Spieltheorie und die Modellierung sozialer Interaktionen

Seminar Artificial Intelligence, Artificial Life, Artificial Societies - Agentenbasierte Modellierung komplexen Verhaltens

Sommersemester 2000

Manuel Klimek[1], Ernst Bachmann[2]

[1]email: klimek@fmi.uni-passau.de
[2]email: bachmann@fmi.uni-passau.de

Inhaltsverzeichnis

1 Die Kooperation

Für uns Menschen ist die Kooperation in Familie, Staat und Bündnissen aller Art seit jeher die ideale Form des Zusammenlebens. Oft genug wird der Grund für die Überlegenheit der menschlichen Rasse in der Entwicklung komplexer Kommunikationsmechanismen gesehen, deren Nutzung die grundlegenste Art von Kooperation darstellt: Den Austausch von Informationen.

Doch werden auch warnende Stimmen laut, die den Verfall der Werte und somit die Abnahme der Kooperationsbereitschaft anprangern. Sehen wir unsere Gesellschaft auf dem Datenhighway in eine Welt von Egoisten und Informationsmonopolisten rasen, so stellt sich die Frage, wie die Kooperation unter rationalen, nutzenoptimierenden Wesen gefördert werden kann. Da die Menschen heute mehr denn je ihre Entscheidungen statt auf moralische Verpflichtungen auf quantitative Berechnungen stützen, liegt der Versuch nahe, ihr Verhalten auch durch mathematische Modelle zu beschreiben. Ein solcher Ansatz soll hier im Folgenden beschrieben werden.

2 Theoretisierung der Kooperation mit Hilfe des Gefangenendilemmas

Um theoretische Erkenntnisse über die Kooperationsbereitschaft von Individuen zu erhalten, muss das zugrundeliegende Problem zuerst genau definiert und mathematisch modelliert werden. Dazu gibt es natürlich eine Vielzahl von möglichen Ansätzen. Eine sehr interessante Möglichkeit ist die Modellierung durch das Gefangenendilemma [1], welches von A.W.Tucker formalisiert wurde.

2.1 Das Gefangenendilemma

Die Grundidee des Gefangenendilemmas ist es, eine Situation zu beschreiben, in der sich zwei Individuen unabhängig voneinander für eine von zwei möglichen Handlungsweisen entscheiden müssen. Diese machen sich bei dem jeweils anderen Spieler als Kooperation bzw. Nichtkooperation (Defektion) bemerkbar.

Auf eine solche Situation trifft man als Mensch alltäglich. Um nun den Erfolg des Handelns messen zu können, gibt es je nach Entscheidung der Mitspieler gewisse Erfolgsbestätigungen, deren Restriktionen am folgenden Beispiel deutlich gemacht werden soll:

Klaus und Detlef begehen gemeinsam ein Verbrechen. Sie werden vom wachsamen Auge des Gesetzes erfasst und schließlich getrennt voneinander verhört. Jedem der beiden schlägt der Staatsanwalt einen Deal vor: Gestehen beide, so werden mildernde Umstände berücksichtigt und jeder bekommt 4 Jahre. Wenn nur einer gesteht, wird dieser freigesprochen, doch der andere sitzt die vollen 5 Jahre. Sollten beide nicht gestehen, so reichen die Indizien, um beide für 2 Jahre hinter Gitter zu bringen.

Dieses Problem kann nun wie folgt in einer Payoff-Matrix dargestellt werden:

[1] 1950 von Merill Flood und Melvin Dresher erfunden

Klaus / Detlef	Schweigen	Gestehen
Schweigen	(-2; -2)	(-5; 0)
Gestehen	(0; -5)	(-4; -4)

Allgemein können für ein Spiel dieser Art folgende Parameter gesetzt werden:

- R (= reward)
  Die Belohnung für Kooperation. Hier ist R gleich den zwei Jahren Gefängnis bei beiderseitigem Schweigen.

- T (= temptation)
  Die Versuchung zu defektieren. In diesem Beispiel ist T gleich dem Freispruch.

- S (= sucker's payoff)
  Die Bestrafung für den Gutgläubigen. Hier die Höchststrafe von 5 Jahren.

- P (= punishment)
  Die Bestrafung für zwei Dickköpfe. Beide müssen 4 Jahre ins Gefängnis.

In der Matrixschreibweise ergibt sich allgemein für zwei Spieler A und B

A / B	Kooperation	Defektion
Kooperation	$(R; R)$	$(S; T)$
Defektion	$(T; S)$	$(P; P)$

Mit den Nebenbedingungen:

1. $T > R > P > S$
   Diese Zusammensetzung ist leicht einsichtig, führt aber zu einer interessanten Schlussfolgerung:

   Wenn das Spiel nur einmal gespielt wird, so ist defektion die dominierende Strategie. Denn geht man davon aus, dass der Gegner Kooperiert, so erhält man wenn man selbst kooperiert R, wenn man defektiert jedoch T. Aus $T > R$ folgt nun sofort, dass in diesem Fall defektion die bessere Wahl ist. Wenn jedoch der Gegner defektiert und man selbst kooperiert, so wird man mit S bestraft, wogegen man durch defektion immerhin noch P erhält. Also ist es beim einmaligen Spiel immer besser zu defektieren, obwohl bei beiderseitiger Kooperation beide Spieler bessergestellt wären. Aus diesem Grund interessiert im folgenden nur noch das iterierte Gefangenendilemma.

