Spieltheorie (Game Theory)

Spieltheorie umfasst ein weit größeres Gebiet als nur die Behandlung des Gefangenendilemmas. Zwei Namen sind untrennbar mit ihrer Entstehung verbunden: John von Neumann und Oskar Morgenstern mit in ihrem 1944 erschienenen Buch *Theory of Games and Economic Behavior*. Darin wurde zum ersten Mal ein Modell einer Entscheidungstheorie (rational choice theory) auf mathematischer Basis entworfen. Die Spieltheorie gilt heute sowohl als Teilgebiet der Mathematik wie der Wirtschaftswissenschaften.
Die Entscheidungen der „Spieler“ wurden dabei als eine rationale Wahl von zu erwartenden Ergebnissen (outcomes) verschiedener Handlungsalternativen (sog. lotteries) aufgefasst. Interessant waren die Spiele, in denen mehrere (meist zwei) Spieler verschiedenen Strategien ihrer Nutzenmaximierung verfolgen. Von Mathematikern formuliert und als Disziplin der Wirtschaftswissenschaften zuhause (Selten, Maynard, Harsanyi, Nash), wurde erst mit der Diskussion des Gefangenendilemmas die sehr technische Diskussion der Spieltheorie auch für Soziologen, Politologen und Philosophen interessant.

Um nur einen kurzen Einblick in die Rational Choice Theory (RCT) zu geben, die als Grundlage der Spieltheorie gilt, sollen folgende Bemerkungen genügen.
Ausgang bildet der das Modell des *homo oeconomicus* und das **Nutzenmaximierungstheorem**.
Demnach wird angenommen, dass sich jeder Mensch grundsätzlich (wenn auch bei steigender Gütermenge mit abnehmenden Grenznutzen) nutzenmaximierend verhält, d.h. jeder Mensch will von einem knappen und begehrten Gut grundsätzlich mehr als weniger erlangen.

<u>Das Konzept des Erwartungsnutzens und die Nutzenfunktion</u>
Die RCT fragt nach der richtigen Wahl zw. zwei oder mehreren Handlungsalternativen in einer gegeben Situation
Gegeben seien zwei Handlungsalternativen A und B in Situation s.
Jede Handlung soll nun drei Folgen besitzen: ua1, ua2 und ua3 bzw. ub1, ub2, ub3
Die jeweiligen Folgen besitzen verschiedene Wahrscheinlichkeiten: pa1, pa2 und pa3 bzw. pb1, pb2 und pb3
Der Gesamtnutzen GN(i) jeder Handlung ergibt sich dann aus der Summe der Produkte von Nutzen der Folgen mit ihren Wahrscheinlichkeiten
GN(A) = ua1*pa1 + ua2*pa2 + ua3*pa3
GN(B) = ub1*pb1 + ub2*pb2 + ub3*pb3
Falls GN(A) > GN(B), ist es richtig Handlung A wählen.
Falls GN(B) > GN(A), ist es richtig Handlung B wählen.
Falls GN(A) = GN(B), ist es egal, welche Handlung A ich wähle. Beide Handlungen sind moralisch richtig.
Neumann/Morgenstern haben dies mit einer Wahl über Lotterien verglichen. Die Summe aller Handlungsfolgen müssen dabei die Gesamtwahrscheinlichkeit von 1 ergeben:
Bsp.: Sei Handlung A eine Lotterie, die mir folgende erwartbaren Gewinne offeriert:
1000 € mit der Wahrscheinlichkeit von ¼
100 € mit der Wahrscheinlichkeit von ¼
10 € mit der Wahrscheinlichkeit von ½
dann beträgt der Erwartungsnutzen dieser Handlung A
GN(A) = 1000*1/4 + 100*1/4 + 10*1/2 = 280 €
Sei nun die Handlung B mit folgenden erwartbaren Gewinnen:

