

## Spieltheorie

Prof. Dr. B. Nebel, Dr. M. Helmert  
Wintersemester 2007/2008

Universität Freiburg  
Institut für Informatik

### Übungsblatt 9 — Lösungen

#### Aufgabe 9.1 (TPGs und Minimax-Algorithmus, 4 Punkte)

Sei  $\Gamma = \langle \{1, 2\}, H, P, (u_i) \rangle$  ein endliches extensives Spiel mit perfekter Information mit den folgenden Eigenschaften:

- $\Gamma$  ist ein Nullsummenspiel, d. h.  $u_1(h) = -u_2(h)$  für alle  $h \in Z$ .
  - Die Spieler sind abwechselnd am Zug (beginnend mit Spieler 1), d. h.  $P(h) = 1$  gdw.  $|h|$  gerade ist für alle nichtterminalen Historien  $h \in H \setminus Z$ .
  - Alle Ausgänge werden unterschiedlich bewertet, d. h. für alle terminalen Historien  $h, h' \in Z$  mit  $h \neq h'$  gilt  $u_1(h) \neq u_1(h')$ .
- (a) Zeigen Sie: In  $\Gamma$  gibt es ein *eindeutiges* teilspielperfektes Gleichgewicht  $s^*$ .

#### Lösung:

Wir können  $s^*$  wie im Beweis des Satzes von Kuhn konstruieren. Da in jedem Konstruktionsschritt die zu wählende Aktion eindeutig ist, ist auch das resultierende teilspielperfekte Gleichgewicht eindeutig.

Die Konstruktion von  $s^*$  erfolgt induktiv über die Länge  $\ell(\Gamma(h))$  aller Teilspiele  $\Gamma(h)$  von  $\Gamma$ . Parallel dazu werden Funktionen  $t_i : H \rightarrow \mathbb{R}$  für alle Spieler  $i \in N$  so konstruiert, dass  $t_i(h^*)$  die Auszahlung für Spieler  $i$  in einem teilspielperfekten Gleichgewicht in  $\Gamma(h)$  angibt.

Um die Eindeutigkeit von  $s^*$  zu beweisen, müssen wir induktiv über die Konstruktion der Funktionen  $t_i$  zeigen, dass für alle nichtterminalen Historien  $h^*$  mit  $P(h^*) = i$  die Nutzenwerte  $t_i(\langle h^*, a \rangle)$  für alle  $a \in A(h^*)$  paarweise verschieden sind. Damit erhält man sofort einen Induktionsbeweis dafür, dass  $\arg \max_{a \in A(h^*)} t_i(\langle h^*, a \rangle)$  für alle nichtterminalen Historien  $h^*$  mit  $P(h^*) = i$  eindeutig ist. Wegen  $s_i^*(h^*) := \arg \max_{a \in A(h^*)} t_i(\langle h^*, a \rangle)$  ist dann auch  $s^*$  eindeutig.

Zur Konstruktion: Falls  $\ell(\Gamma(h)) = 0$ , dann ist  $t_i(h) = u_i(h)$  für alle  $i \in N$ . Alle  $t_i(h)$  sind paarweise verschieden nach Voraussetzung.

Sei  $t_i(h)$  für alle  $h \in H$  mit  $\ell(\Gamma(h)) \leq k$  bereits definiert. Betrachte nun  $h^* \in H$  mit  $\ell(\Gamma(h^*)) = k + 1$  und  $P(h^*) = i$ . Für alle  $a \in A(h^*)$  ist  $\ell(\Gamma(h^*, a)) \leq k$ . Setze  $s_i^*(h^*) := \arg \max_{a \in A(h^*)} t_i(\langle h^*, a \rangle)$  und  $t_j(h^*) := t_j(h^*, s_i^*(h^*))$  für alle  $j \in N$ . Der einzige Unterschied zum Beweis des Satzes von Kuhn ist nun, dass nach Induktionsvoraussetzung die Werte  $t_i(h^*, a) = t_i(\langle h^*, a \rangle, s_{P(\langle h^*, a \rangle)}^*(\langle h^*, a \rangle))$  für alle  $a \in A(h)$  paarweise verschieden sind, woraus die Eindeutigkeit von  $s_i^*(h^*)$  folgt. Außerdem gilt die paarweise Verschiedenheit der  $t_j$ -Werte induktiv auch am Elternknoten von  $h^*$  (sofern ein solcher existiert) nach der induktiven Definition der  $t_j$  (dazu benötigt man noch die Nullsummeneigenschaft, das sich die unterschiedliche Bewertung der Historien durch Spieler 1 sonst nicht notwendigerweise auf Spieler 2 übertragen würde).

