Informatik I: Einführung in die Programmierung

8. Exkurs: Spieltheorie

Albert-Ludwigs-Universität Freiburg

UNI FREIBURG

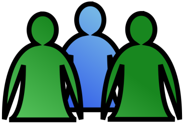
Bernhard Nebel

7. November 2017

1 Was ist Spieltheorie?

Was ist Spieltheorie?

- Spieltheorie beschäftigt sich mit Entscheidungen von rationalen Agenten in Gruppensituationen:



  - Gesellschaftsspiele
  - Entscheidungen in Politik und Wirtschaft
  - Auktionen
  - Wahlen
- Entstanden in der Mathematik (von Neumann 1928, Nash 1950).
- Heute aber auch wichtiges Hilfsmittel in der Informatik:
  - Multi-Agenten-Systeme (und KI allgemein)
  - Internet-Routing (und theoretische Informatik allgemein)
  - Internet-Auktionen

2 Strategische Spiele

- Motivation
- Nutzenmatrix
- Beispiele
- Nash-Equilibrium

Was ist Spieltheorie?
Strategische Spiele
Motivation
Nutzenmatrix
Beispiele
Nash-Equilibrium
Wiederholte strategische Spiele
Zusammenfassung & Ausblick

Strategische Spiele – Motivation

Spiele, bei denen alle Spieler gleichzeitig eine Entscheidung treffen und das Resultat sich aus der gleichzeitig getroffen Entscheidung ergibt.

Beispiele:

- „Schere-Stein-Papier“,
- „Elfmeter-Schießen“,
- Auktion mit verdeckten Geboten,
- Wahlen

Wichtig: Jeder Spieler macht sich Gedanken darüber, wie die anderen wohl spielen würden um dann zu einer Entscheidung zu kommen. Dabei weiß jeder Spieler, dass die anderen genauso vorgehen.

Der Nutzen von Entscheidungen

- Jeder Agent entscheidet über eine auszuführende Aktion (Schere, Stein, . . .)
- Der Nutzen (engl. *Utility*, *Payoff*) der Aktion ist abhängig von den Aktionen, die die anderen gewählt haben, wobei der Nutzen immer eine reelle Zahl ist:
  - Hat man selbst „Schere“ gewählt, so ist der Nutzen 1, falls der andere „Papier“ wählt.
  - Er ist 0, falls der andere auch „Schere“ wählt.
  - Er ist -1, falls der andere „Stein“ wählt.
- Rationale Agenten versuchen ihren Nutzen zu maximieren (und sonst nichts).

Nutzenmatrix

Man kann im Falle von 2 Spielern die Nutzenwerte in einer Matrix angeben.

		Spielerin 2	
		L	R
Spieler 1	T	u_1, u_2	v_1, v_2
	B	x_1, x_2	y_1, y_2

Spieler 1 ist der Zeilenspieler, der zwischen den Aktionen T und B wählen kann. Seine Nutzenwerte sind die jeweils linken Werte. Spielerin 2 ist die Spaltenspielerin. Sie kann hier zwischen den Aktionen L und R wählen. Ihre Nutzenwerte sind die jeweils rechten Werte.
Wählen die Spieler T und R, so ist die Auszahlung für Spieler 1 v_1 und für Spielerin 2 v_2.

Was ist Spieltheorie?
Strategische Spiele
Motivation
Nutzenmatrix
Beispiele
Nash-Equilibrium
Wiederholte strategische Spiele
Zusammenfassung & Ausblick

Beispiel: Elfmeterspiel

Was ist Spieltheorie?
Strategische Spiele
Motivation
Nutzenmatrix
Beispiele
Nash-Equilibrium
Wiederholte strategische Spiele
Zusammenfassung & Ausblick

		Spieler 2 (Torwart)	
		L	R
Spieler 1 (Schütze)	L	$-1, +1$	$+1, -1$
	R	$+1, -1$	$-1, +1$

Wenn Spieler 1 (der Schütze) sich für L entscheidet und der Torwart für L, dann gibt es kein Tor und Spieler 1 erhält -1 und Spieler 2 erhält $+1$. Entscheidet sich Spieler 1 für R und Spieler 2 für L, erhält der Schütze $+1$ und der Torwart -1.