2. $R > \frac{1}{2}(T + S)$
   Um das iterierte Gefangenendilemma interessant zu halten, ist festgelegt, dass dauernde Kooperation besser sein muss als wechselseitige Ausbeutung. Wenn beide Spieler immer kooperieren, erhält jeder pro Zug R Punkte. Falls jedoch immer einen Zug A defektiert und B kooperiert, und im nächsten dann B defektiert und A kooperiert, so erhält jeder Spieler immer abwechselnd S und T, was im Mittel pro Zug $\frac{1}{2}(T + S)$ ergibt.

Es ist auch intuitiv einsichtig, dass ein Individuum einen Gewinn in der Zukunft anders bewerten kann, als einen Gewinn in der Gegenwart. Meist ist es wohl so, dass sofortige Gewinne den eventuell eintreffenden Gewinnen vorgezogen werden, und dass auch ein sofortiger Verlust unangenehmer ist als ein in der Zukunft erwarteter. Zu diesem Zweck wird ein Diskontparameter

w eingeführt, der angibt, wie stark der folgende Zug im Vergleich zum aktuellen wiegt. Für die Summe E der Punkte eines Spiels über n Züge ergibt sich dann bei E_i Punkten im i-ten Zug:

$$E = \sum_{i=0}^{n} w^i \cdot E_{i+1}$$

Falls z.B. beide Spieler stets kooperieren, und das Spiel unendlich fortgesetzt wird, so erhält jeder Spieler konstant R und es ergibt sich

$$E = R \cdot \sum_{i=0}^{\infty} w^i = R \cdot \frac{1}{1-w}$$

2.2 Strategien für das iterierte Gefangenendilemma

Eine Strategie gibt eine Regel an, mit deren Hilfe man in jedem Zug des iterierten Gefangenendilemmas die Entscheidung für den nächsten Zug bestimmen kann. Diese Entscheidung beruht auf einem Gedächtnis über den vergangenen Verlauf des Spiels, welches je nach Regel von 0 bis n Zügen im n-ten Zug reichen kann. Die Regel muss nicht deterministisch sein, d.h. der nächste Zug kann beispielsweise auch komplett zufällig gewählt werden.
Eine grundsätzliches Bewertungskriterium zur Einteilung von Algorithmen ist die **Freundlichkeit**. Ein Algorithmus ist **freundlich**, wenn er nie defektiert, bevor der Mitspieler mindestens einmal defektiert hat. Wie sich Freundlichkeit auf die Bewertung der Algorithmen auswirkt, wird sich später noch zeigen.

2.2.1 Permanente Defektion (Kooperation)

- **Gedächtnis:** 0
- **Algorithmus:** Defektiere (Kooperiere).
- **Bemerkung:**
  Permanente Kooperation ist offensichtlich freundlich, permanente Defektion nicht.

2.2.2 Random(p)

- **Gedächtnis:** 0
- **Algorithmus:** Defektiere mit Wahrscheinlichkeit p.
- **Bemerkung:**
  Random ist im Allgemeinen nicht freundlich, auch wenn es sich im Einzelfall wie eine freundliche Strategie verhalten kann.

2.2.3 Permanente Vergeltung

- **Gedächtnis:** 1
- **Algorithmus:** Falls der Gegner oder man selbst im letzten Zug defektiert hat, so defektiere, ansonsten kooperiere.

- **Bemerkung:**
  Der Algorithmus lässt sich etwas informeller auch wie folgt beschreiben: Kooperiere, bis der Gegner zum ersten Mal defektiert, dann defektiere immer.

2.2.4 Tit for Tat

- **Gedächtnis:** 1

- **Algorithmus:** Kooperiere (Defektiere) im ersten Zug. Handle danach immer wie der Gegner im letzten Zug.

- **Bemerkung:**
  Hier wird das Sprichwort "Wie du mir, so ich dir" direkt umgesetzt. Da man im ersten Zug keine Erfahrungswerte über den Gegner besitzt, gibt es eine freundliche und eine restriktive Variante (Defektion / Kooperation im ersten Zug).

2.2.5 Gradual

- **Gedächtnis:** n (bei n Zügen)

- **Algorithmus:** Kooperiere im ersten Zug. Beantworte die erste Defektion des Gegners mit einer Defektion und zwei Kooperationen, die n-te Defektion mit n Defektionen und zwei Kooperationen.

- **Bemerkung:**
  Diese Strategie wurde von Mathieu und Delahaye speziell für das iterierte Gefangenendilemma entwickelt und baut auf einer Reihe von Erkenntnissen auf, die im Folgenden noch genauer vorgestellt werden.

2.3 Wie erkenne ich erfolgreiches Verhalten?

Das Gefangenendilemma ist ein klassisches Nicht-Nullsummen-Spiel. Bei einem Nullsummenspiel kann ein Spieler nur dann gewinnen, wenn der andere Spieler verliert (z.B. Schach). Dies führt dazu, dass Spieler A weiß, dass der Gegner (falls es in seiner Macht steht) immer den für A schlechtesten Zug macht, da dieser automatisch der beste für B ist.
Bei einem Nicht-Nullsummen-Spiel wie dem iterierten Gefangenendilemma unterscheidet sich nun der relative Erfolg gegenüber dem Mitspieler vom absoluten Erfolg des Einzelnen:
Angenommen, Spieler A spielt **Tit-for-Tat mit Defektion im ersten Spielzug**. Falls Spieler B dagegen mit **permanenter Defektion** antritt, und das Spiel $n + 1$ Spielzüge dauert, so erhalten beide Spieler $(P + n \cdot P)$ Punkte. Relativ gesehen ist dies ein Unentschieden. Verfolgt jedoch Spieler B die Strategie **permanenter Kooperation**, so erhält Spieler A $(T + n \cdot R)$ und Spieler B $(S + n \cdot R)$ Punkte. Im zweiten Fall hat Spieler B also relativ zu A schlecht abgeschnitten. Beobachtet man jedoch nur den Gesamterfolg von Spieler B, so ist die zweite Variante vorzuziehen, falls n groß genug ist.
Dieses Beispiel macht deutlich, dass eine erfolgreiche Strategie nicht versuchen muss, bei jeder Konfrontation mehr Punkte als der Gegner zu erzielen.