Das so induktiv gewonnene (eindeutige) Strategieprofil  $s^*$  erfüllt die Ein-Schritt-Abweichungs-Eigenschaft und ist damit ein teilspielperfektes Gleichgewicht.

- (b) Zeigen Sie: In allen nichtterminalen Historien  $h \in H \setminus Z$  ist die Gleichgewichtsaktion  $s_{P(h)}^*(h)$  genau die Aktion, die auch der Minimax-Algorithmus in dem von  $H$  definierten Zustandsraum unter der Bewertungsfunktion  $u_1$  (bzw.  $u_2$ ) wählen würde. (*Hinweis:* Der Minimax-Algorithmus wird z. B. in dem Lehrbuch *Artificial Intelligence: A Modern Approach* von Russell und Norvig beschrieben.)

**Lösung:**

Der Minimax-Algorithmus wählt in jeder nichtterminalen Historie  $h \in H \setminus Z$  die Aktion  $a \in A(h)$  aus, die unter den Aktionen aus  $A(h)$  den Minimax-Wert  $v(\langle h, a \rangle)$  maximiert (minimiert), falls  $|h|$  gerade (ungerade) ist, wobei der Minimax-Wert  $v(h)$  für Historien  $h \in H$  wie folgt definiert ist:

$$v(h) = \begin{cases} u_1(h), & \text{falls } h \in Z \\ \max_{a \in A(h)} v(\langle h, a \rangle) & \text{falls } h \in H \setminus Z \text{ und } P(h) = 1 \\ \min_{a \in A(h)} v(\langle h, a \rangle) & \text{falls } h \in H \setminus Z \text{ und } P(h) = 2 \end{cases}$$

Wir beweisen durch Induktion über  $\ell(\Gamma(h))$ , dass  $t_1(h) = -t_2(h) = v(h)$  für alle Historien  $h \in H$  und dass

$$s_{P(h)}^*(h) = \begin{cases} \arg \max_{a \in A(h)} v(\langle h, a \rangle) & \text{falls } P(h) = 1 \\ \arg \min_{a \in A(h)} v(\langle h, a \rangle) & \text{falls } P(h) = 2 \end{cases}$$

für alle nichtterminalen Historien  $h \in H \setminus Z$ .

Zum Induktionsanfang müssen wir nur den ersten Teil zeigen, weil bei  $\ell(\Gamma(h)) = 0$  nur terminale Historien möglich sind. Es gilt dann  $t_1(h) = u_1(h) = v(h)$  nach Definition von  $t_1$  und  $v$  und ferner  $t_2(h) = u_2(h) = -u_1(h) = -t_1(h)$ .

Im Induktionsschritt sei nun  $\ell(\Gamma(h)) = n+1 > 0$ ,  $h$  ist also keine terminale Historie und  $s_{P(h)}^*(h)$  ist definiert als  $\arg \max_{a \in A(h)} t_{P(h)}(\langle h, a \rangle)$ . Falls  $P(h) = 1$ , ist nach Induktionsvoraussetzung  $t_{P(h)}(\langle h, a \rangle) = v(\langle h, a \rangle)$ , also  $s_{P(h)}^*(h) = \arg \max_{a \in A(h)} v(\langle h, a \rangle)$ . Falls  $P(h) = 2$ , ist nach Induktionsvoraussetzung  $t_{P(h)}(\langle h, a \rangle) = -v(\langle h, a \rangle)$ , also  $s_{P(h)}^*(h) = \arg \max_{a \in A(h)} t_{P(h)}(\langle h, a \rangle) = \arg \max_{a \in A(h)} -v(\langle h, a \rangle) = \arg \min_{a \in A(h)} v(\langle h, a \rangle)$ .

