

# Knowledge Engineering und Lernen in Spielen



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

## Die Evolution der Kooperation und Strategien des Iterierten Gefangenendilemma

Prof. Fürnkranz  
Knowledge Engineering Group



Christoph Thörle

# Agenda

---



- **Einleitung**
- **Das Gefangenendilemma**
- **Iteriertes Gefangendilemma**
- **Strategien für das iterierte GD**
- **Die Suche nach der Optimalen Strategie**
  - + **Axelrods Computertunier**
  - + **Heuristische Strategiesuche**
  - + **Andere Ansätze: Multi Objective Approach**
- **Fazit/Ausblick**
- **Quellen**



- Die Evolution basiert stets auf dem Kampf um Leben und dem Überleben des Stärksten („survival of the fittest“)
- Bis 1960 wurde in der Evolutionstheorie eine mögliche Kooperation zwischen Individuen nicht beachtet
- In der Evolution streben genetische Ausprägungen danach ihren Fortbestand zu sichern durch Stärkung und Erhöhung der Überlebenschancen der Ausprägung und deren Merkmale

- In den meisten Fällen der Kooperation in der Natur profitieren beide Individuen oder Individuengruppen von der Zusammenarbeit, z.B. Gemeinsames Brüten bei Vögeln, Jagen bei Raubtieren..
- Wenn interagierende Individuen ausreichen nah verwandt sind, kann Altruismus das Überleben der genetischen Gemeinsamen Merkmale sichern, z.B. Der Stachel der Honigbiene
- Das Problem bei der Kooperation liegt darin, dass ein Individuum von der Kooperation profitieren kann, jedoch ist jeder noch besser gestellt, wenn er die Kooperationsbereitschaft des anderen Individuum ausnutzt
- Diese spieltheoretische Problemstellung ist durch das Gefangenendilemma modelliert...

# Das Gefangenendilemma



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

- Zwei-Personen-Nicht-Nullsummenspiel
- Wurde in den 1950er Jahren formuliert
- Beschreibt soziales Dilemma, das zeigt wie individuell rationale Entscheidungen zu kollektiv schlechteren Ergebnissen führen kann
- Dilemma: Nicht kooperieren eines Spielers ist rational, da egoistisch, jedoch kein Nullsummenspiel, somit kann nicht nur einer gewinnen -> Maximaler Gesamt-Gewinn nur durch Kooperation

# Das Gefangenendilemma



		<u>Player B</u>	
		C Cooperation	D Defection
<u>Player A</u>	C Cooperation	R=3 Reward for mutual cooperation	S=0 Sucker's payoff
	D Defection	T=5 Temptation to defect	P=1 Punishment for mutual defection

Die Grafik zeigt den Nutzen für Spieler A in den verschiedenen Situationen.

Das Spiel ist durch  $T > R > P > S$  und  $R > (T+S)/2$  definiert.

Außerdem reizt  $T > R$  und  $P > S$  die Spieler zur Defektion an.

# Das Gefangenendilemma

- Bei einem einmal gespielten Gefangenendilemma, kann keine andere Entscheidung die Defektion eines Spielers ausnutzen und ist hier somit die optimale Strategie
- In der Natur kommt es jedoch vor, dass sich Individuen mehrmals begegnen und erinnern sich etwa an ein vorheriges Treffen und das Verhalten des anderen
- Das Gefangendilemma wird zum iterierten Gefangenendilemma, man kann also eine vorherige Kooperation belohnen oder eine Defäktion bestrafen.

# Iteriertes Gefangenendilemma



- Annahme bei möglichem unendlichen Spiel: Mit der Wahrscheinlichkeit  $w$  treffen sich die Individuen wieder.
- Um der Unvorhersehbarkeit eines Wiedertreffens für die Strategieentwicklung Rechnung zu tragen, wird die Auszahlung der 1. Runde normal, die der  $i+1$ -ten Runde ( $i \in \mathbb{N}$ ) mit  $w^i$  ( $0 \leq w \in \mathbb{R}$ ) gewichtet. Im weiteren sei:  $0 < w \leq 1$
- Verschiedene Wahrscheinlichkeiten haben großen Einfluss auf die Wahl der besten Strategie, wenn  $w$  nahe 0 ist, lohnt es sich wie beim 1-Stufigen GD nicht mehr zu kooperieren
- Die Auszahlung für 3 Kooperationen wäre nach diesem Modell für beide Spieler  $R + wR + w^2R$



# Strategien für das Iterierte GD

- „ALL C“ und „ALL D“
  - sture Strategien
  - defektieren bzw. kooperieren immer
  - ALL C lässt sich immer ausbeuten
  - ALL D verscherzt es sich mit ziemlich allen Strategien
  
- „RANDOM“
  - gemischte Strategie
  - defektiert mit  $p$  und kooperiert mit  $1 - p$
  - ALWAYS D ist RANDOM mit  $p = 1$
  - ALWAYS C ist RANDOM mit  $p = 0$