2.4 Nichtexistenz einer optimalen Strategie unabhängig von der Umgebung

Man kann leicht verdeutlichen, dass “die” optimale Strategie für das iterierte Gefangenendilemma nicht existieren kann, falls die Zukunft stark genug bewertet wird (d.h. dass w groß genug ist): Spielt der andere Spieler permanente Defektion, so ist es am besten, selbst permanente Defektion zu spielen. Spielt der andere Spieler jedoch permanente Vergeltung, so erhält man, wenn man permanente Defektion spielt

$$E = T + \sum_{i=1}^{\infty} w^i P = T + wP \sum_{i=0}^{\infty} w^i = T + wP \cdot \frac{1}{1-w}$$

Falls man selbst permanente Kooperation spielt, so ergibt sich

$$E = R \cdot \frac{1}{1-w}$$

Das heißt, permanente Kooperation zahlt sich aus, wenn gilt:

$$R \cdot \frac{1}{1-w} > T + wP \cdot \frac{1}{1-w}$$

Aufgelöst nach w:

$$w > \frac{T-R}{T-P}$$

3 Die Suche nach einer guten Strategie in einer statischen Umgebung

Auch wenn es keine optimale Strategie geben kann, will man doch eine möglichst gute Strategie finden. In 2.3 hat sich schon gezeigt, dass es nicht einfach ist, ein Bewertungskriterium für eine Strategie festzulegen. Ein möglicher Ansatz ist es, aus einer Vielzahl von Strategien eine Gruppe zu bilden und diese Strategien jeweils gegeneinander antreten zu lassen. Wenn sich die Zusammensetzung der Gruppe während des Experiments nicht ändert, spielt sich die Situation in einer statischen Umgebung ab. In einer solchen Umgebung hat auch Axelrod 1980 die ersten empirischen Ergebnisse gewonnen.

3.1 Empirische Untersuchung mit Hilfe von Computerturnieren

Dazu organisierte er ein Computerturnier, bei dem führende Wissenschaftler aus den Bereichen Psychologie, Ökonomie, Politische Wissenschaft, Mathematik und Soziologie dazu aufgerufen wurden, ein Programm einzureichen, das die von ihnen favorisierte Strategie realisiert. Axelrod ließ die eingereichten Algorithmen gegeneinander antreten, wobei jeder Teilnehmer auch gegen sich selbst und gegen Random(0.5) spielen musste.
Das Ergebnis war: Tit-for-Tat (mit Kooperation im ersten Zug), das einfachste aller eingereichten Programme, gewann das Turnier. Auch in einem zweiten, von Axelrod durchgeführten Turnier gewann wieder Tit-for-Tat. Bei diesem Turnier erhielten alle Teilnehmer eine genaue Beschreibung der Ergebnisse des ersten Turniers, doch Tit-for-Tag wurde abermals nur einmal eingereicht. Viele anderen Teilnehmer versuchten den Tit-for-Tat Algorithmus geringfügig zu verbessern, jedoch konnte keine der Modifikationen ein besseres Ergebnis erzielen als das Original. Es stellt sich nun die Frage, welche Eigenschaften Tit-for-Tat besitzt, die diesen Erfolg ermöglicht haben.

3.2 Folgerungen für eine “möglichst gute” Strategie

Die nachfolgenden Forderungen an einen guten Algorithmus gelten nur, wenn der Diskontparameter w genügend groß ist. Ansonsten gibt es nur eine Forderung: Versuche so schnell wie möglich so viel wie möglich mitzunehmen. Dies führt zu ständiger Defektion. Ist w jedoch hinreichend groß, kann man einige grundlegene Verhaltensforderungen an einen guten Algorithmus angeben.

- **Sie nicht neidisch!**
  Neidische Regeln sind solche, die es nur darauf anlegen, in jedem Spiel besser davonzukommen als der Gegner. Solche Regeln erreichen **relativ** zu anderen Regeln oft bessere Punktzahlen, doch da das Gefangenendilemma kein Nullsummenspiel ist, maximieren sie dadurch nicht zwangsläufig ihre **absoluten** Punktzahlen.

  Es kann theoretisch sogar vorkommen, dass eine Strategie in **jedem einzelnen** Spiel besser ist als ihr Gegner, jedoch insgesamt den letzten Platz belegt. Dies soll nun anhand eines Beispiels verdeutlicht werden.

