

Serjoscha Wiemer

## Von der Matrix zum Milieu. Zur Transformation des Entscheidungsbegriffs zwischen *homo oeconomicus* und evolutionärer Auslese (Preprint).

### Spiele als Entscheidungsmedien

Die Folgenden Überlegungen und Darstellungen setzen sich mit der mathematischen Spieltheorie auseinander, um innerhalb der Spieltheorie Veränderungen des Entscheidungsbegriffs aufzuzeigen, wie sie exemplarisch in der Gegenüberstellung von klassischer und evolutionärer Spieltheorie deutlich werden.

Ausgangspunkt ist dabei ein doppeltes Staunen über die Rolle von Spielen und Algorithmen und ihre enge Beziehung zu Entscheidungstheorien. Gefragt wird speziell nach der Bedeutung von Algorithmen für Veränderungen der theoretischen und experimentellen Untersuchung von Entscheidungen in Konfliktsituationen. Es geht also um Konstellationen von Spielen, Algorithmen und Entscheidungen und ihre historische Veränderung innerhalb der Spieltheorie. Im Zentrum steht dabei ein einzelnes Spiel, das so genannte Gefangenendilemma.

Spiele und Algorithmen bilden unterschiedliche temporär stabilisierte epistemische Konstellationen.<sup>1</sup> Wissenschaftsgeschichtlich bemerkenswert ist, dass sich Diskurse von Algorithmen und Spielen nicht nur innerhalb der mathematischen Spieltheorie kreuzen, sondern mit erstaunlichen Konsequenzen etwa auch in der frühen Computerforschung nach dem 2. Weltkrieg sowie im Forschungsfeld der *artificial intelligence*.<sup>2</sup> Der Computerhistoriker Nathan Ensmenger hat exemplarisch in seiner Studie zur Funktion des Minimax-Algorithmus für die Konstitution des *artificial-intelligence*-Diskurses aufgezeigt (ders. 2011, 10), wie Spiele als »ideale experimentelle Technologie« fungieren, um Problemstellungen zu definieren, Ergebnisse zu quantifizieren und Lösungsvorschläge wiederholt durchzuspielen. Diese wissenschaftliche Funktion von Spielen als formalisierbare und algorithmisierbare Versuchsanordnungen ist ein wichtiges verbindendes Element zwischen dem Computerdiskurs und der modernen Spieltheorie.

Spiele sind Medien der Entscheidung. Das gilt insbesondere für formalisierte Regelspiele, wie sie in der mathematischen Spieltheorie behandelt werden. Dass Spiele als Entscheidungsmedien betrachtet werden können, ist bei klassischen Gesellschaftsspielen wie Schach, Dame, Skat oder Mühle leicht einsehbar. Die Spielenden sind innerhalb des Spiels als Entscheidungsträger aufgefordert, Züge zu tätigen, und dabei aus einem definierten Optionsraum zu wählen.

Genau in diesem Sinne werden Spiele von der modernen Spieltheorie untersucht, deren mathematische Grundlagen auf John von Neumanns Arbeiten aus den 1920er Jahren zurückgehen, als von Neumann formalisierte Analysen von Gesellschaftsspielen durchführte (von Neumann 1928). Seine gemeinsam mit dem Ökonomen Oskar Morgenstern verfasste Studie *Theory of Games and Economic Behavior* von 1944 gilt als Standardreferenz<sup>3</sup> der klassischen Spieltheorie (vgl. Rieck

<sup>1</sup> Während die Unterscheidung zwischen Code und Software in der medienwissenschaftlichen Diskussion fest etabliert ist (vgl. Berry 2011, 29–56), steht eine vergleichbare medientheoretische Differenzierung im Hinblick auf Algorithmen und ihre Materialisierungen, respektive Artikulationen als Code, Software und kulturelle wie politische Praktiken noch weitgehend am Anfang. Vgl. für die Diskussion divergierender Konzepte von Algorithmen mit kultur- und sozialwissenschaftlichem Schwerpunkt Gillespie 2016.

<sup>2</sup> So ist es bezeichnend, dass Alan Turing in den 1950ern die Programmierung von Spielen als zentrale Aufgabe der Computerforschung benennt und Claude Shannon Beitrag zur Softwaregeschichte eng mit der Programmierung von Schach verbunden ist: »Unser Problem besteht nun darin, wie die Maschinen für das Spiel zu programmieren sind«, schrieb Alan Turing 1950 in *Computing Machinery and Intelligence* (ebd., 455). Vgl. weiterführend zur Bedeutung von Spielen für den Computerdiskurs der 1950er: Wiemer 2014.

<sup>3</sup> Luce / Raiffa (1957) weisen darauf hin, dass nicht die Erstausgabe von 1944, sondern die überarbeitete Fassung von 1947 als Standardreferenz gelten kann, weil in der Überarbeitung eine Nutzentheorie [utili-

2006, 107).<sup>4</sup> Ein Blick in die zahlreichen Einführungen zur Spieltheorie verdeutlicht, wie zentral der Begriff der Entscheidung innerhalb der Spieltheorie besetzt ist. Dort findet man, dass jeder *Zug* im Spielverlauf formal als eine »Entscheidung« definiert wird, die ein Spieler im Spielverlauf an einem »Entscheidungsknoten« trifft (ebd., 197).

In der klassischen Spieltheorie wird Spiel von Beginn an als ein Konfliktfeld begrenzter Handlungsoptionen und konkurrierender Interessen bestimmt, an dem »mehrere vernunftbegabte Entscheider« beteiligt sind, die ihre je »eigenen Interessen« verfolgen (ebd., 17). Solche strategischen Spiele können als Modelle für das »rationale Verhalten von Menschen« in sozialen Konfliktsituationen interpretiert werden (Rapoport 1976, 128). Rationale Spieler vermögen die Konsequenzen ihrer Entscheidungen abzuschätzen und tragen dabei dem Umstand Rechnung, dass die Ergebnisse nicht nur von ihren eigenen, sondern auch durch die Entscheidungen der anderen bestimmt werden (ebd.). Die Spieltheorie ist in diesem Sinn eine Entscheidungstheorie unter den Bedingungen von Konflikt und kalkulierender Nutzenmaximierung. Sie tritt mit dem Versprechen an, Entscheidungen analysieren, optimieren und effektiv »rationalisieren« zu können. Wissensgeschichtlich kann die Spieltheorie als ein großangelegtes Projekt der *Formalisierung von Entscheidung* begriffen werden.