# Strategien für das Iterierte GD

- „FRIEDMAN“
  - Trigger-Strategie
  - Beginn mit Kooperation
  - kooperiert, solange der andere kooperiert
  - rachsüchtige Strategie: bei der ersten Defektion übt FRIEDMAN ewige Vergeltung
  
- „TIT FOR TAT“
  - Beginn mit Kooperation
  - erwidert sowohl Kooperation wie auch Defektion
  - Gedächtnis von einer Zeiteinheit
  - TIT FOR  $n$  TATS: Vergeltung erst nach  $n$  Defektionen
  - Suspicious TIT FOR TAT: Beginn mit Defektion

# Strategien für das Iterierte GD

- „Hard Majority“
  - Defektiert beim ersten Zug und defektiert wenn die Anzahl der Defektionen des Gegners größergleich der Anzahl der Kooperationen des Gegners ist
- „Soft Majority“
  - Kooperiert beim ersten Zug und kooperiert wenn die Anzahl der Kooperationen des Gegners größergleich der Anzahl der Defektionen des Gegners ist

# Strategien für das Iterierte GD

- „JOSS“
  - TIT FOR TAT-Derivat (Versuch einer Verbesserung)
  - versucht, mit gelegentlicher Defektion ( $p = 0.1$ ) ungestraft davonzukommen
- „M TIT FOR TAT“
  - Defektierst nach einer Defektion  $m$  mal.
- „Pavlov“
  - Kooperiert beim ersten Zug und defektiert nur, wenn beide Spieler beim letzten Zug nicht kooperiert haben.



- **Frage nach der besten Strategie** ist vom Umfeld, d.h. **von den anderen Strategien abhängig**
- **Beispiel:** eine Strategie spielt ALWAYS D, eine andere FRIEDMAN. Was ist die beste Strategie? Die Antwort auf ALWAYS D ist ALWAYS D und die auf FRIEDMAN ist immer zu kooperieren.



## Axelrod's Computertunier:

- Im Vorfeld der Veröffentlichung von „The Evolution of Cooperation“ hat Robert Axelrod anfang der 1980er ein Computertunier zum iterierten GD veranstaltet.
- **Computerprogramme** mit verschiedenen Strategien von Personen aus unterschiedlichsten Fachgebieten (Mathematik, Psychologie, Wirtschaft, Biologie,...) aus mehreren Ländern der Welt
- die Strategien spielen **jeder gegen jeden** eine **zufällige Anzahl von Runden** (im Mittel etwa 150)



## Die Ergebnisse des Turniers:

- **TIT FOR TAT** war die **beste Strategie im Turnier** und gleichzeitig auch die einfachste
- zwischen den erfolgreichen Strategien lassen sich **Gemeinsamkeiten** finden



## TIT FOR TAT

- Entscheidungsregel: Kooperiere im ersten Zug und tue dann das, was der andere Spieler im Zug zuvor getan hat („Auge um Auge, Zahn um Zahn“)
- Freundlichkeit ermöglicht Kooperation
- Bestrafung von Defektion schützt vor Ausnutzung
- kombiniert perfekt gute Eigenschaften der Strategien





## TIT FOR TAT- Varianten: JOSS

- versucht mit einer gelegentlichen Defektion ungestraft davonzukommen (Defektion in 10 % der Fälle nach einer Kooperation des anderen Spielers)
- **Vorteil:** schneidet bei sehr gutmütigen Strategien besser ab
- **Nachteil:** führt im Spiel mit TIT FOR TAT oder sich selber zu Echos, und zwei Echos führen zur permanenten Defektion beider Seiten



## TIT FOR TAT- Varianten: TIT FOR n TAT

- erst nach n Defektionen wird defektiert
- **Vorteil:** Größere Toleranz führt in einigen Fällen zu mehr Erfolg (JOSS)
- **Nachteil:** leichter ausnutzbar (z.B. von RANDOM)

# Die Suche nach der Optimalen Strategie



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

TIT FOR TAT- Varianten:  $m$  TIT FOR TAT

- Defektiere nach einer Defektion  $m$  mal
- **Vorteil:** Nicht so leicht ausnutzbar
- **Nachteil:** gegen JOSS noch schlimmer



## Eigenschaften guter Strategien:

- (1) Sei nicht neidisch.
- (2) Defektiere nicht als erster.
- (3) Erwidere sowohl Kooperation als auch Defektion.
- (4) Sei nicht zu raffiniert.



## (1) Sei nicht neidisch

- Iteriertes Gefangenendilemma ist **kein Nullsummenspiel!**
- konzentriere dich auf deinen eigenen Erfolg und nicht darauf, besser als der andere zu sein
- Man kann zusammen Erfolg haben. Der eigene Erfolg ist sogar oft an den Erfolg des anderen gebunden.