Ferner ist (nach Definition von  $t_1$  und wegen der Nullsummeneigenschaft von  $\Gamma$ )

$$t_1(h) = t_1(\langle h, s_{P(h)}^*(h) \rangle) = \begin{cases} \max_{a \in A(h)} t_1(\langle h, a \rangle) & \text{falls } P(h) = 1 \\ \min_{a \in A(h)} t_1(\langle h, a \rangle) & \text{falls } P(h) = 2 \end{cases}$$

Nach Induktionsvoraussetzung ist das wiederum

$$\begin{cases} \max_{a \in A(h)} v(\langle h, a \rangle) & \text{falls } P(h) = 1 \\ \min_{a \in A(h)} v(\langle h, a \rangle) & \text{falls } P(h) = 2 \end{cases}$$

und das ist nach Definition von  $v$  gerade  $v(h)$ .

Analog kann man zeigen, dass  $t_2(h) = -v(h)$  gilt, denn

$$t_2(h) = t_2(\langle h, s_{P(h)}^*(h) \rangle) = \begin{cases} \min_{a \in A(h)} t_2(\langle h, a \rangle) & \text{falls } P(h) = 1 \\ \max_{a \in A(h)} t_2(\langle h, a \rangle) & \text{falls } P(h) = 2 \end{cases}$$

Nach Induktionsvoraussetzung ist das wiederum

$$\begin{cases} \min_{a \in A(h)} -v(\langle h, a \rangle) & \text{falls } P(h) = 1 \\ \max_{a \in A(h)} -v(\langle h, a \rangle) & \text{falls } P(h) = 2 \end{cases}$$

bzw.

$$\begin{cases} -\max_{a \in A(h)} v(\langle h, a \rangle) & \text{falls } P(h) = 1 \\ -\min_{a \in A(h)} v(\langle h, a \rangle) & \text{falls } P(h) = 2 \end{cases}$$

und das ist nach Definition von  $v$  gerade  $-v(h)$ .

**Aufgabe 9.2** (Zufallszüge, 4 Punkte)

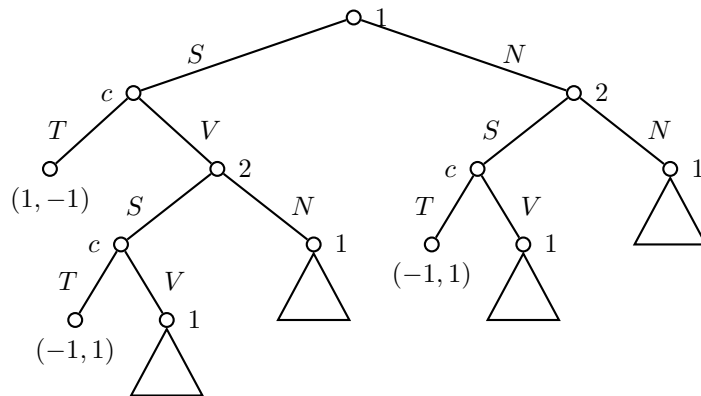
Ein *extensives Spiel mit perfekter Information und Zufallszügen* ist ein Tupel  $\Gamma = \langle N, H, P, f_c, (u_i) \rangle$ , wobei  $N$ ,  $H$  und  $u_i$  wie bei extensiven Spielen definiert sind, die Spielerfunktion  $P : H \rightarrow N \cup \{c\}$  auch den Wert  $c$  für einen Zufallsknoten annehmen kann, und für jedes  $h \in H \setminus Z$  mit  $P(h) = c$  die Funktion  $f_c(\cdot|h)$  ein Wahrscheinlichkeitsmaß über  $A(h)$  ist, wobei die Wahrscheinlichkeitsmaße  $f_c(\cdot|h)$  für alle  $h \in H$  unabhängig voneinander sind.