  Die Auszahlungsmatrix sei wie bei dem von Axelrod durchgeführten Computer - Turnier mit

$$T = 5, R = 3, P = 1, S = 0$$

bestimmt. Nun kann man die Ergebnisse des Spiels Tit-For-Tat gegen permanente Defektion bei einer Spiellänge von $n = 200$ Zügen wie folgt in einer Matrix darstellen:

	permanente Defektion	Tit-for-Tat
permanente Defektion	(200; 200)	(204; 199)
Tit-for-Tat	(199; 204)	(600; 600)

Spielt Tit-for-Tat gegen sich selbst, so erhält jeder n mal R, also 600 Punkte. Spielt permanente Defektion gegen sich selbst, ist das Ergebnis für beide $n \cdot P = 200$ Punkte. Spielt Tit-for-Tat gegen permanente Defektion, so erhält der Defektion - Spieler im ersten Zug T, während der Tit-for-Tat Spieler im ersten Zug gar nichts bekommt ($S = 0$). Danach verhält es sich 199 Züge lang wie beim Spiel Defektion vs. Defektion. Also erhält Tit-for-Tat in diesem Match 199 Punkte, permanente Defektion 5 Punkte mehr. Das heißt permanente Defektion ist in jedem Spiel besser oder zumindest gleich gut wie der Gegner. Wenn nun ein Tit-for-Tat Spieler gegen einen Defektion - Spieler antritt, so gewinnt der Defektion - Spieler. Wenn aber zwei Defektions - Regeln (PD1, PD2) und zwei Tit-for-Tat Regeln (TT1, TT2) in einer Gruppe zusammengefasst werden, wobei jeder gegen jeden gespielt wird, dann zeigt sich folgendes Ergebnis:

Ergebnis von TT1	Ergebnis von PD1
vs. TT2: 600	vs. TT1: 204
vs. PD1: 199	vs. TT2: 204
vs. PD2: 199	vs. PD2: 200
= 998	= 808

Hier zeigt sich deutlich der Vorsprung von Tit-for-Tat, obwohl es in **keinem einzigen Spiel** besser abschneidet als der Gegner. Tit-for-Tat ist sogar komplett neidlos und wird niemals mehr Punkte erreichen als der Gegner. Doch wie man an diesem Beispiel nochmal schön erkennt, kann ein relatives Maß keinen Aufschluss über die Güte eines Algorithmus liefern. Für viele Menschen ist es aber nicht leicht, sich vom Denken in Nullsummenspielen zu lösen, wie das Computerturnier gezeigt hat. Axelrod bestätigt diese Ansicht auch durch Experimente mit Studenten.

- **Freundlichkeit**
  In beiden Turnieren zeichnete sich ein klares Ergebnis ab. Freundliche Algorithmen erreichten wesentlich bessere Platzierungen als provokative. Der Plan der Teilnehmer, die Regeln einreichten, die sich nicht freundlich verhalten, ist offensichtlich. Durch geschicktes Einstreuen von Defektionen soll oftmals T erreicht werden. Dadurch kann man Algorithmen, die zu nachsichtig sind, leicht ausbeuten. Das Problem liegt hier darin, dass die meisten Regeln schlau genug sind, sich nicht ausbeuten zu lassen. Regeln die nicht freundlich sind, wollen sowieso den anderen ausbeuten, werden sich also selbst nur schwer als Opfer eignen. Dadurch erhält ein provokativer Algorithmus gegen eben solche Regeln oft nur P, da sich die Regeln gegenseitig provozieren und bestrafen. Aber auch freundliche Regeln sind zumeist nicht “dumm”, das heißt, sie sind zwar nicht provokativ, aber doch provozierbar. Und auch gegen solche Regeln wird ein provokativer Algorithmus selten mehr als $n \cdot R$ erhalten, da er zwar ab und zu T einstreichen kann, jedoch meistens sofort wieder bestraft wird, was oft in einer Serie von Bestrafungen endet, wobei beide Spieler nur P bekommen.

- **Erwidere Defektion und Kooperation**
  Mit anderen Worten: Sei nachsichtig, aber bestrafe den Gegner für jede Defektion. Wichtig ist es hierbei die Waage zwischen den zwei Extremen zu halten. Denn Regeln, die sich zu nachsichtig verhalten, werden oft ausgenützt, wogegen Regeln, die zu wenig Nachsicht zeigen, oft in einem Teufelskreis von gegenseitigen Bestrafungen enden. Auch hier zeigt Tit-for-Tat wieder eine denkbar einfache Lösung: Die absolute Gerechtigkeit. Dieser Ansatz spiegelt sich schon seit jeher in Sprichwörtern wie “Wie du mir, so ich dir!” und “Auge um Auge, Zahn um Zahn” wider. Und auch wenn solche Verhaltensweisheiten aus unserer sozialpolitischen Sicht vielleicht ein wenig nach einem Relikt aus dem Mittelalter klingen und bestimmt aus ethnischer Sicht von vielen Seiten attackierbar sind, so zeigt sich hier doch eine gewisse Eleganz in der Simplizität, die dem Algorithmus auf den ersten Blick anhaftet.

  Diese Trivialität von Tit-for-Tat scheint auch den Teilnehmern der Computer - Turniere suspekt gewesen zu sein, denn immerhin dachten auch im zweiten Turnier alle Teilnehmer bis auf einen, es besser machen zu können, also das Prinzip durch Spezialisierung zu schlagen. Und mittlerweile haben Mathieu und Delahaye auch durch Gradual eine Strategie entwickelt, die in einigen Versuchen Tit-for-Tat deutlich geschlagen hat. Diese Strategie widerspricht der Forderung nach Gerechtigkeit, denn sie bestraft den Gegner härter als Tit-for-Tat, und ist damit auch noch erfolgreicher. Im Sinne der sozialen Gerechtigkeit, die sich gerade heutzutage eher durch zuviel Nachsicht auszeichnet, ist diese Regel natürlich nicht optimal. Aber Mathieu und Delahaye wollten wohl auch keine allgemeingültige Aussage mit dieser Regel belegen, sondern eher die Suche nach einer “möglichst” guten Strategie weitertreiben. Und trotz allem ist Gradual ja eine freundliche Regel. Es wird also nur der bestraft, der provoziert. In welchem Maß zurückgeschlagen werden muss, ist natürlich wieder in jeder Umgebung unterschiedlich.