In einer medien- und wissensgeschichtlichen Perspektive wird anhand der Spieltheorie nicht nur die erstaunliche Relevanz von Entscheidungstheorien im 20. Jahrhundert nachvollziehbar, die ausgehend von Mathematik und Ökonomie in unterschiedlichen Wissensfeldern und Disziplinen wie Konfliktforschung, Psychologie, Soziologie, Biologie, etc. aufgegriffen werden. Darüber hinaus wird, so meine These, anhand der historischen Entwicklung der Spieltheorie zugleich eine markante Verschiebung im Begriff der Entscheidung selbst erkennbar, verbunden mit Veränderungen von Rationalitätsbegriffen sowie den zugehörigen Rollen von Individuen und Algorithmen als wechselnde Träger und Funktionen von Entscheidungen.

Während in der klassischen Spieltheorie jede Entscheidung erst durch den Verweis auf den »Spieler« als kalkulierendes Entscheidungssubjekt als eine rationale Entscheidung qualifiziert ist, wird in der evolutionären Spieltheorie ein ökologisch-evolutionäres Konzept von *Population* verwendet, das grundsätzlich ohne rasonierende Individuen auskommt. An die Stelle der Entscheidung als Handlung bewusster rationaler Individuen tritt die relative Verteilung von Vielheiten, verstanden als Gleichgewicht erfolgreich evolvierender Strategien oder als Verhaltensdispositionen in einem dynamisch veränderlichen Milieu.

Ein Wechsel der Protagonisten auf der geschichtlichen Bühne des Denkens: Es tritt ab die Figur des Individuums als Ausgangspunkt von Entscheidungsfreiheit und Vernunft; in den Vordergrund drängen konkurrierende Konzepte von Vielheiten (Populationen), technobiologische Simulationen und Algorithmen als Repräsentationen post-souveräner Akteure.

Von diesem Prozess wird auch die Funktion des Spiels erfasst. Stellt es in der klassischen Spieltheorie noch das Medium dar, das die Formalisierbarkeit von Entscheidungen und die algorithmische Berechenbarkeit<sup>6</sup> der besten Strategien gewährleistet, wird es in der evolutionären Spieltheorie zum Modell für die Konstruktion simulierter Evolution.

---

ty theory] ergänzend expliziert wird, die in der Originalausgabe fehlt (Luce / Raiffa 1957, 3, Anm.1).

<sup>4</sup> Als »klassische Spieltheorie« wird im Folgenden der Mainstream der mathematischen Spieltheorie ab Mitte der 1940er Jahre bis in die 1970er Jahre bezeichnet. Der Ansatz der Spieltheorie, wie er durch von Neumann & Morgenstern entwickelt wurde, ist für diese Zeit prägend und wird durch zahlreiche Lehrbücher kanonisiert. Im Unterschied dazu verweist der Ansatz der »evolutionären Spieltheorie« auf Ergänzungen und Reformulierungen der Spieltheorie seit den 1970er Jahren, die mit dem wachsenden Einfluss biologischer, evolutionärer Theorien und neuer algorithmischer Methoden in Zusammenhang stehen. Vgl. für die Unterscheidung zwischen klassischer und evolutionärer Spieltheorie weiterführend Müller 1990, 3-8.

<sup>6</sup> Das Konzept der Formalisierung von Entscheidungen in der klassischen Spieltheorie zielt auf die *Berechenbarkeit* von (optimalen) Entscheidungen. Innerhalb der Mathematik gibt es unterschiedliche Ansätze zur Definition von Berechenbarkeit. Nach der Church-Turing-These können die Begriffe berechenbar und algorithmisierbar synonym gebraucht werden, insofern die Turing-Maschine den intuitiven Begriff von Berechenbarkeit ausdrückt. Das heißt, eine mathematische Funktion ist berechenbar, wenn für sie eine Berechnungsanweisung (Algorithmus) formuliert werden kann. Eine Funktion ist genau dann berechenbar, wenn sie Turing-berechenbar ist (vgl. Ziegenbalg 2010, 225f. sowie Artikel: Berechenbarkeit, in: Wikipedia, Die freie Enzyklopädie; letzter Abruf 7. 7. 2015)

## Gefangenendilemma

Das Gefangenendilemma (engl. *Prisoner's Dilemma*) gilt als das bekannteste Spiel der Spieltheorie überhaupt (Holler/Illing 2009, 2; Rieck 2006, 40). Robert Axelrod hat es als *E. Coli* der Sozialwissenschaft bezeichnet, andere haben es als »Urpflanze« (Pias 2000, 217) oder »ewigen Kauknochen« (Mérő 1998, 28) der Spieltheorie beschrieben; tausende von MathematikerInnen, PsychologInnen, PolitologInnen, SoziologInnen, PhilosophInnen und ÖkonomInnen haben sich seit den 1950er Jahren damit befasst und bis heute ist es Gegenstand aktueller Forschungen (Axelrod 1997, xi).<sup>7</sup>

Der grundlegende Konflikt des Gefangenendilemmas<sup>8</sup> wird mit kleinen Variationen immer wieder gleich erzählt. Es handelt sich um eine Art Kriminalgeschichte, deren dramatische Entscheidungssituation das getrennte Verhör zweier Gefangener bildet:

»Zwei Verdächtige werden getrennt vom Distriktstaatsanwalt vernommen. Sie sind des Verbrechens schuldig, dessen sie verdächtigt werden, aber der Staatsanwalt hat nicht genügend Beweise, um den einen oder den anderen zu überführen. Der Staatsanwalt hat jedoch ausreichend Beweise, um beide wegen eines geringeren Vergehens zu verurteilen. Den Verdächtigen A und B stehen die Alternativen offen, das schwere Verbrechen zu gestehen oder nicht zu gestehen. Sie sind getrennt und können sich nicht miteinander absprechen. Die möglichen Ergebnisse sind folgende: Falls beide gestehen, erhalten beide strenge Urteile, die jedoch wegen des Geständnisses etwas ermäßigt werden. Wenn einer aussagt, (d.h. Beweise gegen den anderen liefert), bekommt der andere alle möglichen Strafen, und der Informant geht straflos aus. Wenn keiner aussagt, können sie nicht des schweren Verbrechens überführt werden, werden aber mit Sicherheit verhört und wegen des geringeren Verbrechens verurteilt werden.« (Rapport 1976, 192)

Die mathematische Darstellung erfolgt charakteristisch durch die so genannte Normalform als Spiel-Matrix oder in der Extensivform. Gegenüber der Extensivform, in der die einzelnen Entscheidungsmöglichkeiten aller SpielerInnen vollständig und in chronologischer Abfolge der Züge (Entscheidungen) nachvollziehbar dargestellt werden, gibt die typische Matrix der Normalform eine Überblicksdarstellung, in der alle möglichen Zugkombinationen und die zugehörigen Auszahlungen ablesbar sind.