Strategien sind wie zuvor definiert. Der Ausgang eines Spieles bei gegebenem Strategieprofil ist ein Wahrscheinlichkeitsmaß über der Menge der terminalen Historien und  $U_i$  ist die erwartete Auszahlung für Spieler  $i$ .

- (a) Modellieren Sie die folgende Situation: In einem *sequentiellen Duell* haben zwei Spieler (1 und 2) abwechselnd die Gelegenheit, aufeinander zu schießen. Spieler 1 beginnt. Beide haben unendlich viele Kugeln zur Verfügung. In jedem Schritt kann der Spieler, der an der Reihe ist, entweder schießen oder nichts tun. Jeder Schuss von Spieler  $i$  trifft und tötet den Gegner mit Wahrscheinlichkeit  $p_i$ . Beide Spieler wollen ihre eigene Überlebenswahrscheinlichkeit maximieren.

**Lösung:**

Da beide Spieler unendlich viele Kugeln zur Verfügung haben, hat das Spiel einen unendlichen Horizont. Der Spielbaum ist rekursiv aufgebaut und Situationen wiederholen sich nach jedem Paar von Aktionen der beiden Spieler.



Formal ist  $\Gamma = \langle \{1, 2\}, H, P, f_c, (u_i) \rangle$ . Dabei besteht die Menge  $Z$  der terminalen Historien aus genau den Tupeln  $\langle a_1, \dots, a_k \rangle$ , für die gilt:

- $k \in \mathbb{N} \cup \{\infty\}$ ,
- $a_1 \in \{S, N\}$  („schießen“ bzw. „nicht schießen“),
- wenn  $a_i = S$ , dann  $a_{i+1} \in \{T, V\}$  („Treffer“ bzw. „verfehlt“), für alle  $i \in \mathbb{N}$  mit  $i \geq 1$ ,
- wenn  $a_i \in \{N, V\}$ , dann  $a_{i+1} \in \{S, N\}$ , für alle  $i \in \mathbb{N}$  mit  $i \geq 1$ ,

- $a_i = T$  gdw.  $i = k$  und  $k < \infty$  (Spiel ist beendet genau dann, wenn ein Spieler getroffen hat, und in unendlichen Historien trifft keiner der beiden Spieler den anderen).

Die Menge der Historien  $H$  ist dann der Abschluss von  $Z$  unter Präfixbildung, besteht also aus genau den Historien  $\langle a_1, \dots, a_j \rangle$ ,  $j \in \mathbb{N} \cup \{\infty\}$ , die Anfangsstück eines  $h \in Z$  der Form  $h = \langle a_1, \dots, a_j, \dots, a_k \rangle$  sind. Die Spielerfunktion ist definiert durch  $P(\langle \rangle) = 1$ ,  $P(\langle S \rangle) = c$ ,  $P(\langle N \rangle) = 2$  und

$$P(\langle h, a, a' \rangle) = \begin{cases} c & \text{falls } a' = S \\ 3 - P(\langle h, a \rangle) & \text{falls } a' = N \\ 3 - P(h) & \text{falls } a' = V \end{cases}$$

für alle Historien  $h \in H$  und Aktionen  $a \in A(h)$  und  $a' \in A(\langle h, a \rangle)$  mit  $\langle h, a, a' \rangle \notin Z$ . Die Wahrscheinlichkeitsmaße an den Zufallsknoten sind

$$\begin{aligned} f_c(T|\langle h, S \rangle) &= p_{P(h)} \quad \text{und} \\ f_c(V|\langle h, S \rangle) &= 1 - f_c(T|\langle h, S \rangle) \end{aligned}$$

für alle Historien  $h$  mit  $S \in A(h)$ .