- **Sei nicht zu raffiniert!**
  Bei dem Computer - Turnier gab es häufig Regeln, die zumindest versuchten, Random zu erkennen und entsprechend mit lauter Defektionen zu antworten. Jedoch gab es auch sehr komplexe Algorithmen, die versuchten, auf Grund des vergangenen Spielverlaufs die Wahrscheinlichkeit zu berechnen, mit der der Gegner kooperiert. Solche Regeln waren für ihren Gegner oft nicht mehr von Random zu unterscheiden und beide Parteien erzielten suboptimale Ergebnisse. Das Problem, welches die Programmierer der komplexen Algorithmen übersehen haben, ist, dass alles, was man zieht, wieder auf einen selbst zurückfällt. Das heißt, wenn man das Verhalten des Gegners modellieren will, dann muss man zwangsweise auch bestimmen, wie er sein Verhalten dem Gegner anpasst. Da dies aber offensichtlich zu weit führt, kann es nicht gut sein, eine zu komplexe Strategie zu entwickeln.

  Auf der anderen Seite kennen wir mittlerweile Gradual. Gradual ist eine relativ komplizierte Regel, schon allein wegen ihres großen Gedächtnisaufwandes. Andererseits ist Gradual aber auch eine sehr erfolgreiche Regel. Die Erklärung ist, dass Gradual eine hochangepasste Lösung ist. Wahrscheinlich gibt es noch andere, komplexe Lösungswege, die unter ähnlichen Voraussetzungen besser abschneiden als Tit-for-Tat. Jedoch ist der Vorsprung von Gradual gegenüber Tit-for-Tat nur marginal, während der Entwicklungsaufwand von Gradual gegenüber Tit-for-Tat wohl doch erheblich höher war. Die Raffinesse hat aber noch einen zweiten Nachteil. Tit-for-Tat kann man in vielen verschiedenen Varianten des Gefangenendilemmas ohne weiteres übernehmen. Je spezialisierter eine Strategie ist, desto anfälliger ist sie aber auch gegenüber Änderungen der Spielparameter.

4 Die Suche nach einer guten Strategie in einer evolutionären Umgebung

Nachdem die verschiedenen Strategien in einer statischen Gruppe bewertet wurden, stellt sich nun die Frage, wie es um die "Überlebenschancen" der einzelnen Algorithmen wohl in einer evolutionären Umgebung bestellt wäre. Axelrod motiviert diese Idee mit der Überlegung, dass bei einem weiteren Turnier wohl Tit-for-Tat aufgrund seines enormen Erfolges öfter eingereicht worden wäre als jede andere Strategie. Er geht dieser Frage nach, indem er die vom Turnier schon bekannten Algorithmen wieder gegeneinander antreten lässt, diesmal aber über mehrere Generationen hinweg. Erhält Algorithmus A in Generation n p-mal so viele Punkte wie Algorithmus B, so wird er in Generation $(n + 1)$ p-mal so oft vertreten sein wie sein Konkurrent.
Das Ergebnis eines solchen Simulationslaufes sieht man in Abbildung 1.

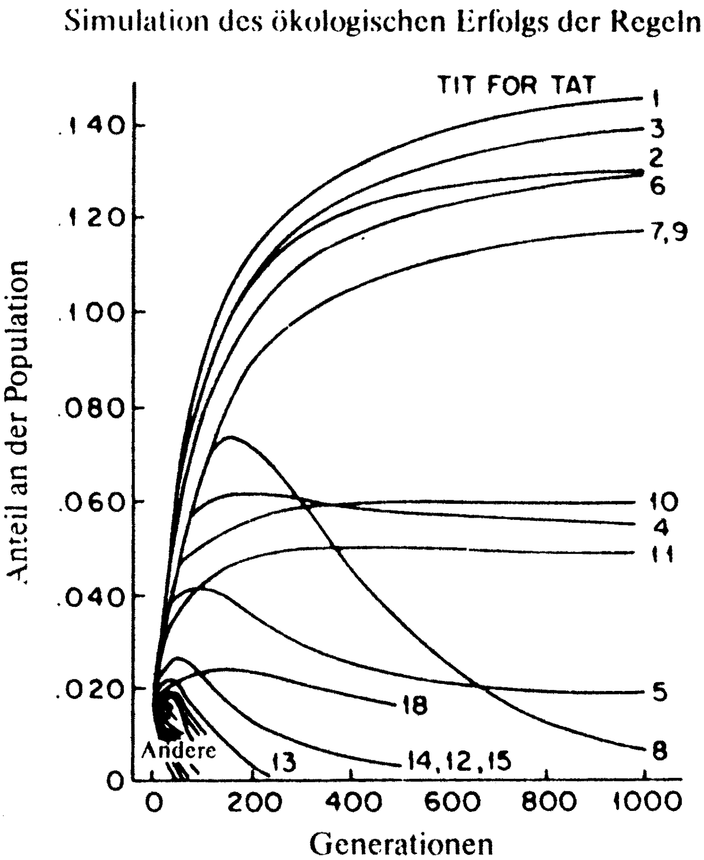

Abbildung 1: Aus [**?**, Seite 45]

Der Verlauf lässt sich im Wesentlichen in drei Teile zerlegen:

1. Aussonderung
   In dieser Phase sind noch alle Strategien vertreten. Es vermehren sich also hauptsächlich Regeln, die mit einer vielzahl verschiedener Gegner gut umgehen können. Zu gutmütige Strategien werden ausgebeutet und verschwinden.