100.000 € mit der Wahrscheinlichkeit von ¼
100 € mit der Wahrscheinlichkeit von ¼
-1000 € mit der Wahrscheinlichkeit von ½ (ich muss etwas zahlen)
dann beträgt der Erwartungsnutzen dieser Handlung A
GN(A) = 100.000*1/4 + 100*1/4 - 1000*1/2 = 2025 €
Da GN(B) > GN(A), ist nach dem Nutzenmaximierungsprinzip Handlung B (Lotterie B) zu wählen.
Einwände und Kritik:
Man sieht schnell, dass im Modell des Erwartungsnutzens die bloßen Zahlen an ihre Grenzen stoßen. Eine riesiger Nutzen mit einer geringen Wahrscheinlichkeit kann geringere Verluste mit höheren Wahrscheinlichkeiten mehr als aufwiegen. Man kann darauf unakzeptable Extremfälle konstruieren, etwa: 10.000.000 mit 0,1% und –1.000 € mit 99,9% (was einen annähernden Erwartungsnutzen von etwas mehr als 9.000 € ergibt). Eine solche Lotterie würde jedoch keiner der Lotterie A vorziehen, in der nur Gewinne winken.
Die RCT hat deswegen Risikostrategien und Risikoregeln entwickelt, wonach etwa in Fällen extremer Unsicherheit (0,1%) die Handlung vorzuziehen ist, in denen die Minima maximal werden (Maximin-Regel).

Die Spieltheorie nimmt dieses Konzept nun zur Grundlage und entwickelt Situationen, in denen die Nutzenmaximierungen zweier oder mehrerer Agenten (von engl. agents) in Konflikt geraten können. Dazu hat die Spieltheorie verschiedene Spiele entwickelt, in denen es meist darum geht, ein Gleichgewicht herzustellen, das heißt Strategien (was eine Strategie ist, dazu unten) zu finden, die dazu führen, dass alle Seiten soweit wie möglich zufrieden sind. Im Nash-Gleichgewicht etwa kann kein Spieler eine Verbesserung seiner Auszahlung (Nutzens) mehr durch einen Spielzug erreichen. Alle Strategien sind damit auch noch nicht einmal pareto-optimal. Keiner kann sich verbessern, selbst wenn er einem anderen schadet.

Die Mutter aller Spiele der Spieltheorie ist das Gefangenendilemma, das in unzähligen Variationen und Versionen vorliegt. Das Gefangenendilemma (PD = Prisoner’s dilemma)soll zeigen, dass rationales nutzenmaximierendes Verhalten sich gegen sich selbst kehren kann. Andere interpretieren das PD als Konflikt zwischen egoistischer Rationalität und altruistischer Moral.

Das PD

Ausgangslage:
Die klassische Ökonomie (Smith, Ricardo, Bentham, Mill) nimmt an, dass die Verfolgung des Eigeninteresses automatisch – durch eine unsichtbare Hand – zur Steigerung des Allgemeinwohls führt. Ein Konflikt zwischen Eigeninteresse und Gemeinwohl besteht demnach nicht. Diese Kongruenzannahme hat sich jedoch in der Praxis als sehr brüchig erwiesen. Das PD stellt nun eine Situation vor, in der dies in Form eines Dilemmas klar wird.

Man stelle sich zwei Gefangene Kain und Abel vor, die gerade festgenommen und in getrennte Zellen verbracht wurden. Man verdächtigt sie mehrerer Räubereien, kann ihnen jedoch lediglich Diebstahl oder geringere Delikte nachweisen.
Nun versucht die Polizei mit folgender Taktik, die beiden Gefangenen gegeneinander auszuspielen. Sie bietet jedem der beiden getrennt einen Deal an: „Wenn du gegen deinen Kompagnon aussagst und die Räubereien beweist, dann bleibst du straffrei; dein Kumpane jedoch wird 11 Jahre Gefängnis bekommen, weil er alle Verbrechen angelastet bekommt.