Beispiel: Koordinationsspiel

UNI FREIBURG

Ein Paar geht gerne ins Kino, er schaut gerne **S**cience-Fiction-Filme, sie **B**ollywood-Filme. Sie gehen lieber zusamen ins Kino, als dass sie sich den Film alleine anschauen. Sie müssen allerdings ihre Kinokarten kaufen, ohne dass sie den anderen kontaktieren können (Handy kaputt).

		Spieler 2	
		B	S
Spieler 1	B	$1,2$	$0,0$
	S	$0,0$	$2,1$

Auch als BoS, „Bach or Strawinsky“ oder „Battle of Sexes“ bekannt.

Was ist Spieltheorie?
Strategische Spiele
Motivation
Nutzenmatrix
Beispiele
Nash-Equilibrium
Wiederholte strategische Spiele
Zusammenfassung & Ausblick

Beispiel: Gefangenendilemma

Zwei Verbrecher, die zusammen verhaftet wurden, werden einzeln verhört. Sie können mit 0-4 Jahren Gefängnis bestraft werden (Nutzenwert 4 für 0 Jahre, 3 für 1 Jahr, usw.). Wenn einer gesteht (**D**efect), während der andere schweigt (**C**ooperate), so wird ersterer frei gelassen (Wert: 4), während der andere 4 Jahre (Wert: 0) ins Gefängnis muss. Schweigen beide (**C**/**C**), müssen sie für 1 Jahr (Wert 3) ins Gefängnis. Gestehen beide, müssen sie beide für 3 Jahre (Wert 1) ins Gefängnis.

	C	D
C	$3,3$	$0,4$
D	$4,0$	$1,1$

Lösungen für strategische Spiele

- Welche Strategie sollte man spielen?
- Maximin: Das Maximum über alle Worst-Case-Werte.
  - Im Gefangenendilemma, ist der Worst-Case Wert 0 für C, 1 für D (für Spieler 1 hervorgehoben).
  - Für *BoS* und *Elfmeter* bekommen wir keine Lösung.
- Dominante Strategien: Ist eine Entscheidung immer besser, egal was der andere wählt, dann nimm diese.
  - Im Gefangenendilemma ist der Nutzen für D *immer höher* als für C (für Spieler 1 hervorgehoben).
  - Bei den anderen Spielen nicht vorhanden.

	C	D
C	3,3	0,4
D	4,0	1,1

Was ist Spieltheorie?
Strategische Spiele
Motivation
Nutzenmatrix
Beispiele
Nash-Equilibrium
Wiederholte strategische Spiele
Zusammenfassung & Ausblick

Nash-Equilibrium

- John Nash schlug vor, Kombinationen von Aktionen (*Aktionsprofile*) als Lösungen zu betrachten, bei denen sich kein Spieler durch eine Abweichung verbessern kann: **Nash-Equilibrium** (*NE*)
  - Im Gefangenendilemma, ist das einzige Nash-Equilibrium (D,D).
  - Bei BoS gibt es zwei Nash-Equilibria: (B,B), (S,S).
  - Im Elfmeterspiel gibt es kein Gleichgewicht.
  - Erweitert man die wählbaren Aktionen auf *Wahrscheinlichkeitsverteilungen* über den Aktionen, so gibt es (in endlichen strategischen Spielen) immer ein Nash-Equilibrium (Satz von Nash)!

	C	D
C	3,3	0,4
D	4,0	1,1

Was ist Spieltheorie?
Strategische Spiele
Motivation
Nutzenmatrix
Beispiele
Nash-Equilibrium
Wiederholte strategische Spiele
Zusammenfassung & Ausblick

3 Wiederholte strategische Spiele

- Wiederholte Spiele mit unsicherem Abbruch
- Strategien = Moore-Automaten
- Alternative Nash-Equilibria
- Welche Strategie ist die Beste?

Was ist Spieltheorie?
Strategische Spiele
Wiederholte strategische Spiele
Wiederholte Spiele mit unsicherem Abbruch
Strategien = Moore-Automaten
Alternative Nash-Equilibria
Welche Strategie ist die Beste?
Zusammenfassung & Ausblick

Das Dilemma der Gefangenen

- Wählt man die Maximin-Aktion um das schlechtest mögliche Ergebnis zu maximieren, erhält man D.
- Außerdem ist D eine dominante Aktion, da der Nutzen, egal was der andere Spieler macht, immer maximal ist.
- Außerdem ist (D,D) ein Nash-Equilibirum.
- Wünschenswert wäre aber natürlich für beide Spieler, dass (C,C) gespielt wird.
- Tatsächlich spielen Menschen auch oft (C,C):
  - weil sie Erfahrungen gesammelt haben,
  - weil sie Vertrauen in den anderen Spieler haben,
  - weil sie sich vor Bestrafung fürchten.

