zum jeweiligen Parameter gehörigen „*Beta prior*“-Verteilung bewertet.

$$Beta(\theta, \omega)_{Parameter} : \mathbb{N} \times \mathbb{N} \longrightarrow [0, 1]$$

θ = Häufigkeit des Auftretens des jeweiligen Parameters

ω = Häufigkeit des Nichtauftretens des jeweiligen Parameters

($\theta, \omega = 10$ gibt bspw. mehr „Vertrauen“ als $\theta, \omega = 1$)

²Hoen, Southey, Holte & Bulitko [1]



Weiterführende Überlegungen:

- ▶ Anmerkung: „best response“-Strategie liefert das günstigste Ergebnis für einen Spieler anhand eines bekannten gegnerischen Modells.

³Johanson, Zinkewich & Bowling [2]



Weiterführende Überlegungen:

- ▶ Anmerkung: „best response“-Strategie liefert das günstigste Ergebnis für einen Spieler anhand eines bekannten gegnerischen Modells.
- ▶ Wir können nun (einfach) die beste Antwort („best response“-Strategie) spielen → **Frequentist Best Response**³ (FBR),

³Johanson, Zinkewich & Bowling [2]



Weiterführende Überlegungen:

- ▶ Anmerkung: „best response“-Strategie liefert das günstigste Ergebnis für einen Spieler anhand eines bekannten gegnerischen Modells.
- ▶ Wir können nun (einfach) die beste Antwort („best response“-Strategie) spielen → **Frequentist Best Response³** (FBR),
- ▶ oder wir wägen zwischen einer (pessimistischen) „Nash-Gleichgewicht“-Strategie und der „best response“-Strategie ab → **Restricted Best Response³** (RBR).

³Johanson, Zinkewich & Bowling [2]



Weiterführende Überlegungen:

- ▶ Anmerkung: „best response“-Strategie liefert das günstigste Ergebnis für einen Spieler anhand eines bekannten gegnerischen Modells.
- ▶ Wir können nun (einfach) die beste Antwort („best response“-Strategie) spielen → **Frequentist Best Response³** (FBR),
- ▶ oder wir wägen zwischen einer (pessimistischen) „Nash-Gleichgewicht“-Strategie und der „best response“-Strategie ab → **Restricted Best Response³** (RBR).
- ▶ **FBR** erzielt i.A. sehr gute Ergebnisse für einen bekannten Gegner, verliert jedoch einschlägig gegen Gegner mit anderen Strategien.
- ▶ **RBR** erzielt i.A. nicht so gute Ergebnisse wie FBR für einen bestimmte Gegner, verliert jedoch auch nicht gegen unbekannte Gegner.

³Johanson, Zinkewich & Bowling [2]

Gegner-Beobachtung

Strategy Learning



TECHNISCHE
UNIVERSITÄT
DARMSTADT

- ▶ Strategien als **Experten** (Hoen, Southey, Holte & Bulitko [1])

- ▶ Strategien als **Experten** (Hoen, Southey, Holte & Bulitko [1])
- ▶ Verwendung von **Exp3** (sog. „*bounded regret*“ Algorithmus)

1. Initialize the scores for the K strategies: $s_j = 0$
2. For $t=1, 2, \dots$ until the game ends:

(a) Let the probability of playing the i th strategy for hand t be

$$p_i(t) = (1 - \psi) \frac{(1+\rho)^{s_i(t)}}{\sum_{j=1}^K (1+\rho)^{s_j(t)}} + \frac{\psi}{K}$$

(b) Select the strategy to play u according to the distribution p and observe the hand's winnigs w .

$$(c) s_j(t+1) = A \begin{cases} s_j(t) + \frac{\psi w}{K p_j(t)} & \text{if } u = i \\ s_j(t) & \text{if } u \neq i \end{cases}$$

Lernrate $\rho > 0$ und Forschungsrate $0 \leq \psi \leq 1$



- ▶ Strategien als **Experten** (Hoen, Southey, Holte & Bulitko [1])
- ▶ Verwendung von **Exp3** (sog. „*bounded regret*“ Algorithmus)