Wenn du jedoch weiterhin schweigst, und dein Kumpel ebenso, dann werdet ihr zwei Jahre hinter Gitter wandern. Wenn ihr euch beide gegenseitig verpfeift, bekommt ihr 7 Jahre, weil die Gesamtschuld auf beide Schultern verteilt wird. Ich werde diesen Deal auch deinem Kumpanen anbieten.“
Kain überlegt nun:
„Wenn ich Abel verpfeife, dann wandere ich hier als freier Mann heraus, aber eben nur dann, wenn Abel dicht hält. Falls der mich auch verpfeift, bekommen wir beide 7 Jahre. Wenn ich dicht halte und er auch, sind wir in zwei Jahren draußen.“
Dann schießt es ihm jedoch durch den Kopf: „Und wenn Abel jetzt das Gleiche wie ich denkt? Er wird auch zuerst denken: Super, ich verpfeife Kain, um frei herauszukommen.“
Das Dumme an der Situation ist die Tatsache, dass beide für sich genommen, dann nutzenmaximierend verhalten, wenn sie den anderen verpfeifen, genau dies aber zum Resultat führt, dass sie beide zusammengenommen am schlechtesten wegkommen. Zudem müssen sie aus Nutzenmaximierungsgründen den anderen fast verpfeifen, weil für den Fall, dass sie das nicht tun, sie riskieren, 11 Jahre im Knast zu verbringen.
Folgende Übersicht gibt die Alternativen für beide wieder

	Abel			
Kain		Dicht halten Kooperation (K)	Verpfeifen Defektion (D)	
	Dicht halten Kooperation (K)	2 /2 K/K ok/ok	11/0 K/D schlechtest/best	
	Verpfeifen Defektion (D)	11/0 D/K best /schlechtest	7/7 D/D schlecht/schlecht	

In der Spieltheorie bezeichnet das Dichthalten nun als Kooperation (C oder K) und das Verpfeifen als Defektion (D).
Schaut man sich nun die Tabelle genauer an, so sieht man, dass es ein Dilemma für jeden Gefangenen gibt: Handelt er nach Nutzenmaximierungsgesichtspunkte, sollte er defektieren, wählen jedoch beide diese Nutzenmaximierung, kommt das Ergebnis mit dem schlechtesten Gesamtwert (in Gefängnisjahren) heraus.
Man kann also sagen: Wenn sich beide nutzenmaximierend verhalten, kommt das insgesamt schlechteste Ergebnis, nämlich 14 Jahre in summa für D/D, heraus. Würden sich beide nicht nutzenmaximierend verhalten, käme das beste Gesamtergebnis in summa heraus, nämlich 4 Jahre für K/K. Das hieße aber, dass die Nutzenmaximierung zur Nutzenminimierung führt.

Eine weitere Interpretation des PD ist der Umstand, dass sich Kooperation nicht lohnt, wenn jeder für sich den eigenen Nutzen verfolgt. Er muss nämlich immer damit rechnen, der Gelackmeierte (sucker) am Ende zu sein.

Man kann das PD nun auf beliebige andere Felder anwenden, etwa bei Verträge über zu bezahlende Warenlieferungen. Person A und B haben einen vertrag geschlossen, wonach A 100 Kissen a 1 € an B liefert. Dieser hat versprochen, die Kissen zu bezahlen. Termine wurden nicht abgemacht. Nun ist es aus Sicht von A rational (im Sinne der Nutzenmaximierung), auf die Zahlung von B zu warten und die Kissen nicht zu liefern. Umgekehrt denkt B auch: Lass A erst die Kissen liefern, dann zahle ich nicht. In so einem

Fall kommt eine ähnliche Auszahlungsmatrix zustande, wenngleich hier ein weiterer Nebeneffekt auffällt. In Nutzenmaximierungsumgebungen, in denen abstrakte Gegenseitigkeit (Geld gegen Lieferung, Lieferung gegen Geld) gefragt ist, kommt es zu völliger Passivität (Attentismus = Abwarten).
Mittlerweile gibt Hunderte von Abarten des PD: so findet man solche mit einem Belohnungs- und Bestrafungssystem für K und D, solche die externe Kosten internalisieren (Umweltschutzspiele) usw.