---

<sup>7</sup> Die erste Formulierung des Gefangenendilemmas stammt von Merrill Flood und Melvin Drescher, die in einem Paper der RAND Corporation von einem Spielexperiment berichten, dass sie mit zwei Versuchspersonen durchgeführt haben (Flood 1952). Philip Mirowski (2002) und andere haben darauf hingewiesen, wie effektiv RAND in den 1950er und 60er Jahren als Inkubator fungierte, um Spieltheorie jenseits militärischer Verwendungen in Management-Theorien, sowie Wirtschafts- und Sozialwissenschaften zu verankern. Die Forschung bei RAND war programmatisch darauf ausgerichtet, die Grenzen zwischen Natur- und Sozialwissenschaften abzubauen. Mediengeschichtlich bemerkenswert ist, in welchem Umfang bei RAND Probleme der Rationalisierung und Effizienz von Organisation mit Problemen von Computer Design und Programmierung verschmolzen wurden. (ebd., 188). Merrill Flood war beispielsweise nicht nur an spieltheoretischen, militärischen und sozialpsychologischen Fragestellungen interessiert, er entwickelte auch 1951 einen Computeralgorithmus für einen künstlichen Spieler, die »Stat-Rat« und beansprucht für sich, als einer der ersten den Begriff »Software« geprägt zu haben, und zwar 1946 in einem Memo für das amerikanische Kriegsministerium (ebd., 261). Dass die Spieltheorie bei RAND von zentraler Bedeutung war, geht auf direkte Einflussnahmen durch John von Neumann zurück. Dabei fungierte Spieltheorie zunächst als mathematischer Zweig von Operations Research, wurde im Umfeld ökonomischer Ansätze aber bald auch als *Entscheidungstheorie* oder *decision theory* übersetzt. Nicht zufällig trägt ein amerikanisches Standardwerk zur Einführung in die Spieltheorie von Duncan Luce und Howard Raiffa (1957) den Titel *Games and Decisions*.

<sup>8</sup> Der Name Gefangenendilemma geht vermutlich auf den Mathematiker Albert W. Tucker zurück, der in Stanford bereits 1950 ein Seminar mit dem Titel »A two-person dilemma« abhielt (vgl. Mirowski 2002, 354-358).

Abb. 1: Auszahlungsmatrix von A. W. Tucker nach Luce/Raiffa 1957, 95.

	b1	b2
a1	(0.9, 0.9)	(0,1)
a2	(1,0)	(0.1, 0.1)

Abb. 2: Originale Auszahlungsmatrix von Flood & Drescher im originalen Experiment nach Mirowski 2002, 358.<sup>9</sup>

	<b>JW Defect</b>	<b>JW Cooperate</b>
AA Cooperate	(-1, 2)	(1/2, 1)
AA Defect	(0, 1/2)	(1, -1)

Abb. 3: Zuordnungen von Auszahlungswerten und Spielentscheidungen in der Gewinnmatrix bei Rapoport 1976, 192.

	b1 (Defektion/Geständnis)	b2 (Kooperation)
a1 (Defektion/Geständnis)	-5,-5	10,-10
a2 (Kooperation)	-10,10	5,5

In der Matrix ist ablesbar, welche Strategiepaarungen welche Auszahlungswerte ergeben, wie also das Ergebnis von den kombinierten Entscheidungen von a und b abhängt. So ergibt beispielsweise in der Matrix von Rapoport die Entscheidung a1 entweder 5 Verlustpunkte (-5) für a oder aber 10 Gewinnpunkte (10), je nachdem ob b sich für b1 oder b2 entscheidet, also entweder mit a kooperiert oder gegenüber dem Staatsanwalt ein Geständnis ablegt.<sup>10</sup>

### *Die mediale Form der Spiele*

Die Doppelstruktur von abstrakter Matrix und Konflikterzählung ist charakteristisch für die spieltheoretische Darstellungsweise. Die mediale Doppelstruktur des Gefangenendilemmas stützt die Übersetzbarkeit zwischen einer mathematischen »Normalform« in tabellarischer Anordnung und einer dramatischen Erzählform mit sozialen wie psychologischen Qualifizierungen. Erzählung und Matrix repräsentieren unterschiedliche Strukturen und Modi von Wissen. Es lässt sich vermuten, wie man im Anschluss an Markus Krajewskis Thesen zum Medium der Tabelle formulieren könnte, dass die Kombination beider Formate (Erzählung und Matrix) den medialen »Funktionsmodus« (ders. 2007, 38) des Spiels innerhalb der Spieltheorie determiniert. Ihr käme damit eine entscheidende Rolle für die theoretische Leistungsfähigkeit der Spieltheorie zu. Wenn dies zutrifft, dann kann die mediale Doppelstruktur des Spiels in der Gleichzeitigkeit von Matrix und dramati-

<sup>9</sup> Die Kürzel AA und JW stehen für die Versuchspersonen, die RAND-Analysten John Williams (JW) und Armen Achian (AA), die Flood und Drescher rekrutierten, um 100 Wiederholungen des Spiels gegen bzw. miteinander zu spielen. Die Auszahlungen erfolgten in Cents (Mirowski 2002, 358.).

<sup>10</sup> Rapoport 1976, 192.

scher Situation bzw. Spielnarrativ als ein wesentliches Element der epistemologischen Produktivität der Spieltheorie gelten.<sup>11</sup>

### *Rationalitätsprobleme der Entscheidung*

Die anhaltende Popularität und Faszination des Gefangenendilemmas innerhalb der Spieltheorie ist eng mit dem verbunden, was man als ›egoistische Rationalitätsfalle‹ beschreiben könnte. Der Psychologe und Spieltheoretiker Andrew Colman hat diese Eigenschaft als »self-defeating rationality« bezeichnet (ders. 2003, 139).