## (2) Defektiere nicht als erster.

- die Antwort auf unbegründete Defektion ist **Vergeltung**
- **Schaden für beide** ( $t + p < 2r$ )
- **die besten 8 Strategien** waren freundlich (Beginn mit Kooperation), die schlechtesten 8 unfreundlich (Beginn mit Defektion)



## (3) Erwidere sowohl Kooperation als auch Defektion.

- Ziel ist Kooperation
- übertriebene Rachsucht führt zu gegenseitiger Defektion (FRIEDMAN als extremes Beispiel)
- Verzeihung als Mittel um Eskalation vorzubeugen (**vgl. JOSS gegen TIT FOR TAT**)
- große Toleranz von Defektion verleitet die andere Strategie dazu sich immer mehr zu erlauben
- die Rangfolge unter den freundlichen Regeln ist davon abhängig, wie sie mit den unfreundlichen Strategien klarkommen



## (4) Sei nicht zu raffiniert.

- undurchschaubare Strategien sind RANDOM sehr ähnlich
- optimale Strategie gegen RANDOM ist ALWAYS D!
- einfache, durchschaubare Strategien ermöglichen es dem anderen sich anzupassen





## Exkurs: Vergleich mit Schach

- Nullsummenspiel
- Maximierung des eigenen Gewinns führt zu Verlusten des Gegners
- Kooperation keine gewinnbringende Lösung
- versuche möglichst undurchschaubar zu sein, um den anderen zu Fehlern zu verleiten



## Heuristische Strategiesuche

- Mit **Heuristischen Verfahren** werden bekannte Strategien untersucht um mögliche Verbesserungen zu erhalten
- **Genetische Algorithmen**, die mittels Mutation und Crossover versuchen lokale Optima bei einer Strategie zu verlassen
  - **Mutation**: z.B. JOSS Mutation von TIT FOR TAT
  - **Crossover**: Kombination von Strategien z.B. FRIEDMAN
- **Simulated Annealing**: Mit zunehmender Rundenzahl wird die Wahrscheinlichkeit für Abweichungen in den Entscheidungen der eingesetzten Strategie erhöht

# Die Suche nach der Optimalen Strategie



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

Andere Ansätze: Multi-Objective Approach

Ansatz von *Shashi Mittal* und *Kalyanmoy Deb*

*Indian Institute of Technology, Kanpur, Indien:*

IGD wird nicht mehr als single-objective game betrachtet (Maximierung der eigenen Punkte, **oder** Minimierung der gegnerischen Punkte), sondern nun als multi-objective game (Maximierung der eigenen Punkte, **sowie** Minimierung der gegnerischen Punkte)

# Die Suche nach der Optimalen Strategie



- Ergebnisse von S. Mittal und K. Deb:

Player	Average score
Tit for Tat	387
Soft Majority	379
Tit for two tats	379
Spiteful	376
Hard Tit For Tat	370
Always Cooperate	359
Periodic Player CCD	354
Naive Prober	353
Pavlov	351
Periodic Player CD	351
Remorseful Prober	351
Random Player	323
Hard Majority	317
Suspicious Tit for Tat	310
Periodic Player DDC	309
Always Defect	305

Single-objective Approach

Player	Average score
Strategy MO	448
Tit for Tat	391
Hard Tit For Tat	375
Soft Majority	370
Tit for two tats	370
Spiteful	363
Naive Prober	358
Remorseful Prober	344
Always Cooperate	337
Periodic Player CCD	336
Periodic Player CD	334
Pavlov	334
Suspicious Tit for Tat	319
Hard Majority	312
Random Player	310
Periodic Player DDC	296
Always Defect	296

Multi-Objective Approach

- Gefangenendilemma ist formalistisches Modell für Kooperation bei Reziprozität
- Die Frage nach der Optimalen Strategie ist nicht eindeutig beantwortbar, da die Strategie immer von verschiedenen Einflüssen der Umwelt abhängt
- Verschiedene Ansätze in der Informatik bilden einige Szenarien bei der Kooperation von Individuen in der Evolution ab

# Quellen

The Evolution of Cooperation, (Axelrod, Robert; Hamilton, William D.),  
Science 211: 1390–96, 1981.

The Evolution of Strategies in the Iterated Prisoner's Dilemma (Robert Axelrod), in Lawrence Davis (ed.), Genetic Algorithms and Simulated Annealing, London: Pitman, and Los Altos, CA: Morgan Kaufman, 1987, pp. 32-41.

Optimal Strategies of the Iterated Prisoner's Dilemma Problem for Multiple Conflicting Objectives, (Mittal, Shashi and Deb, Kalyanmoy), Boston, 2006

The Selfish Gene (Dawkins, Richard), Oxford, 1989 pp 202-231.

# Sind noch Fragen offen?



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