Eine wesentliche Änderung ergibt sich jedoch durch sog. Mehrfach- oder Superspiele. Menschen begegnen sich ja selten nur einmal im Leben und sie sind ja auch nicht nach einer Interaktion „aus der Welt". So wird Kain, sollte K gewählt haben und von seinem Kumpel verpfiffen worden sein, sich diesen nach 11 Jahren „vorknöpfen" und ihn „zusätzlich bestrafen". Person A, der Lieferant, wird Person B beim nächsten Mal genauso betrügen wie der ihn.
Diesem Umstand wiederholter Zusammentreffen wird die Spieltheorie gerecht, indem sie die Spieler mehrere PDs nacheinander spielen lässt. Solche *Superspiele* verlangen eine anderes Spielverhalten, denn jetzt könnte sich Kain für die Defektion Abels nachher rächen. Während man die K und D Strategien nennt, werden die Strategien in Superspielen manchmal auch *Superstrategien* genannt, meist aber nur Strategien.
Eine Superstrategie legt fest, ob und wie man auf die das Wahlverhalten seiner Mitspieler reagiert.
Man unterscheidet demnach:

- nicht-reaktive und
- reaktive Strategien
- konstante
- variable

Andere Einteilungen unterscheiden zwischen

- reinen und
- gemischten

STRATEGIEN. Reine Strategien liegen dann vor, wenn sich ein Spieler auf eine Spielweise festlegt, also nach einem festen Spielplan spielt (etwa indem immer betrügt oder immer kooperiert). Bei gemischten Strategien „mischt" ein Spieler verschiedenen Spielvarianten nach einem Wahrscheinlichkeitskalkül. So beginnt ein gemischter Spieler mit Dauerdefektion, wechselt dann „urplötzlich" zu Dauerkooperation. Gemischte Strategien sind – wie unstete Zeitgenossen – schwer zu durchschauen. Alle im Folgenden genannten Strategien sind reine Strategien. Beispiele für Spiele, in denen typischerweise gemischte Strategien vorkommen, sind das Schere-Papier-Stein(-Brunnen)-Spiel (im Bayer. Schnick-Schnack-Schnuck). Hier erfolgt die Wahl einer Option in der Regel vollkommen zufällig, da eine reine Strategie aufgrund der Nachteile gar nicht möglich ist. Eine IMMER-Schere Strategie kann im ersten Zug zu einem Punkt führen, dann wird sie jedoch durch den Stein geschlagen.

Die einfachsten nicht-reaktiven und konstanten STRATEGIEN sind etwa:
Ich betrüge immer (IMMER-D) oder Ich kooperiere immer (IMMER-K)
Immer-Strategien sind nicht sehr intelligent. Stellt man sich nur vor, IMMER-D trifft auf IMMER-K, dann wird klar, wie vernichtet IMMER-K enden wird. Wandeln wir unsere Gefängnisjahre in Geldbeträge um, so müssen wir, je höher die Jahre sind, einen niedrigeren Betrag ansetzen, da schließlich mehr Jahre hinter Gittern schlechter als weniger Jahre sind. Am einfachsten, dreht man dazu die Seiten um: Für D/D gibt es 2/2, für K/D bzw. D/Kgibt es 0/11 bzw. 11/0 und für K/K 7/7. So ergibt für die beiden nicht-reaktiven, konstanten Strategien:

Spiel	IMMER-D	IMMER-K
1	11	0
2	11	0
3	11	0
Gesamt	33	0

Andere STRATEGIEN etwa sind variabler, entweder nach einem festen Algorithmus oder nach einem Zufallsmechanismus: Die sind etwa KDK oder noch kompliziertere STRATEGIEN.