2. Es trennt sich die Spreu vom Weizen
   Nun sind nur noch Regeln vorhanden, die den Aussonderungsprozess überlebt haben. Regeln, die im ersten Teil dadurch dominiert haben, dass sie andere Strategien ausgebeutet haben, finden nun keine Opfer mehr. Sie sterben aus.

3. Die Elite unter sich
   Die kooperativen Algorithmen dominieren. Es stellt sich eine stabile Situation ein.

4.1 Stabilität einer evolutionären Umgebung

Wenn man die Ergebnisse der Simulation im letzten Drittel betrachtet, so stellt man fest, dass sich ein relativ stabiles Gleichgewicht zu bilden scheint. Keine der Strategien verbessert sich noch relativ zu den anderen. Was passiert aber, falls ein andersgearteter Algorithmus versucht in die Gruppe einzudringen bzw. wie definiert man "eindringen" sinnvoll?
Um diese Frage zu beantworten geht man zunächst von einer Population aus, in der alle Individuen die selbe Strategie verfolgen. Man spricht nun vom **Eindringen** einer Strategie, "wenn der Neuling einen höheren Punktwert mit einem Einheimischen erhält als ein Einheimischer mit einem anderen Einheimischen" ([**?**], S. 50). Aus ökonomischer Sicht würde das dazu führen, dass andere Einheimische den Vorteil des Eindringlings bemerken und seine Strategie übernehmen. Andererseits muss auch untersucht werden, ob es Populationen gibt, in die keine Strategie eindringen kann. Die Strategie dieser Population nennt man **kollektiv stabil** ([**?**], S.51 Abs. 1 und S.62 Fußnote 1). Die Stabilität einer Strategie wird nun an verschiedenen Beispielen untersucht.

4.2 Untersuchung einer "gutwilligen" Umgebung

Eine "gutwillige" Umgebung zeichnet sich dadurch aus, dass die vorherrschende Strategie im ersten Zug mit einer positiven Wahrscheinlichkeit kooperiert. Dies ist eine stark abgeschwächte Version der Freundlichkeit und ist wohl auch hauptsächlich für die folgenden theoretischen Erläuterungen interessant, da eine Strategie, die im ersten Zug kooperiert und dann immer defektiert intuitiv wohl kaum als "gutwillig" bezeichnet werden kann.

4.2.1 Tit-for-Tat

Da Tit-for-Tat in den Computer-Turnieren und -Simulationen sehr erfolgreich war, liegt es nahe, diesen Algorithmus auf seine Stabilität zu untersuchen. Ein weiterer Grund Tit-for-Tat als Beispiel auszusuchen ist, dass Tit-for-Tat sehr einfach ist und somit die Beweise leicht verständlich sind.
Es sei nun eine Population von Tit-for-Tat gegeben. Wenn eine andere Strategie in Tit-for-Tat eindringen will, so muss sie gegen Tit-for-Tat mehr Punkte erhalten, als Tit-for-Tat gegen Tit-for-Tat. Da Tit-for-Tat gegen sich selbst stets kooperiert, muss der andere Algorithmus mindestens einmal T gegen Tit-for-Tat erhalten, da T das einzige Ergebnis $> R$ ist. Das heißt

aber, dass die Strategie mindestens einmal defektieren muss. Da Tit-for-Tat diese Defektion im nächsten Zug sofort imitiert, hat der Herausforderer zwei Möglichkeiten:

- **Kooperation**
  Jetzt erhält Tit-for-Tat T, die andere Strategie S. Tit-for-Tat wird im nächsten Zug wieder kooperieren. Falls die Defektion im ersten Zug optimal war, so muss auch hier die Defektion optimal sein, da Tit-for-Tat nur ein Gedächtnis von einem Zug besitzt. Im nächsten Zug muss dann, falls Kooperation im zweiten Zug optimal war, auch wieder kooperiert werden. Es ergibt sich also als optimale Strategie eine ständige Wiederholung von Defektion und Kooperation.

  Bei dieser Strategie erhält der Herausforderer

  $$E_H = T + wS + w^2T + ... = (T + wS) \cdot \sum_{i=0}^{\infty} w^2 i = (T + wS)\frac{1}{1 - w^2}$$

  Tit-for-Tat erhält gegen sich selbst: (siehe 2.4)

  $$E_T = R \cdot \frac{1}{1 - w}$$

  Wenn man nun sicherstellen will, dass $E_T \geq E_H$, so muss für w gelten

  $$w \geq \frac{T - R}{R - S}$$

  Die Strategie DC kann also nicht gegen Tit-for-Tat eindringen, wenn w hinreichend groß ist.

- **Defektion**
  Da nun beide defektieren, erhält jeder P. Tit-for-Tat wird nun im nächsten Zug auch defektieren. Wenn in diesem Zug Defektion optimal war, so muss im nächsten Zug Defektion ebenfalls wieder die bessere Wahl ergeben usw. Es ergibt sich also im weiteren Verlauf als optimale Strategie permanente Defektion.