Das eigentliche Dilemma des Spiels offenbart sich bei dem Versuch, analytisch die rational beste Entscheidung für jeden Spieler zu bestimmen. Wie für die Spieltheorie üblich, kann man dabei voraussetzen, dass die Spieler als Entscheidungssubjekte Erwartungen über das Verhalten ihrer Mitspieler bilden und deren erwartbare Entscheidungen in ihre ›Kalkulation‹ einbeziehen (Holler/illing 2009, 1). Vorausgesetzt wird also, dass sich die Spielenden als Entscheidungsträger der Interdependenz des Ergebnisses ihrer Entscheidungen bewusst sind (ebd.).

Ein ›rationaler‹ Spieler wird also Überlegungen ungefähr folgender Art anstellen: Wie soll ich mich verhalten, um eine möglichst geringe Gefängnisstrafe zu erhalten? Ist es am besten für mich zu gestehen? Vermutlich ja, denn falls B ein Geständnis ablegt und mich belastet, wäre ich schlecht beraten, die ganze Gefängnisstrafe auf mich zu nehmen. Wenn B aber das Verbrechen leugnet, dann ist es vorteilhaft für mich, ein Geständnis abzulegen und B zu belasten, denn in diesem Fall kann ich als Informant mit einer reduzierten Gefängnisstrafe rechnen. Wie ich es auch betrachte, die vernünftige Alternative ist das Geständnis.

Nach den Voraussetzungen der klassischen Spieltheorie muss man also davon ausgehen, dass die ›rational‹ beste Entscheidung darin besteht, den Mitgefangenen zu verraten. Verhalten sich aber beide Spieler nach dieser Kosten-Nutzen-Rechnung, werden beide in der Konsequenz ein schlechteres Ergebnis erzielen, als wenn sie beide das Verbrechen geleugnet und miteinander kooperiert hätten. Die Entscheidungsrationalität der Spieltheorie scheitert nicht nur bei der Suche nach dem besten Ergebnis, sie läuft sogar zwangsläufig darauf hinaus, dass die Gefangenen sich gegenseitig belasten. Was nach dem Rationalitätskonzept von Rational Choice, das der Spieltheorie zugrunde liegt, ›vernünftig‹ und damit zwingend als beste Strategie erscheint, führt zu einer irrationalen Entscheidung.

Für die Spieltheorie ist damit nicht nur die Frage aufgeworfen, unter welchen Bedingungen Kooperation ›vernünftig‹ erscheint, sondern weitergehend werden die impliziten Rationalitätsannahmen der Spieltheorie in Frage gestellt.

Wissenschaftsgeschichtlich wirft die Kränkung des Gefangenendilemmas auch einen Schatten auf die ökonomische Theorie, in der bis heute die Idee verfolgt wird, in der Spieltheorie ein mathematisch-theoretisches Fundament finden zu können. Tatsächlich wurde nach dem Erfolg von *Theory of games and economic behavior* die Wirtschaftstheorie eines der wichtigsten Anwendungsgebiete der Spieltheorie.

Die Nähe von Spieltheorie und Wirtschaftstheorie tritt vielleicht an keinem Punkt so deutlich hervor wie im Konzept der (egoistischen) Nutzenmaximierung. Was in der klassischen Spieltheorie

---

<sup>11</sup> Der Mathematiker und Biologe Anatol Rapoport hat ausgehend von Problemen der Spieltheorie die Möglichkeit der Interpretation von mathematischen Gleichungen am Beispiel der mathematischen Analysen von Populationswachstum erläutert. In der Darstellung als System von Gleichungen werde vom »Inhalt« (Rapoport 1976, 95) der charakteristischen biologischen Situation abgesehen, um die »mathematischen Eigenschaften« der Gleichungen zu untersuchen (ebd.). Wenn die mathematischen Gleichungen dann jedoch biologisch interpretiert werden, muss dieser ›biologische Inhalt‹ wieder hinzugefügt werden; erst dadurch kann überhaupt eingeschätzt werden, ob und in welchen Fällen die mathematische Formalisierung tatsächlich zum besseren Verständnis der biologischen Situation beiträgt oder nicht. Anders formuliert: Die Interpretation spieltheoretischer Gleichungen (oder Algorithmen) als Aussagen über wirkliche (biologische, soziale, ökonomische, psychologische, etc.) Sachverhalte ist nicht aus der Binnenlogik mathematischer Kompetenz zu leisten, sondern ist wesentlich auf die prüfende Rückbindung an Inhalte angewiesen, die in der mathematischen Abstraktion eliminiert werden. Übertragen auf das Gefangenendilemma wird damit eine Funktion der doppelten medialen Repräsentation spieltheoretischer Problemstellungen verständlich. Die Spielnarrative der Spieltheorie lassen sich in dieser Hinsicht als flexible Narrative betrachten, die über illustrative oder pädagogische Funktionen hinaus einen Deutungsspielraum für die Übertragbarkeit und Überprüfbarkeit mathematischer Erkenntnisse hinsichtlich konkreter Sachverhalte stiften.

als rationales Kalkül jedes Spielers vorausgesetzt wird, das Eigeninteresse zu verfolgen und den Eigennutzen zu maximieren, gründet auf dem gleichen Rationalitätskonzept, das der klassischen Nationalökonomie seit Adam Smith zu Grunde liegt. Das Gefangenendilemma kratzt mit dem Zweifel, den es an der ›Rational-Choice-Theorie‹ weckt, an einem Eckpfeiler neoklassischer Wirtschaftstheorie.

Auch jenseits der Kritik einer affirmativen Haltung gegenüber der Spieltheorie in den Wirtschaftswissenschaften gab es von Beginn an gute Gründe, das Dilemma einer »self-defeating rationality« ernst zu nehmen. Schließlich gehörten in der Zeit des Kalten Krieges neben der Wirtschaftstheorie spieltheoretische Strategieberatungen im Feld von Militär- und Politikberatung zu den wichtigsten Anwendungsgebieten der Spieltheorie. Strategien von Abschreckung, Vergeltung oder Erstschlag waren Teil der Militärlogik des nuklearen Wettrüstens. Das Dilemma einer sich selbst widerlegenden Entscheidungsrationalität angesichts der Wahl zwischen Konfrontation und Kooperation betraf in dieser historischen Situation nicht nur die Haltbarkeit wirtschaftswissenschaftlicher Rationalitätsannahmen oder die Vereinbarkeit von Moral und instrumenteller Vernunft, sondern das Schicksal von Nationen in einem Spiel auf Leben und Tod im Zeitalter atomarer Massenvernichtungswaffen. In den Worten Anatol Rapoport ist das Gefangenendilemma ein schlagendes Beispiel dafür, wie ein »Mangel an gegenseitigem Vertrauen in Verbindung mit völlig ›rationalen‹ Überlegungen zum Verhängnis führt« (Rapoport 1976, 192).<sup>12</sup>