1. Initialize the scores for the K strategies: $s_j = 0$

2. For $t=1,2,\dots$ until the game ends:

(a) Let the probability of playing the i th strategy for hand t be

$$p_i(t) = (1 - \psi) \frac{(1+\rho)^{s_i(t)}}{\sum_{j=1}^K (1+\rho)^{s_j(t)}} + \frac{\psi}{K}$$

(b) Select the strategy to play u according to the distribution p and observe the hand's winnings w .

$$(c) s_j(t+1) = A \begin{cases} s_j(t) + \frac{\psi w}{K p_j(t)} & \text{if } u = i \\ s_j(t) & \text{if } u \neq i \end{cases}$$

Lernrate $\rho > 0$ und Forschungsrate $0 \leq \psi \leq 1$

- ▶ Verbesserungen (da wenig Samples):

- ▶ Strategien als **Experten** (Hoen, Southey, Holte & Bulitko [1])
- ▶ Verwendung von **Exp3** (sog. „*bounded regret*“ Algorithmus)

1. Initialize the scores for the K strategies: $s_j = 0$

2. For $t=1,2,\dots$ until the game ends:

(a) Let the probability of playing the i th strategy for hand t be

$$p_i(t) = (1 - \psi) \frac{(1+\rho)^{s_i(t)}}{\sum_{j=1}^K (1+\rho)^{s_j(t)}} + \frac{\psi}{K}$$

(b) Select the strategy to play u according to the distribution p and observe the hand's winnings w .

$$(c) s_j(t+1) = A \begin{cases} s_j(t) + \frac{\psi w}{K p_j(t)} & \text{if } u = i \\ s_j(t) & \text{if } u \neq i \end{cases}$$

Lernrate $\rho > 0$ und Forschungsrate $0 \leq \psi \leq 1$

- ▶ Verbesserungen (da wenig Samples):
 - ▶ mehrere „Scores“ pro Strategie (abhängig von den eigenen Karten) und einen Counter für jede Karte

- ▶ Strategien als **Experten** (Hoen, Southey, Holte & Bulitko [1])
- ▶ Verwendung von **Exp3** (sog. „*bounded regret*“ Algorithmus)

1. Initialize the scores for the K strategies: $s_j = 0$

2. For $t=1,2,\dots$ until the game ends:

(a) Let the probability of playing the i th strategy for hand t be

$$p_i(t) = (1 - \psi) \frac{(1+\rho)^{s_i(t)}}{\sum_{j=1}^K (1+\rho)^{s_j(t)}} + \frac{\psi}{K}$$

(b) Select the strategy to play u according to the distribution p and observe the hand's winnings w .

$$(c) s_j(t+1) = A \begin{cases} s_j(t) + \frac{\psi w}{K p_j(t)} & \text{if } u = i \\ s_j(t) & \text{if } u \neq i \end{cases}$$

Lernrate $\rho > 0$ und Forschungsrate $0 \leq \psi \leq 1$

- ▶ Verbesserungen (da wenig Samples):
 - ▶ mehrere „Scores“ pro Strategie (abhängig von den eigenen Karten) und einen Counter für jede Karte
 - ▶ „Belohnung“ für Experten, die dieselbe Aktion in einer gegebenen Situation vorschlagen



- ▶ Strategien als **Experten** (Hoen, Southey, Holte & Bulitko [1])
- ▶ Verwendung von **Exp3** (sog. „*bounded regret*“ Algorithmus)

1. Initialize the scores for the K strategies: $s_j = 0$

2. For $t=1,2,\dots$ until the game ends:

(a) Let the probability of playing the i th strategy for hand t be

$$p_i(t) = (1 - \psi) \frac{(1+\rho)^{s_i(t)}}{\sum_{j=1}^K (1+\rho)^{s_j(t)}} + \frac{\psi}{K}$$

(b) Select the strategy to play u according to the distribution p and observe the hand's winnings w .

$$(c) \quad s_j(t+1) = A \begin{cases} s_j(t) + \frac{\psi w}{K p_j(t)} & \text{if } u = i \\ s_j(t) & \text{if } u \neq i \end{cases}$$