  Da sich Tit-for-Tat gegen permanente Defektion ebenso verhält wie permanente Vergeltung, können hier die Ergebnisse aus 2.4 übernommen werden. Das heißt, wenn w hinreichend groß ist, kann auch diese Strategie nicht in eine Tit-for-Tat Population eindringen. Genauer gilt, wie wir wissen

  $$w \geq \frac{T - R}{T - P}$$

  Wir benötigen hier nur ein $\geq$, da der gegnerischen Strategie das Eindringen auch schon verwehrt ist, falls sie ebensoviele Punkte erhält wie Tit-for-Tat.

Für jede der beiden Möglichkeiten gibt es eine untere Schranke für w, ab der die implizierte optimale Strategie nicht mehr in die Tit-for-Tat Population eindringen kann. Insgesamt ergibt sich also, dass das Eindringen in Tit-for-Tat unmöglich ist, falls

$$w \geq \max\left(\left\{\frac{T - R}{T - P}, \frac{T - R}{R - S}\right\}\right)$$

Das heißt aber, ab dieser Schranke ist Tit-for-Tat kollektiv stabil. Eine gutartige Umgebung **kann** also kollektiv stabil sein, muss aber nicht. Nun ist noch die Frage zu klären, wie es um die Möglichkeiten zum Eindringen in eine “bösartige” Umgebung bestellt ist.

4.2.2 Andere Strategien

Eine beliebige gutwillige Strategie kann nur dann kollektiv stabil sein, wenn w hinreichend groß ist. Dies ist intuitiv leicht einsichtig, da eine gutwillige Strategie immer im ersten Zug ausgebeutet werden kann. Die Strafe muss dann schwer genug wiegen, um dem Gegner im Durchschnitt weniger als R Punkte zu bescheren. Der formale Beweis ist jedoch etwas komplexer und in [**?**] nachzulesen.

Weiter muss jede **freundliche** Regel, um kollektiv stabil sein zu können, von der ersten Defektion des Gegner provoziert werden. Würde sie das nicht, könnte eine Strategie eindringen, die genau einmal defektiert.

Da eine freundliche Strategie immer gutwillig ist, gilt für jede freundliche Regel, dass sie nur dann kollektiv stabil sein kann, wenn w hinreichend groß ist und sie durch die erste Defektion des Gegners provoziert wird.

4.3 Untersuchung einer "böswilligen" Umgebung

Bei gutwilligen Strategien hängt es immer von den Parametern des Spiels ab, ob sie überhaupt kollektiv stabil sein können. Böswillige Strategien können auch unabhängig von der Parametrisierung kollektiv stabil sein. Permanente Defektion zum Beispiel ist kollektiv stabil. Denn wenn eine andere Strategie gegen permanente Defektion spielt, so muss mindestens einmal kooperieren, sonst wäre sie ja selbst permanente Defektion. Bei dieser Kooperation erhält der Herausforderer aber nur S, ist also schon schlechter dran als permanente Defektion gegen sich selbst.

Dies führt zu dem Ergebnis, dass in einer Welt ohne Kooperation niemals ein einzelnes Individuum anfangen wird zu kooperieren, da es dadurch nur schlechter gestellt würde. Da dadurch der Entstehung von Kooperation praktisch ein Riegel vorgeschoben wird, hat Axelrod die Situation noch weiter untersucht. Denn in eine Population von permanenter Defektion kann nur dann niemand eindringen, wenn die Herausforderer alle einzeln ankommen. Sobald nur eine kleine Gruppe auf einmal ankommt, ergibt sich eine völlig andere Situation. Die Eindringlinge können nämlich jetzt untereinander interagieren, und somit eventuell als Gruppe einen vorteil gegenüber den Einheimischen erlangen.

Als Beispiel sei nun die selbe Auszahlungsmatrix wie in 3.2 gegeben (siehe auch [**?**], S. 57ff). In einer Population von permanenter Defektion werden immer Durchgänge mit je zehn Zügen gespielt. Die Wahrscheinlichkeit, auf den selben Spieler wieder zu treffen sei $w := 0.9$. Jeder Einheimische erhält also $10 \cdot P = 10$ Punkte pro Runde. Nun trifft eine kleine Gruppe von Tit-for-Tat Spielern ein. Ein Tit-for-Tat Spieler erhält gegen permanente Defektion im ersten Zug $S = 0$, danach jeweils $P = 1$, also insgesamt 9 Punkte. Untereinander erhalten Tit-for-Tat Spieler aber 30 Punkte pro Runde. Somit ist Tit-for-Tat gegen permanente Defektion nur leicht schlechter, erhält jedoch mit sich selbst viel mehr Punkte als permanente Defektion.

Weiter sei nun angenommen, die Tit-for-Tat Spieler sind eine so kleine Minderheit, dass für den Durchschnitt der Defektions Spieler die Konfrontation mit Tit-for-Tat Spielern vernachlässigbar ist. Das heißt, dass permanente Defektion im Wesentlichen weiterhin 10 Punkte bringt. Nun ist es offensichtlich, dass Tit-for-Tat Spieler diesen 10 Punkte Durchschnitt überbieten können, wenn sie oft genug mit anderen Tit-for-Tat Spielern interagieren. Wenn man nun einen einzelnen Tit-for-Tat Spieler betrachtet, dann sei p der Anteil seiner Spiele gegen andere Tit-for-Tat Spieler und entsprechend $1 - p$ der Anteil der Interaktion mit permanenter Defektion.