Claus Pias hat mit Blick auf die zeitgeschichtliche Situation des Kalten Krieges die Relevanz von Spieltheorie für die politische und militärische Beratung als »Steuerinstrument eines alltäglich auszubalancierenden ›Gleichgewichts der Kräfte‹« beschrieben (Pias 2000, 221) und die Payoff-Matrix des Gefangenendilemmas als Rüstungskampf umgedeutet:

Abb. 4: Tabelle aus Pias 2000, 217.

	abrüsten	aufrüsten
abrüsten	5 / 5	-10 / 10
aufrüsten	10 / -10	-5 / -5

## Axelrods Turniere - eine neue Phase der Spieltheorie

Die bis in die Gegenwart anhaltende Relevanz des Gefangenendilemmas ist vielfältig mit den Interessen und Theoriebewegungen innerhalb unterschiedlicher wissenschaftlicher Disziplinen verbunden. Robert Axelrods Forschungen zum Gefangenendilemma sollten nach seiner eigenen Aussage dazu beitragen Kooperation zu fördern, besonders angesichts der Blockkonfrontation zwischen

<sup>12</sup> Es wurden vielfach unterschiedliche Auswege aus diesem Dilemma diskutiert. Anatol Rapoport hat vorgeschlagen, die »Unzulänglichkeit der individuellen Rationalität« (1976, 196) anzuerkennen, die das beste Ergebnis deshalb verfehlt, weil sie vom Prinzip des individuell-egoistischen Maximalnutzens ausgeht. Würde das rationale Kalkül der Gefangenen dagegen auf den gemeinsamen Maximalnutzen zielen, könnten die Spieler angesichts der Gewinnmatrix zu der Einsicht gelangen, dass sie kooperativ das beste Ergebnis für beide erzielen. Die egoistische Kalkulation sei zu kurzfristig, um dieses Ergebnis zu erreichen. Die Grundlagen mathematischer Entscheidungstheorie werden, das wird hier erkennbar, durch politische und soziologische Vorannahmen beeinflusst, die das Verhältnis von Individuum und Gesellschaft, Eigeninteresse und Kollektivnutzen betreffen. In der Psychologie und der Soziologie wurde das Gefangenendilemma zum Ausgangspunkt für alternative Rationalitätsmodelle, etwa durch die Berücksichtigung sozialer Prägungen oder psychologischer Einstellungen für die Entscheidungsfindung. Bis heute einflussreich ist das von Herbert Simon Ende der 1950er formulierte Konzept »begrenzter Rationalität« (1957), das zwar weiterhin von dem Ideal ausgeht, dass Entscheidung als vollständig rationaler Prozess betrachtet werden kann, jedoch die Einschränkung vornimmt, dass Entscheidungen auf der Grundlage einer bestimmten Menge verfügbarer Informationen und angesichts begrenzter Ressourcen getroffen werden müssen.

den Supermächten USA und Sowjetunion (Axelrod 1997, xi). Axelrod ist Politikwissenschaftler, hat aber Erfahrungen mit Computerprogrammierung. Durch seine Zeit an der Michigan Universität war er mit der Arbeit von John Holland vertraut, der bereits in den 1960er Jahren Methoden für die dynamische Veränderung von Populationen von Algorithmen entwickelt hatte, und die Technik der so genannten genetischen Algorithmen mit begründete.<sup>13</sup> Beachtlich ist, wie konsequent Axelrod theoretisch und methodisch neue Ansätze erprobte. Er ersann Computerturniere als Experimentieranordnungen, in denen Computerprogramme gegeneinander das Gefangenendilemma spielten, um die erfolgreichste Entscheidungsregel im algorithmischen Wettkampf zu bestimmen. Und er verwendete ökologische und evolutionäre Simulationen, um die dynamische Entwicklung hypothetischer Populationen von Computerprogrammen zu studieren. Seine 1984 veröffentlichte Studie *The evolution of cooperation* wurde nicht nur ein Sachbuch-Bestseller, sondern markiert zugleich eine neue Phase der Spieltheorie an der Schnittstelle von Ökonomie, Sozialwissenschaft, Biologie und Computerforschung.

Um für die Situation des Gefangenendilemmas geeignete Strategien zu finden, hatte Axelrod verschiedene Spieltheoretiker eingeladen Algorithmen zu entwerfen, die Regeln für die Wahl von Kooperation oder Nicht-Kooperation bei jedem Zug definierten und die als Programme für den Turnierwettkampf eingereicht werden sollten. Die 14 Programme des ersten Turniers stammten von Spieltheoretikern aus den Fächern Ökonomie, Psychologie, Soziologie, Politikwissenschaft und Mathematik. Die 62 eingereichten Programme des nachfolgenden Turniers kamen darüber hinaus von Evolutionsbiologen, Physikern, Informatikern sowie einer Anzahl von »Hobbyspielern«, die auf die Ausschreibung des Turniers in einer Heimcomputerzeitschrift geantwortet hatten. Im Turnier trafen die Programme in unterschiedlichen Paarungen aufeinander und erhielten nach jeder Runde die in der Matrix festgelegten Auszahlungen. Im ersten Turnier wurden 120000, im zweiten insgesamt eine Million Züge (= algorithmische Entscheidungen) ausgeführt.

Aus beiden Turnieren ging das Programm TIT FOR TAT (auf Deutsch etwa: Wie-Du-Mir-So-Ich-Dir), eingereicht von Anatol Rapoport, als Sieger hervor. TIT FOR TAT beginnt im ersten Zug stets mit Kooperation und in allen weiteren Zügen tut es das, was der andere Spieler im vorangegangenen Zug getan hat (Axelrod 2005, 28).