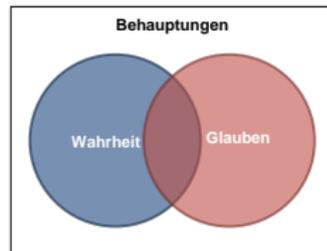
Lernrate $\rho > 0$ und Forschungsrate $0 \leq \psi \leq 1$

- ▶ Verbesserungen (da wenig Samples):
 - ▶ mehrere „Scores“ pro Strategie (abhängig von den eigenen Karten) und einen Counter für jede Karte
 - ▶ „Belohnung“ für Experten, die dieselbe Aktion in einer gegebenen Situation vorschlagen
- ▶ weitere Überlegung: Verwende **RBR** als Experten.

Gegner Beobachtung

Schlussfolgerungen

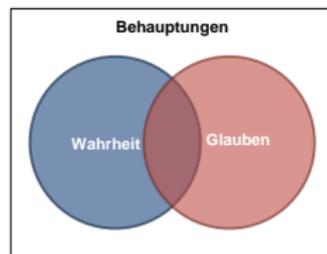
- ▶ Wann gilt der Gegner als durchschaut?



Gegner Beobachtung

Schlussfolgerungen

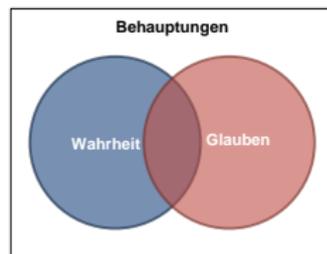
- ▶ Wann gilt der Gegner als durchschaut?
- ▶ Das weiß man nie. Ein guter Spieler wechselt seine Strategie recht oft. (→ausprobieren)



Gegner Beobachtung

Schlussfolgerungen

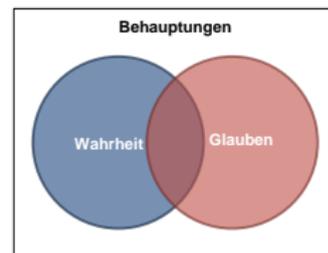
- ▶ Wann gilt der Gegner als durchschaut?
- ▶ Das weiß man nie. Ein guter Spieler wechselt seine Strategie recht oft. (→ausprobieren)
- ▶ Experimente zeigen, dass das Lernen, einen Gegner maximal auszunutzen, anhand weniger Hände kaum möglich ist.



Gegner Beobachtung

Schlussfolgerungen

- ▶ Wann gilt der Gegner als durchschaut?
- ▶ Das weiß man nie. Ein guter Spieler wechselt seine Strategie recht oft. (→ausprobieren)
- ▶ Experimente zeigen, dass das Lernen, einen Gegner maximal auszunutzen, anhand weniger Hände kaum möglich ist.
- ▶ Jedoch ist der **Lernansatz** wesentlich erfolgreicher als der Ansatz einer (pessimistischen) „Nash-Gleichgewicht“-Strategie.



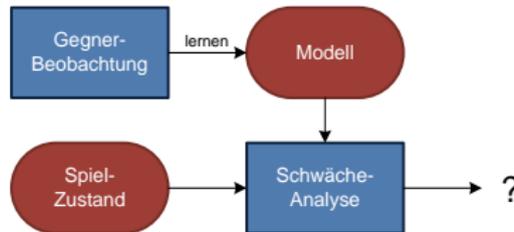


Wir können nun

- ▶ den Gegner **beobachten** und
- ▶ anhand der Beobachtung seine **Schwäche** einschätzen.

Wir können nun

- ▶ den Gegner **beobachten** und
- ▶ anhand der Beobachtung seine **Schwäche** einschätzen.



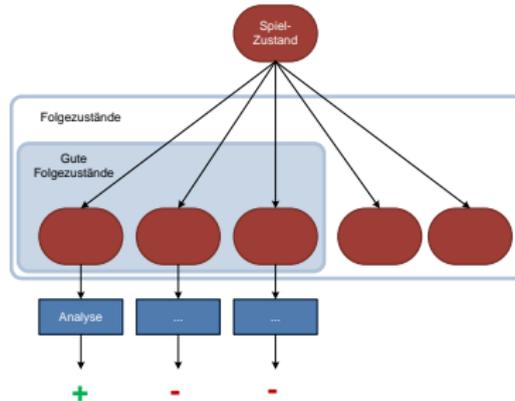
Welchen Nutzen können wir daraus ziehen?