Die durchschnittliche Punktzahl des Tit-for-Tat Spielers beträgt dann

$$30p + 9(1 - p)$$

Der Tit-for-Tat Spieler ist also im Mittel besser als permanente Defektion, wenn gilt:

$$30p + 9(1 - p) > 10 \Leftrightarrow$$

$$p > \frac{1}{21}$$

Dies gilt, wie bereits erwähnt, nur unter Vernachlässigung des Punktanstiegs des durchschnittlichen Defektion Spielers.
Es ergibt sich also, dass eine Invasion von Tit-for-Tat Spielern in eine Gruppe von Defektion Spielern möglich ist, sobald die Tit-for-Tat Spieler nur in 5% ihrer Spiele untereinander interagieren.

4.4 Interpretation realer Situationen

Mit Hilfe dieser Erkenntnisse versucht Axelrod in [**?**, Kapitel 4] das Auftreten "stiller Abkommen" zwischen den verfeindeten Soldaten im Stellungskrieg des ersten Weltkrieges zu erklären. Dazu zitiert er verschiedene historische Quellen, die alle eine gemeinsames Phänomen beschreiben. Denn während der lange andauernden Gegenüberstellung der feindlichen Soldaten bildete sich oftmals eine Waffenruhe nach dem Motto "leben und leben lassen". Dies erklärt Axelrod anhand einer anfangs defektionären Umgebung, in der zufällig durch äußere Umstände, wie zum Beispiel schlechtes Wetter oder Essensausgabe, kurzfristige Waffenruhen entstanden. Die Soldaten wussten, dass sie den Gegner empfindlich treffen könnten, falls sie während der Essenszeit den Gegner bombardierten, wussten aber ebenso, dass der Gegner am nächsten Tag erzürnt zurückschlagen würde. Durch diese etwas an Gradual erinnernde Strategie wurde über lange Zeit Kooperation aufrecht erhalten. Um die Vorgesetzten zu täuschen, die ja nur den Angriff tolerierten, schossen die Soldaten absichtlich daneben. Diese Praktiken wurden schließlich durch die Obrigkeit unterbunden, indem diese in kurzen Abständen Stoßtrupps hinter die feindlichen Linien schickten. Als Beweis für die Ausführung des Befehls musste der Trupp Gefangene mitbringen, oder Verluste unter den eigenen Soldaten vorweisen können. Somit war es den Frontsoldaten unmöglich, weiterhin stille Übereinkommen mit dem Feind zu treffen und ihnen war die Entscheidungsfreiheit als Grundlage der Kooperation entzogen. Axelrod führt auch eine Vielzahl von Beispielen aus biologischer und ökonomischer Sicht an, deren Aufarbeitung den Rahmen dieser Arbeit sprengen würde, die aber sehr nett zu lesen sind.

5 Anwendbarkeit der Modellierung

Die Modellierung im Gefangenendilemma ist mathematisch relativ einfach. Das Problem ist also im Vergleich zu der “realen Welt” sehr abstrakt beschrieben. Es muss also untersucht werden, inwieweit die aus mathematischer Sicht notwendigen Vereinfachungen des Modells noch der Wirklichkeit gerecht werden und ob nur vorhandene Einsichten erläutert werden, oder auch neue Erkenntnisse gewonnen werden können.

Zunächst zu den Vereinfachungen des Modells. Auf den ersten Blick ist die größte Vereinfachung wohl, dass der Spieler nur die Wahl zwischen Kooperation und Defektion hat. Allerdings lassen sich erstaunlicherweise viele Fälle im alltäglichen Leben eines Menschen finden, wo er tatsächlich zur Interaktion mit einem anderen Individuum gezwungen ist. Gerade im Sinne der Wirtschaftlichkeit ist zum Beispiel ein Unternehmen oft an *einen* Zulieferer oder Kunden gebunden. Auch kann man den Wert einer rein quantitativen Output-Funktion kaum mehr in Frage stellen, nachdem die Ökonomie seit geraumer Zeit in der Nutzen-Funktion den selben Mechanismus verwendet.

Am interessantesten sind jedoch die theoretischen Erkenntnisse über die Förderung der Kooperation durch Änderung der Parameter des Spiels (siehe auch [**?**, Kapitel 7]). Denn niemand wird wohl anzweifeln, dass eine Erhöhung der Bestrafung für Defektion die Bereitschaft zur Kooperation erhöht. Ebenso ist es gerade in einer Gesellschaft, in der die Moral unter anderem auf Grund wachsender Anonymität verfällt, einsichtig, dass wiederholte Interaktion mit dem selben Individuum die grundlegene Voraussetzung für Kooperation ist. Somit kann man sagen, dass die Modellierung doch trefflich gelungen ist. Ob jedoch die Aussagen der Theorie wirklich neue Erkenntnisse vermitteln, ist die andere Frage. Denn jeder Mensch erkennt intuitiv sofort die eben beschriebenen Zusammenhänge. Doch auch wenn die Modellierung der Kooperation mit Hilfe des Gefangenendilemmas nicht zur Revolution des menschlichen Selbstverständnisses geführt hat, so ist sie vielleicht ein erster Schritt auf dem Weg.

Literatur

[1] Robert Axelrod. *Die Evolution der Kooperation*. Oldenbourg, 2000.

[2] Tobias Thelen. Spieltheorie und das gefangenendilemma. http://www.cl-ki.uni-osnabrueck.de/˜nntthele/ipd/index.html.