Die Einfachheit des Algorithmus hält Axelrod nicht davon ab, eine feingliedrige theoretische und moralische Interpretation von TIT FOR TAT vorzunehmen. Er hofft dadurch die generellen Erfolgsbedingungen von Kooperation in Konfliktsituationen zu erkennen. TIT FOR TAT wird in der Folge als eine Art universelle moralische Norm und soziale wie evolutionäre Erfolgsstrategie proklamiert. Dazu gehören auch Empfehlungen für nationale und internationale Politik:

»Der [...] an die Spieler des Gefangenendilemmas gerichtete Rat könnte nationalen Führern genauso dienlich sein: Sei nicht neidisch, defektiere nicht als erster, erwidere sowohl Kooperation wie Defektion und sei nicht zu raffiniert. Ähnlich mögen die [...] diskutierten Techniken der Förderung von Kooperation im Gefangenendilemma für den Bereich internationaler Beziehungen nützlich sein.« (ebd., 172)

Der Algorithmus und sein Verhalten werden psychologisch charakterisiert und anthropomorph personalisiert. So sei TIT FOR TAT »freundlich« (»nice«), weil es bei der ersten Begegnung stets Kooperation wähle; es sei »gutmütig«, »verzeihend« und »nachsichtig«, weil es nach einer Konfrontation nur einmal mit Vergeltung antworte, danach aber wieder zur kooperativen Strategie zurückkehre.<sup>14</sup> Andere algorithmische Entscheidungsregeln werden auf ähnliche Weise analysiert, beispielsweise wird das Programm TRANQUILIZER als »tückisch« beschrieben.

<sup>13</sup> Holland und Axelrod waren Mitglieder der so genannten BACH-Gruppe, bestehend aus Arthur Burks, Robert Axelrod, Michael Cohen und John Holland (vgl. Mirowski 2002, 484f.).

<sup>14</sup> »Was den robusten Erfolg von TIT FOR TAT erklärt, ist die Kombination, freundlich zu sein, zurückzuschlagen, Nachsicht zu üben und verständlich zu sein. Freundlichkeit schützt vor überflüssigen Schere-reien. Zurückschlagen hält die andere Seite nach einer versuchten Defektion davon ab, diese unbeirrt fortzusetzen. Nachsicht ist hilfreich bei der Wiederherstellung wechselseitiger Kooperation. Schließlich erleichtert Verständlichkeit die Identifikation und löst dadurch langfristige Kooperation aus« (Axelrod 2005, 48).

## *Bio-Logik der Kooperation*

Axelrods Einsatz des Computers als Forschungsinstrument stellt nicht nur eine methodische Innovation dar, sondern vollzieht sich in enger Kopplung mit einer radikalen theoretischen Umorientierung. Es geht dabei nicht nur um einen Medienwechsel (Gramelsberger 2010, 93), sondern um einen Austausch der grundlegenden Rationalitätskonzepte der Spieltheorie.<sup>15</sup> Die Abwendung vom Menschen als Entscheidungsinstanz ist dabei viel weitreichender, als es in den Computerturnieren zunächst erscheinen mag. Schließlich agierten die Algorithmen dort zumindest noch als Stellvertreter oder Repräsentationen entscheidungsfähiger Individuen. Bald wird jedoch deutlich, dass die Ersetzung der Spieler durch Computerprogramme nur einen ersten Schritt markiert. Darauf folgend werden die Ergebnisse des Turniers durch Simulationen ergänzt und schließlich in einem biologischen Theorierahmen neu interpretiert. Axelrod unterscheidet dabei zwischen *ökologischen* und *evolutionären* Simulationsverfahren.

Die ökologische Perspektive basiert auf einer Folge hypothetischer zukünftiger Runden des Turniers, in denen der Erfolg der Algorithmen in Abhängigkeit zu ihrer Umgebung untersucht wird, das heißt in Bezug auf die Verteilungshäufigkeit der anderen Algorithmen, mit denen sie zusammentreffen können. Dabei werden die Algorithmen in Analogie zu Lebewesen perspektiviert, die zu anderen Algorithmen-Lebewesen in einer biologischen Beziehung stehen.

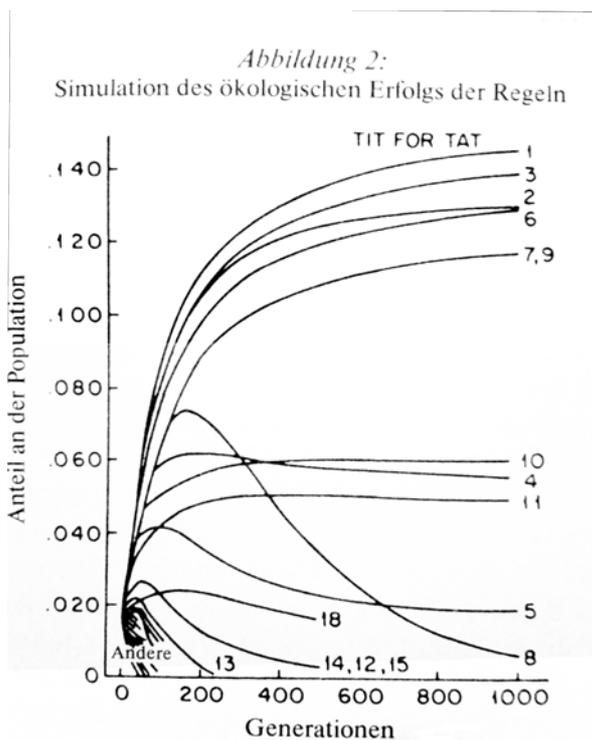
Bei Axelrod ist dieser Perspektivwechsel vom Turnierwettkampf zur biologischen Analogie wie folgt beschrieben:

»Stellen wir uns vor, daß viele Tiere einer einzigen Art ziemlich oft miteinander interagieren. Wir wollen annehmen, daß die Interaktionen die Form eines Gefangenendilemmas haben. [...] Eine Runde des Turniers kann man dann als eine Simulation einer einzelnen Generation dieser Tiere auffassen, wobei jede Entscheidungsregel von einer großen Zahl von Individuen angewendet wird. [...] Der Grundgedanke ist, daß die erfolgreicher Regeln in der nächsten Runde mit größerer und die weniger erfolgreichen mit geringerer Wahrscheinlichkeit erneut angewendet werden. Um das zu präzisieren, setzen wir voraus, daß die Anzahl der Kopien (oder Nachkommen) einer Regel ihrer Punktzahl im Turnier proportional ist. Wir interpretieren einfach die durchschnittliche Auszahlung für ein Individuum als proportional der erwarteten Anzahl seiner Nachkommen. [...] Je besser eine Strategie ist, um so stärker wächst ihr Anteil. [...] Dieser Prozeß simuliert das Überleben des Tüchtigsten.« (Axelrod 2005, 43-45)

---

<sup>15</sup> Axelrods methodisch-theoretische Neuorientierung der Spieltheorie beinhaltet einen epistemologischen Schwenk, der soziale und >evolutionäre< Norm einander annähert und das Soziale als Imitation des Lebendigen betrachtet. Damit wird auf theoretischem Gebiet vollzogen, was politisch mit dem Begriff der Biomacht bei Foucault beschrieben worden ist. Vgl. zur Mimesis des Sozialen an die Biologie im Kontext der Begriffsgeschichte der Biopolitik: Muhle 2013.