Was ist Spieltheorie?
Strategische Spiele
Wiederholte strategische Spiele
Wiederholte Spiele mit unsicherem Abbruch
Strategien = Moore-Automaten
Alternative Nash-Equilibria
Welche Strategie ist die Beste?
Zusammenfassung & Ausblick

Wiederholte Spiele

- Um die Zeit- und Erfahrungsaspekte mit einzubringen, kann man die Spiele mehrfach spielen lassen.
- Also z.B. 10 Runden des Gefangenendilemmas hintereinander spielen.
- Was wäre ein Nash-Equilibrium für dieses neue Spiel?
  - Im letzten Spiel ist das einzige NE das bekannte (D,D).
  - Dann ist allerdings im vorletzten Spiel auch (D,D) das einzige NE …
- Wir könnten allerdings nach jeder Runde mit einer Wahrscheinlichkeit p das Spiel beenden, dann gibt es keine letzte Runde!
- Statt festen Nutzenwerten müssen wir jetzt den Erwartungswert des Nutzens (= erwarteten Nutzen) bestimmen.

Unsicherer Abbruch

Was ist Spieltheorie?
Strategische Spiele
Wiederholte strategische Spiele
Wiederholte Spiele mit unsicherem Abbruch
Strategien = Moore-Automaten
Alternative Nash-Equilibria
Welche Strategie ist die Beste?
Zusammenfassung & Ausblick

- Wir nehmen an, dass nach jeder Runde mit einer Wahrscheinlichkeit von $0 < p < 1$ das Spiel beendet wird.
- Was ist der Erwartungswert für die Anzahl der Runden?
- $1p + 2(1-p)p + 3(1-p)^2 p + \ldots = \sum_{i=1}^{\infty} ip(1-p)^{i-1} = ?$
- Welchen erwarteten Nutzen erhält ein Spieler, wenn er in jeder Runde den konstanten Nutzen u bekommt?
- $u + u(1-p) + u(1-p)^2 + \ldots = \sum_{i=0}^{\infty} u(1-p)^i = ?$
- Dazu müssen wir wissen, welchen Wert die unendlichen Reihen $\sum_{i=1}^{\infty} ix^i$ und $\sum_{i=0}^{\infty} x^i$ (für $|x| < 1$) haben.

Zwei wichtige Reihen

Für $0 < x < 1$:

$$
\begin{array}{rcll}
s & = & 1 + x + x^2 + \ldots & \\
 & = & 1 + x(1 + x + x^2 + \ldots) & (x \text{ ausgeklammert}) \\
 & = & 1 + xs & (s \text{ eingesetzt}) \\
(1-x)s & = & 1 & (-xs \text{ und } s \text{ ausgeklammert}) \\
s & = & \frac{1}{(1-x)} & (\text{durch } (1-x) \text{ geteilt}) \\
\sum_{i=0}^{\infty} x^i & = & \frac{1}{(1-x)} & \\
\sum_{i=0}^{\infty} ix^{i-1} & = & \frac{1}{(1-x)^2} & (\text{auf beiden Seiten differenzieren}) \\
\sum_{i=1}^{\infty} ix^i & = & \frac{x}{(1-x)^2} & (\text{mit } x \text{ multiplizieren})
\end{array}
$$

Erwartete Spiellänge

$\sum_{i=1}^{\infty} ix^i = \frac{x}{(1-x)^2}$ angewandt auf die erwartete Spiellänge:

$$\begin{aligned}
\sum_{i=1}^{\infty} ip(1-p)^{i-1} &= \frac{p}{1-p}\sum_{i=1}^{\infty} i(1-p)^i \\
&= \frac{p}{(1-p)} \cdot \frac{(1-p)}{(1-(1-p))^2} \\
&= \frac{p}{(1-p)} \cdot \frac{(1-p)}{p^2} \\
&= \frac{1}{p}
\end{aligned}$$

Z.B. für $p = 1/10$ ist die erwartete Spiellänge 10.