Die Nutzenfunktionen in den terminalen Historien  $h \in Z$  sind gegeben durch  $u_i(h) = 1$ , falls  $|h| = \infty$ , für  $i = 1, 2$ , und

$$u_i(\langle h', S, T \rangle) = \begin{cases} 1 & \text{falls } P(h') = i \\ -1 & \text{falls } P(h') \neq i \end{cases}$$

für  $i = 1, 2$ , falls  $|h| < \infty$  und  $h$  die Form  $\langle h', S, T \rangle$  hat.

- (b) Zeigen Sie, dass die Strategieprofile, in denen (i) kein Spieler jemals schießt bzw. in denen (ii) beide Spieler immer schießen, teilspielperfekte Gleichgewichte sind.

**Lösung:**

Dass das Strategieprofil  $s^*$ , in dem kein Spieler jemals schießt, ein teilspielperfektes Gleichgewicht ist, ist klar, denn dann ist für alle Spieler  $i = 1, 2$  und Historien  $h$  mit  $P(h) = i$  der einzige mögliche Ausgang  $O_h(s^*|_h)$  die unendliche Historie  $\langle N, N, N, \dots \rangle$  mit (erwartetem) Nutzen  $U_i|_h(O_h(s^*|_h)) = 1$ . Da  $u_i \leq 1$ , ist auch  $U_i|_h \leq 1$ , d.h. Spieler  $i$  kann durch Abweichen keinen höheren Nutzen erzielen.

Um zu zeigen, dass auch das Strategieprofil  $s^*$ , in dem beide Spieler immer schießen, ein teilspielperfektes Gleichgewicht ist, müssen wir zwei Fälle unterscheiden. Falls die Trefferwahrscheinlichkeiten  $p_1 = p_2 = 0$  sind, so ist die resultierende Historie in jedem Teilspiel  $\langle S, V, S, V, \dots \rangle$  und mit demselben Argument wie oben muss deshalb  $s^*$  ein teilspielperfektes Gleichgewicht sein.

Bleibt noch der Fall, dass  $p_1 > 0$  oder  $p_2 > 0$  gilt. Sei  $i \in \{1, 2\}$  und  $h$  eine Historie mit  $P(h) = i$ . Der erwartete Nutzen  $U_i$  von Spieler  $i$  über alle möglichen Ausgänge  $O_h(s^*|_h)$  von  $s^*|_h$  in  $\Gamma(h)$  muss wegen des rekursiven Aufbaus des Spielbaums und weil sich beide Spieler immer gleich verhalten, die folgende Gleichung erfüllen:

$$U_i = p_i + (1 - p_i)(-p_{-i} + (1 - p_{-i})U_i)$$

Löst man diese Gleichung nach  $U_i$  auf, so erhält man

$$U_i = \frac{p_i - p_{-i} + p_i p_{-i}}{p_i + p_{-i} - p_i p_{-i}}.$$

Nach der Definition des erwarteten Nutzens muss  $U_i$  für alle  $0 \leq p_i, p_{-i} \leq 1$  mit  $p_i > 0$  oder  $p_{-i} > 0$  einen Wert zwischen  $-1$  und  $1$  annehmen. Wie erwartet ist  $U_i = 1$ , falls  $p_{-i} = 0$ , und  $U_i = -1$ , falls  $p_i = 0$ .

Wir müssen noch zeigen, dass Spieler  $i$  durch Änderung seiner Strategie seinen Nutzen gegenüber  $U_i$  nicht erhöhen kann.

Der Fall, dass  $p_{-i} = 0$  ist trivial, weil dann der Nutzen von Spieler  $i$  gegeben  $s_i^*$  immer  $1$  ist und Spieler  $i$  nicht profitabel abweichen kann.