Abb. 5: Populationsentwicklung im ökologischen Gefangenendilemma aus Axelrod 2005, 45.



Das Feststellen der ›Fitness‹ der Algorithmen wird dadurch gewährleistet, dass sie in Beziehung zueinander ihr eigenes Milieu konstituieren. Der Unterschied zur »evolutionären Sichtweise« besteht darin, dass in der ökologischen Simulation keine neuen Verhaltensregeln eingeführt werden, es gibt dort keine Mutationen oder Rekombinationen, durch die neue Strategien in die ›Population‹ eingeführt werden könnten (ebd., 459).

## Evolutionäre Spieltheorie

Das Konzept Evolutionär Stabiler Strategien (= ESS) übernimmt von der Spieltheorie die mathematische »Kosten-Nutzen-Rechnung« (Dawkins 2007, 137), wie sie in der Normalform der Auszahlungsmatrix ausgedrückt wird, um die taktischen Entscheidungen, die dem Verlauf eines Kampfes vorausgehen, zu modellieren und Strategie- bzw. Taktik-Paarungen mit Auszahlungen zu versehen.<sup>16</sup> Der Simulierbarkeit des evolutionären Verlaufs wird damit über die Spieltheorie ein spezifisch ökonomisches Kalkül zugrunde gelegt. Als stabil gilt im Sinne der ESS eine Strategie dann, wenn sie durch keine zufällig als Mutation auftauchende neue Strategie verdrängt werden kann, nachdem sie erst einmal von der Mehrzahl der Mitglieder einer Population adaptiert wurde. In seiner evolutionstheoretischen Perspektive greift Axelrod das Konzept der ESS direkt auf, und führt die darin im Ansatz erfolgte Integration von Spieltheorie und Evolutionstheorie fort. Die Akteure der Spieltheorie werden dabei biologisch und evolutionstheoretisch wie folgt uminterpretiert, um den Prozess natürlicher Selektion zu modellieren: Spieler entsprechen individuellen Organismen, Strategien entsprechen dem Genotyp eines Organismus und die Auszahlungen korrespondieren den Veränderungen ihrer Darwinschen ›Fitness‹ (vgl. Colman 2003, 140). In der biologischen Interpretation der evolutionären Spieltheorie wählen Spieler ihre Strategien nicht bewusst oder freiwillig, sondern unterschiedliche Eigenschaften von Strategien führen zu unterschiedlichen Auszah-

<sup>16</sup> Methodisch bemerkenswert ist, dass bereits in dem grundlegenden Artikel zur Evolutionär Stabiler Strategie *Die Logik des Konflikts* (1973) John Maynard Smith und George R. Price ihre Argumentation auf die Ergebnisse einer »Simulationsstudie« gestützt hatten. Sie ließen fünf verschiedene Konfliktstrategien, die als Maus, Falke, Angeber, Vergelter und Wehrhafter bezeichnet wurden, in 2000 simulierten Auseinandersetzungen aufeinandertreffen, um mit den gewonnenen Daten die Bedingungen für evolutionstabile Strategien zu analysieren (vgl. Smith / Price [1973] (1996)).

lungen. Das Prinzip der freien (Wahl-)Entscheidung wird in gewisser Weise durch die Idee natürlicher Selektion als ›Auswahlprozess‹ ersetzt. Statt durch die Rationalität von Entscheidung wird das Spielergebnis durch die Bio-Logik evolutionärer Auswahl und konkurrierender Fitness-Werte bestimmt.

### *Neodarwinistische Logik quasirationaler Entscheidungen*

Im Rahmen der Spieltheorie des Gefangenendilemmas trägt Axelrod entscheidend dazu bei, der Simulation als drittem Modus der Wissensproduktion zur Anerkennung zu verhelfen, neben den etablierten Modi der theoretischen Analyse und des sozialpsychologischen Experiments. Auffällig ist jedoch die Diskrepanz zwischen der vollzogenen methodisch-theoretischen Neuorientierung und der ›humanistischen‹ Interpretation der Ergebnisse: Auf Basis evolutionstheoretischer Annahmen soll die neodarwinistische Logik des Lebendigen erkannt werden, ausgedrückt als Reproduktionserfolg, gemessen als Entwicklung der Verteilungshäufigkeit von Algorithmen innerhalb einer Population. Statt Entscheidungen als Wahl bewusster Individuen zu modellieren, werden Algorithmen verwendet, die bestimmte Strategien als Teil einer ökologischen oder evolutionären Simulation definieren, für die Entscheidungsfreiheit und Bewusstheit keine relevanten Kategorien mehr darstellen. Axelrods anschließende Beschreibung der Algorithmen mit den anthropomorph-psychologisierenden Attributen ›gutmütig‹, ›freundlich‹ etc., ist vor diesem Hintergrund bemerkenswert, da doch die *Ersetzung* von Spielern durch Programme gerade eine der wesentlichen methodischen Neuerungen seines Ansatzes darstellt. Wenn anstelle von rationalem Abwägen Algorithmen die Entscheidungen bestimmen, läuft eine nachträgliche anthropomorphe Beschreibung der Algorithmen nach dem Vorbild menschlicher Charaktereigenschaften der Anerkennung der daraus folgenden Konsequenzen zuwider.

Die oben angesprochene mediale Doppelstruktur des Spiels in der Spieltheorie macht die eigentümliche Spannung und Widersprüchlichkeit zwischen evolutionärer und anthropomorpher Deutung der Computerexperimente verständlich. Sie lässt sich als Spannung zwischen zwei unterschiedlichen Narrativen bei gleichbleibender Matrix betrachten. Das Basisnarrativ des Gefangenendilemmas stützt mit seinen impliziten Annahmen zu individueller Entscheidungskompetenz, -rationalität, -freiheit und -souveränität eine anthropomorphe Interpretation. Dagegen wird durch ein evolutionäres Narrativ die Unterminierung des Rational-Choice-Paradigmas befördert, insofern es als eine Erzählung über die Geschichtlichkeit des Lebens als Wettbewerb von Populationen im Kampf ums Dasein, jenseits von individueller Vernunft, Reflexion und Entscheidungsfähigkeit zu verstehen ist. Die Inkongruenzen zwischen diesen Narrativen offenzulegen ist nicht Axelrods Anliegen. Es würde wohl auch bedeuten, die Überzeugungskraft seiner politischen Rhetorik und der ethischen Imperative (Freundlichkeit, Nachsicht, Kooperation, Wie-Du-mir-so-ich-Dir, etc.) zu schwächen.