Nun hat sich bei Turnieren gezeigt, in denen alle mögliche Strategien gegeneinander angetreten sind, dass reaktive und variable Strategien allen anderen überlegen waren. Eine reaktive STRATEGIEN ist eine, die auf den Spielzug des Gegner im nächsten Zug oder irgendwann später reagiert. So gibt es STRATEGIEN (wie auch Menschen), die nach einem D des Gegners bis zum Ende ebenfalls defektieren. Das sind die IMMER-Rächer. Umgekehrt gibt es auch sehr vertrauensselige STRATEGIEN, die nach dem ersten K des Gegners immer kooperieren IMMER-Vertrauen.
Als optimale Strategien hat sich jedoch die Titfortat-STRATEGIEN erwiesen. Sie ist ebenso simpel wie effektiv:

- Sie beginnt stets mit Kooperation
- Wird sie einmal betrogen, betrügt sie im nächsten Zug zurück (titforD-tat)
- Kooperation beantwortet sie ebenfalls mit Kooperation (titforK-tat)

Man kann sich vorstellen, dass TfT gegen sehr böswillige und hinterlistige STRATEGIEN nicht unbedingt gewinnen wird (also den höchsten Gesamtpunktestand hat). So gewinnt LETZT-D (Defektion im letzten Spiel) auf jeden Fall gegen TitforTat. TitforTat (wie du mir so ich dir)mistz keine aggressive Gewinnerstrategie. Ihr Vorteil besteht darin, dass sie kleinere Punktverluste bei aggressiven STRATEGIEN in Kauf nicht, dafür jedoch bei den meisten anderen STRATEGIEN eifrig Punkte sammelt. TitforTat ist

- freundlich: Defektiere/betrüge nie als Erstes
- friedlich, aber bestimmt: Zeige dem anderen seine Grenzen, aber räche dich nicht übermäßig. Versuche auch nur deine Verluste auszugleichen.
- Langfristig: je länger da Spiel dauert, desto besser wird es für TitforTat, da einmalige Verluste aufgewogen werden können.

Mittlerweile gibt es noch andere Spiele als das PD. Im Grunde sind sie strukturell damit vergleichbar, auch wenn sie anderen Formalisierungen erfordern, die uns hier nicht interessieren müssen, und auch andere Konsequenzen nach sich ziehen.
So nimmt das Betrand-Preisspiel ein Duopol (Angebotsmarkt mit zwei Anbietern) an, die um die Nachfrage durch einen niedrigen Angebotspreis konkurrieren. Der Clou ist, dass sie diesen Preis gleichzeitig (etwa durch Eingabe in einen Computer) bekannt geben müssen. Derjenige mit dem niedrigsten Preis räumt den Markt ab. Das Dilemma liegt auf der Hand: Beide Anbieter haben aus Gewinnabsicht das Interesse an einem möglichst hohe Preis, gleichzeitig müssen sie aber einen möglichst niedrigen Preis bieten, da sie sonst die Nachfrage verlieren.

Natürlich ist das Spiel sehr abstrakt, da in Wirklichkeit die Nachfrage sich niemals nur nach Preiskriterien richtet und auch die absolute Vergleichbarkeit selten herstellbar ist (das gilt höchstens für Güter des täglichen Bedarfs wie Milch, Fleisch, obwohl selbst bei Brot manche Menschen bereit sind, enorme Summen auszugeben, wenn eine best. Marke damit verbunden ist). Angenommen wird, dass für das Betrand-Spiel ein Nash-Gleichgewicht existiert.

Als Darstellungsformen gibt es neben der von mir gewählten Matrixform, auch die Baumstruktur und die Agentennormalform (wie oben bei Superspiel), die tabellarisch für jeden Spieler die Ergebnisse nach unten abträgt.
Arithmetische Matrixdarstellungen machen Spiele quantifizierbar.

Literatur:
Robert Axelrod: *The Evolution of Co-operation,* Penguin 1994.
David Gauthier: Morals by Agreement
...und als nette und einfache Einführung:
Manfred Eugen, Ruthild Winkler: *Das Spiel.* Piper, München 1987