Was ist Spieltheorie?
Strategische Spiele
Wiederholte strategische Spiele
Wiederholte Spiele mit unsicherem Abbruch
Strategien = Moore-Automaten
Alternative Nash-Equilibria
Welche Strategie ist die Beste?
Zusammenfassung & Ausblick

Erwarteter Nutzen

$\sum_{i=0}^{\infty} x^i = \frac{1}{(1-x)}$ angewandt auf den erwarteten Nutzen:

$$\begin{aligned}
\sum_{i=0}^{\infty} u(1-p)^i &= u\sum_{i=0}^{\infty}(1-p)^i \\
&= u\frac{1}{1-(1-p)} \\
&= \frac{u}{p}
\end{aligned}$$

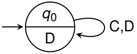
Z.B. für $p = 1/10$ und $u = 4$ ist dann der erwartete Nutzen: 40.

Strategien in wiederholten Spielen

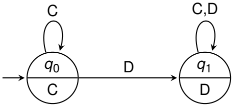
- Wie kann man bei wiederholten Spielen Strategien formulieren?
- Diese müssen potentiell unendlich sein.
- Endliche Automaten mit Ausgabe (= Moore-Automaten) wären da eine Möglichkeit:
  - *Unkooperativ:* Egal was der andere gespielt hat, spiele immer D.
  - *Kooperativ:* Egal was der andere gespielt hat, spiele immer C.
  - *Grimmig:* Spiele C bis der andere das erste Mal D spielt, dann spiele immer D.
  - *Tit-for-tat:* Spiele anfangs C und antworte dann auf jedes D mit D und auf C mit C.
  - *Bipolar/Troll:* Startend mit D, spiele abwechselnd C und D.

Strategien als Automaten

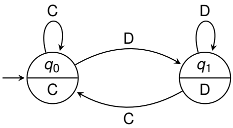
Unkooperativ



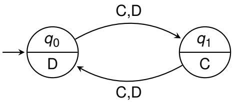
Grimmig



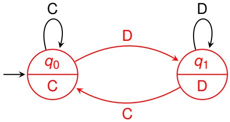
Tit-for-tat

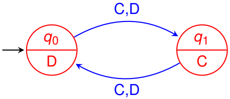


Bipolar



Spieler 1 spielt Tit-for-tat, Spielerin 2 spielt bipolar. 4 Runden.





Runde	Aktionen	Nutzen	Akkumuliert
1	(C,D)	(0,4)	(0,4)
2	(D,C)	(4,0)	(4,4)
3	(C,D)	(0,4)	(4,8)
4	(D,C)	(4,0)	(8,8)

Was ist Spieltheorie?
Strategische Spiele
Wiederholte strategische Spiele
Wiederholte Spiele mit unsicherem Abbruch
Strategien = Moore-Automaten
Alternative Nash-Equilibria
Welche Strategie ist die Beste?
Zusammenfassung & Ausblick

Nash-Equilibria für das wiederholte Gefangenendilemma

- Die Kombination (Unkooperativ,Unkooperativ) ist ein Nash-Equilibrium:
  - Spielverlauf: (D,D), (D,D), ...
  - Erwarteter Nutzen: (1/p,1/p)
  - Kann ein Spieler sich durch Abweichung von der unkooperativen Strategie verbessern?
  - Weicht ein Spieler (z.B. Spieler 1) in einer Runde ab, so verliert er einen Nutzenpunkt.
- Aber es gibt jetzt auch alternative NE.
- Die Kombination (Grimmig, Grimmig) sieht sehr viel versprechend aus.
- Erwarteter Nutzen von (Grimmig, Grimmig) ist $(3/p, 3/p)$.
- Kann ein Spieler durch Abweichung seinen Nutzen steigern?

Was ist Spieltheorie?

Strategische Spiele

Wiederholte strategische Spiele

Wiederholte Spiele mit unsicherem Abbruch

Strategien = Moore-Automaten

Alternative Nash-Equilibria

Welche Strategie ist die Beste?

Zusammenfassung & Ausblick

Ist eine Abweichung von (Grimmig, Grimmig) sinnvoll?

