

Spieltheorie

B. Nebel
Sommersemester 2009

Universität Freiburg
Institut für Informatik

Übungsblatt 6

Abgabe: Dienstag, 16. Juni 2009

Aufgabe 6.1 (Nash-Gleichgewichte und LCPs, 2 Punkte)

Zur Erinnerung: Die Nutzenmatrix des Bach-oder-Strawinsky-Spiels ist

	B	S
B	2, 1	0, 0
S	0, 0	1, 2

- (a) Spezifizieren Sie das LCP, das diesem Spiel entspricht.
- (b) Geben Sie drei verschiedene Lösungen des LCPs an und verifizieren Sie, dass es sich um Lösungen handelt.

Aufgabe 6.2 (Nash-Gleichgewichte in allgemeinen Spielen, 4 + 1 + 1 Punkte)

Neben den in der Vorlesung genannten Algorithmen zum Finden von Nash-Gleichgewichten in allgemeinen Zwei-Personen-Matrixspielen gibt es auch eine Methode, die auf einer Formulierung des Spiels als gemischt-ganzzahligem Programm (mixed integer program oder kurz MIP) basiert¹. Die MIP-Formulierung ist der LCP-Formulierung ähnlich, argumentiert aber über den *Regret*, der wie folgt definiert ist: Der Regret einer (reinen) Strategie $a_i \in A_i$ gegeben die (gemischte) Strategie $\alpha_{-i} \in \Delta(A_{-i})$ des anderen Spielers ist definiert als

$$\text{regret}_i(a_i, \alpha_{-i}) = \left(\max_{a'_i \in A_i} u_i(a'_i, \alpha_{-i}) \right) - u_i(a_i, \alpha_{-i}).$$

Das Rückrichtung des Support-Lemmas besagt, dass ein Profil (α_1, α_2) ein Nash-Gleichgewicht ist, wenn für jedes i und jede reine Strategie $a_i \in \text{supp}(\alpha_i)$ gilt, dass $a_i \in B_i(\alpha_{-i})$. Für den Regret heißt das: Wenn $a_i \in \text{supp}(\alpha_i)$, muss $\text{regret}_i(a_i, \alpha_{-i}) = 0$ sein. Die MIP-Formulierung fordert also, dass jedes reine a_i entweder nicht im Support liegt oder einen Regret von 0 hat.

Für jede reine Strategie $a_i \in A_i$ gibt es eine Boolesche Variable $f_{a_i} \in \{0, 1\}$, die anzeigt, ob $a_i \in \text{supp}(\alpha_i)$ gelten kann. Wenn $f_{a_i} = 1$, so muss $a_i \notin \text{supp}(\alpha_i)$ und damit $\alpha_i(a_i) = 0$ gelten. Wenn $f_{a_i} = 0$, so darf a_i in $\text{supp}(\alpha_i)$ liegen (muss es aber nicht) und es muss $\text{regret}_i(a_i, \alpha_{-i}) = 0$ gelten.

Der Regret einer Strategie wird mit r_{a_i} bezeichnet. Die anderen Variablen der MIP-Formulierung sind von der LCP-Formulierung bekannt: Für jeden Spieler i ist U_i der höchstmögliche erwartete Nutzen, den er gegeben α_{-i} erreichen kann. Für jede reine Strategie a_i ist $\alpha_i(a_i)$ die Wahrscheinlichkeit, mit der sie gespielt

¹<http://www.cs.cmu.edu/~sandholm/MIPNash.aaai05.pdf>

wird, und u_{a_i} der erwartete Nutzen von a_i gegen α_{-i} . Die Konstante Δ_i ist die maximale Differenz zwischen zwei Nutzenwerten von Spieler i :

$$\Delta_i = \max_{a, a' \in A} (u_i(a) - u_i(a')).$$

Die MIP-Formulierung für Zwei-Personen-Matrixspiele (ohne Optimierungsbedingung) sucht dann nach $\alpha_i(a_i), U_i, u_{a_i}, r_{a_i}$ und f_{a_i} , so dass:

$$\begin{aligned} \sum_{a_i \in A_i} \alpha_i(a_i) &= 1 && \text{für alle } i \\ u_{a_i} &= \sum_{a_{-i} \in A_{-i}} \alpha_{-i}(a_{-i}) u_i(a_i, a_{-i}) && \text{für alle } i, a_i \in A_i \\ U_i &\geq u_{a_i} && \text{für alle } i, a_i \in A_i \\ r_{a_i} &= U_i - u_{a_i} && \text{für alle } i, a_i \in A_i \\ \alpha_i(a_i) &\leq 1 - f_{a_i} && \text{für alle } i, a_i \in A_i \\ r_{a_i} &\leq \Delta_i f_{a_i} && \text{für alle } i, a_i \in A_i \end{aligned}$$

mit $\alpha_i(a_i) \geq 0, U_i \geq 0, u_{a_i} \geq 0, r_{a_i} \geq 0, f_{a_i} \in \{0, 1\}$.

Gegeben sei nun das folgende Spiel $G = \langle \{1, 2\}, (A_i), (u_i) \rangle$ mit $A_1 = \{a_1, a_2\}$ und $A_2 = \{b_1, b_2\}$ und der folgenden Auszahlungsmatrix:

	b_1	b_2
a_1	3, 4	1, 2
a_2	1, 3	2, 6

- Stellen Sie das entsprechende MIP auf und bestimmen Sie die Optimierungsbedingung für ein wohlfahrtsmaximierendes Nash-Gleichgewicht. Ein Nash-Gleichgewicht ist wohlfahrtsmaximierend, wenn die Summe der erzielten Nutzenwerte unter allen Nash-Gleichgewichten maximal ist. Lösen Sie es mittels `lp_solve`² und schicken Sie die dazugehörige Datei an Robert Mattmüller (mattmuel@informatik.uni-freiburg.de).
- Bestimmen Sie ein Nash-Gleichgewicht mit maximaler „Fairness“ und erläutern Sie kurz, wie Sie vorgegangen sind. Fairness ist definiert als minimale Nutzendifferenz.
- Bestimmen Sie ein Nash-Gleichgewicht mit maximaler Anzahl reiner Strategien in den beiden Supports (zusammen) und geben Sie die entsprechende Optimierungsbedingung des MIP an.

Die Übungsblätter dürfen und sollten in Gruppen von zwei Studenten bearbeitet werden. Bitte schreiben Sie beide Namen auf Ihre Lösung.

²Unter <http://lpsolve.sourceforge.net/5.5/integer.htm> finden Sie Hinweise zum Umgang mit Integer-Variablen.