Idee: Wir versuchen einen Spiel-Zustand zu erreichen, in dem der Gegner schwach ist.

Durch geschickte Wahl der Strategie vermeiden wir, dass wir selbst in einen unvorteilhaften Zustand landen. Drei mögliche Verfahren sind⁴:

- ▶ Fensterbasierte Methode
- ▶ Modifizierte Nutzenfunktion
- ▶ Eine Mischung aus beiden

⁴Markovitch & Reger [3]



- ▶ Bei der **Fensterbasierten Methode** werden aus der Menge der möglichen Strategien nur die beste und eine Anzahl (Fenster) der **nahezu besten** Strategien ausgewählt.
- ▶ Diese Strategien werden dann nach ihrer **Schwäche-Klassifizierung** selektiert.



Zur Erinnerung: Die bisher verwendete Nutzenfunktion bildet vom Zustandsraum auf ein Zahlenintervall $[0, 1]$, welches den Nutzwert darstellt, ab.

- ▶ Anstatt der zuvor beschriebenen zweistufigen Strategie-Selektion, kann auch gleich die **Nutzenfunktion modifiziert** werden.

Zur Erinnerung: Die bisher verwendete Nutzenfunktion bildet vom Zustandsraum auf ein Zahlenintervall $[0, 1]$, welches den Nutzwert darstellt, ab.

- ▶ Anstatt der zuvor beschriebenen zweistufigen Strategie-Selektion, kann auch gleich die **Nutzenfunktion modifiziert** werden.
- ▶ Die modifizierte Nutzenfunktion ergibt sich aus einer Linearkombination der ursprünglichen Nutzenfunktion $u(s)$ und der **Schwächeabschätzung** $Weakness(s) \in [0, 1]$:

$$u'(s) = (1 - \delta_u) \cdot u(s) + \delta_u \cdot Weakness(s)$$

Wobei $\delta_u \in [0, 1]$ beeinflusst, wie stark die Gegner-Schwäche gegenüber der ursprünglichen Nutzenfunktion **gewichtet** wird.

Es könnte sein, dass uns die ausgewählte Strategie in einen Zustand versetzt, in dem sowohl der Gegner als auch der **eigene Agent** Schwäche zeigt.

Es könnte sein, dass uns die ausgewählte Strategie in einen Zustand versetzt, in dem sowohl der Gegner als auch der **eigene Agent** Schwäche zeigt.

Eine zusätzliche Selektion soll diese Strategien ausschließen. Dazu muss neben den Gegnern auch der eigene Agent modelliert werden.

Fragen?!



Quellen

Hauptquellen

-  Bret Hoen, Finnegan Southey, Robert C. Holte, and Valeriy Bulitko.
Effective short-term opponent exploitation in simplified poker.
Technical report, University of Alberta, Athabasca University, 2005.
<http://cs.ualberta.ca/~holte/Publications/aaai2005poker.pdf>.
-  Micheal Johanson, Martin Zinkevich, and Micheal Bowling.
Computing robust counter-strategies.
Technical report, University of Alberta, 2008.
<http://www.cs.ualberta.ca/~games/poker/papers/Papers/NIPS07-rnash.pdf>.
-  Shaul Markovitch and Ronit Rubinfeld.
Learning and exploiting relative weaknesses of opponent agents.
Technical report, Israel Institute of Technology, 2005.
<http://www.springerlink.com/content/v1pq0023768762u1/>.

Titelseite	<i>Louis Wain</i> . Cats playing poker (1915, gemeinfrei)
Seite 3	<i>Microsoft Visio</i> . Clipart ©
Seite 7	Eigene Infografik basierend auf einem „Nuvola icon“ (KDE, LGPL)
Seite 9, 12, 16, 17, 19	Eigene Infografiken
Seite 13	<i>unbekannter Author</i> . Why Dogs Make Lousy Poker Players ©
Seite 23	<i>Thomas Bros.</i> (<i>cartoonstock.com</i>). Online Poker Night ©