- Die Strategiekombination (Grimmig,Grimmig) führt zu dem Spielverlauf $(C,C),(C,C),(C,C),\ldots$
- Kann Spielerin 2 abweichen um einen höheren Nutzenwert zu erhalten?
- Wenn sie in der Runde k auf D abweicht, muss sie danach immer D spielen, ansonsten verschenkt sie Punkte.
- Wie hoch ist die erwartete Nutzensteigerung ab Schritt k?

$$
\begin{aligned}
+1-2\cdot(1-p)-2\cdot(1-p)^2-\ldots \quad &= \quad 1-2\cdot\textstyle\sum_{i=1}^{\infty}(1-p)^i \\
&= \quad 3-2\cdot\textstyle\sum_{i=0}^{\infty}(1-p)^i \\
&= \quad 3-2\cdot\frac{1}{1-(1-p)} \\
&= \quad 3-\frac{2}{p}
\end{aligned}
$$

- Falls $p=\frac{2}{3}$ oder kleiner ist, gibt es keinen Nutzenzuwachs.
- $(Grimmig, Grimmig)$ ist damit ein NE für $p\leq\frac{2}{3}$.

Tit-for-tat führt auch zu einem Nash-Equilibrium

- Spielverlauf für (Tit-for-tat, Tit-for-tat) genauso wie für (Grimmig, Grimmig).
- Wir betrachten den Fall, dass ein Spieler in der Runde k einmal nach D wechselt und danach wieder C spielt.
- Nutzensteigerung im Schritt k:

$$+1-2(1-p) \quad = \quad 2p-1$$

- D.h. für alle $p \leq 1/2$ gibt es keine Nutzensteigerung.
- Wird öfter als einmal bei $p \leq 1/2$ abgewichen, kann es keine Nutzensteigerung geben.
- D.h. auch (Tit-for-tat,Tit-for-tat) ist ein Nash-Equilibrium – so lange die Abbruchwahrscheinlichkeit klein genug ist.

Was ist Spieltheorie?
Strategische Spiele
Wiederholte strategische Spiele
Wiederholte Spiele mit unsicherem Abbruch
Strategien = Moore-Automaten
Alternative Nash-Equilibria
Welche Strategie ist die Beste?
Zusammenfassung & Ausblick

Und jetzt?

- Die unkooperative Strategie ist weder dominante Strategie noch ist sie die einzige Equilibriumsstrategie.
- D.h. es gibt rationale Kooperationsstrategien (mit höheren Auszahlungen).
- Allerdings gibt es sehr viele NEs!
- Welche Strategie sollte man spielen?
- Das kann man ja auch empirisch bestimmen: Sie dürfen ihre eigenen Strategien entwerfen, die dann gegeneinander im Wettkampf antreten um möglichst hohe Auszahlungen zu erhalten.
- Wie man einen Moore-Automat als Python-Programm realisiert, haben wir ja bereits gesehen.

Wetere Beispielstrategien

- Strenges Tit-for-tat: Bestrafe 2- oder 3-mal, bevor zur Kooperation zurück gekehrt wird.
- Missvertrauen: Beginne mit D und spiele dann Tit-for-tat.
- Majorität: Spiele den meistbenutzten Zug des Gegners (bei Gleicheit Kooperation).
- Schnorrer: Probiere irgendwann D und mache weiter damit, solange der andere C spielt, ansonsten Tit-for-Tat.
- Spätes Abweichen: Weiche in einer sehr späten Runde ab und spiele D (in der Hoffnung, dass das die letzte Runde ist).
- Einige dieser Strategien sind NE, andere nicht. Aber darauf kommt es ja gar nicht an, wenn man gegen viele verschiedene Agenten spielen muss …

4 Zusammenfassung & Ausblick

Zusammenfassung & Ausblick

- Spieltheorie beschäftigt sich mit Entscheidungen von rationalen Agenten in Gruppen.
- Spieltheorie ist in der Mathematik entstanden, ist mittlerweile aber ein wichtiges Werkzeug innerhalb der Informatik.
- Strategische Spiele sind die einfachsten Spiele, die untersucht werden.
- Das Gefangenendilemma modelliert das Problem, dass Kooperation zwar sinnvoll ist, aber unkooperatives Verhalten höheren Nutzen bringen kann.
- Wiederholte Spiele bringen den Aspekt von Zeit und Erfahrungen in die spieltheoretische Analyse.
- Im wiederholten Gefangenendilemma existieren rationale Kooperationsstrategien, aber es existieren sehr viele davon.