Da das Spiel einen unendlichen Horizont hat, können wir in den verbleibenden Fällen das Ein-Schritt-Abweichungs-Lemma *nicht* ohne weiteres anwenden (das Lemma und der Satz von Kuhn gelten auch für Spiele mit Zufallszügen).

Im Beweis des Ein-Schritt-Abweichungs-Lemmas wird aus einer profitablen Abweichung  $s_i^+$  von  $s_i^*|_h$ , die sich in nur *endlich vielen* Historien (nämlich entlang dem Ausgang des Spiels, wenn  $s_i^+$  gespielt wird) von  $s_i^*|_h$  unterscheidet, eine abweichende Strategie in einem geeigneten Teilspiel konstruiert, die sich nur im ersten Schritt von der alten Strategie unterscheidet.

Wenn wir wüssten, dass es immer dann, wenn es in einem Teilspiel  $\Gamma(h)$  *irgendeine* profitable Abweichung  $s_i^+$  von  $s_i^*|_h$  (insbesondere mit unendlich vielen Teilspielen  $\Gamma(\langle h, h' \rangle)$  von  $\Gamma(h)$ , in denen  $s_i^+$  die Aktion  $N$  vorschreibt, sich also von  $s_i^*|_h$  unterscheidet), auch eine profitable Abweichung  $s_i^+$  von  $s_i^*|_h$  mit nur endlich vielen Abweichungen gibt, so könnten wir von diesem  $s_i^+$  ausgehend einen Rückwärtsinduktionsbeweis analog zum Ein-Schritt-Abweichungs-Lemma und dem Satz von Kuhn führen. Die Anwendung des Ein-Schritt-Abweichungs-Lemmas bestünde dann darin, zu zeigen, dass der Nutzen  $U_i'$ , den man erhält, wenn man nur in der initialen Historie eines Teilspiels von  $S$  zu  $N$  abweicht, höchstens  $U_i$  ist. Das wäre aber wegen  $U_i' = p_{-i}(-1) + (1 - p_{-i})U_i \leq p_{-i}U_i + (1 - p_{-i})U_i = (p_{-i} + (1 - p_{-i}))U_i = U_i$  der Fall (denn  $U_i \geq -1$ ).

Bleibt also noch zu zeigen, wie aus einer profitablen Abweichung mit unendlich vielen Historien, in denen  $N$  statt  $S$  gespielt wird, eine profitable Abweichung mit nur endlich vielen Historien erzeugt werden kann, in denen  $N$  gespielt wird.

Sei  $s_i^+$  eine profitable Abweichung von  $s_i^*|_h$  in  $\Gamma(h)$  für eine Historie  $h$ . Zunächst können wir wie beim Beweis des Ein-Schritt-Abweichungs-Lemmas ohne Beschränkung der Allgemeinheit annehmen, dass alle Abweichungen auf einem Pfad im Spielbaum liegen. Es ist  $U_i(s_i^+, s_{-i}^*|_h) > U_i(s_i^*|_h, s_{-i}^*|_h)$ . Sei  $\epsilon = U_i(s_i^+, s_{-i}^*|_h) - U_i(s_i^*|_h, s_{-i}^*|_h)$ . Wegen  $p_{-i} > 0$  ist  $(1 - p_{-i}) < 1$ , d. h. es gibt eine natürliche Zahl  $m$  mit der Eigenschaft, dass  $(1 - p_{-i})^m < \epsilon/2$ . Damit fallen aber Abweichungen, die auf dem relevanten Pfad mehr als  $4m$  Schritte ( $m$  Zufallsentscheidungen, ob Spieler  $3 - i$  trifft) unterhalb von  $h$  liegen, für die Profitabilität nicht mehr ins Gewicht und man kann die Änderungen wieder rückgängig machen, ohne die Profitabilität zu beeinträchtigen. Dann bleiben aber nur endlich viele Abweichungen entlang des Pfades übrig.