Richard Dawkins, populärer Evolutionsbiologe und Autor von *Das egoistische Gen*, hat die Abkehr vom Konzept bewusster Entscheidung in geradezu euphorischem Tonfall begrüßt. Als Reaktion auf Axelrods Studie notiert er:

»Ein Computerprogramm kann sich strategisch verhalten, ohne sich seiner Strategie oder überhaupt irgendeines Dinges bewusst zu sein. Wir sind natürlich mit der Vorstellung unbewußt agierender Strategen völlig vertraut, oder zumindest solcher Strategen, deren Bewußtsein, falls sie es haben, irrelevant ist. Dieses Buch [Das egoistische Gen – SW] ist voll von Strategen, denen ein Bewußtsein fehlt. Axelrods Programme sind ein hervorragendes Modell für die Art und Weise, wie wir uns in den vorangegangenen Kapiteln mit Tieren und Pflanzen, ja in der Tat mit Genen befaßt haben. [...] Niemand würde jemals behaupten, eine Bakterie sei ein bewußt handelnder Stratege, und doch spielen bakterielle Parasiten mit ihren Wirten wahrscheinlich unaufhörliche ›Gefangenendilemma-Spiele‹.« (Dawkins 2007, 375f.)

Was Dawkins Begeisterung auslöst, ist letztlich die *Kompatibilität* sozialer, ökonomischer und biologischer Forschungsansätze, die er in Axelrods Arbeit verwirklicht sieht. Methodisch zentral ist da-

bei die *algorithmische Modellierung* von Strategien, und das impliziert ein de-personalisiertes, a-subjektives und bewusstloses (programmiertes) Entscheidungshandeln.

Als Zwischenschritt kann hier festgehalten werden, dass in der Geschichte des Gefangenendilemmas ein folgenreicher Austausch zwischen ökonomischen, biologischen und sozialwissenschaftlichen Theorien und Modellvorstellungen beobachtet werden kann, der den Begriff der Entscheidung selbst an den Rand seiner Auflösung führt.

Verbunden ist dies mit einer Abwendung vom tradierten Begriff der Entscheidungsrationalität und dessen Kopplung an das Individuum, verstanden als ›Insel der Rationalität‹ (Vogl 2010, 38).<sup>17</sup> In der evolutionären Spieltheorie zeigt sich eine deutliche Verschiebung vom Pol des Individuums hin zur Population, von der Idee einer individuell zu verankernden Vernunft hin zu einer (neodarwinistisch verstandenen) Produktivität des Lebendigen. Die Idee von Rational-Choice und das damit verbundene Konzept von Rationalität werden weitgehend aufgegeben. Zugleich wird der biogeografische Begriff der *Population* entscheidend aufgewertet, neodarwinistisch und ökonomisch qualifiziert als eine Situation von Selektions- und Differenzierungsdruck, Wettbewerb und Ressourcenkampf.

### *Umstrittene Rationalitätsbegriffe: Selektion als bio-logik quasirationaler Entscheidungen*

Weil evolutionäre Spieltheorie im Ansatz von nicht-bewussten strategischen Interaktionen ausgeht und keine rationale Entscheidung im Sinne bewusster Individuen modelliert, ist umstritten, ob sie überhaupt zum Kanon der Spieltheorie zählen soll (vgl. Colman 2003, 140). Zwar argumentiert die evolutionäre Spieltheorie, wie oben dargelegt wurde, sehr wohl im Kern *ökonomisch* als Kosten-Nutzen-Analyse, der Theorie-Figur des rationalen Agenten im Sinne des homo oeconomicus verschließt sie sich aber dennoch. Sie führt eine andersartige Ontologie der Entscheidung ein, in der Populationen von Insekten, Pflanzen und sogar Populationen von Computer-Programmen evolutionäre Gleichgewichte erreichen können und Kooperationen jenseits rationaler Entscheidungsprozesse beschreibbar werden (ebd.).

Die Theorie evolutionärer Selektion versucht gerade die Bedingungen von Verhalten zu beschreiben, das »keine Vernunftfähigkeit« (Rieck 2006, 234) voraussetzt. Der Begriff der *Selektion als Auswahl* folgt im epistemologischen Modell der evolutionären Spieltheorie einem strikten (biologischen) *Funktionalismus*,<sup>18</sup> im Unterschied zum Begriff der *Wahl als Entscheidung*. Dagegen gehört in der klassischen Spieltheorie die Verkettung von Individuum, Entscheidung und Vernunft, also die Idee, dass Vernunft im Hinblick auf Entscheidungsprozesse als *Wahl* zu verstehen und letztlich auf die Rationalität eines ›freien Individuums‹ zu beziehen ist, zum Kern ihres Rationalitäts- und Entscheidungsbegriffs. Dies gilt auch oder gerade unter den Voraussetzungen des Berechenbarkeitspostulats, des Primats der Nutzenmaximierung und der Unterordnung unter einen strikten Instrumentalismus, die die klassische Spieltheorie kennzeichnen.

Nachfolgende Forschungen zum Gefangenendilemma haben versucht, die Spannung zwischen klassischer und evolutionärer Spieltheorie, die bei Axelrod als Differenz zwischen evolutionärer und anthropomorpher Interpretation von Algorithmen deutlich wurde, in die eine oder andere Richtung aufzulösen. Während einige Ansätze sich verstärkt der *psychologischen Rationalität* individueller Entscheidungsträger zuwenden, in Verbindung mit Ansätzen der Entscheidungstheorie zu ›beschränkter Rationalität‹ (Colman 2003; Colman et al. 2012) forcieren andere die ökologische und evolutionäre Simulationsperspektive, für die individuelle Rationalität und Entscheidungsfähigkeit grundsätzlich keine tragenden Konzepte mehr darstellen. Hierzu zählen beispielsweise Si-

<sup>17</sup> Wie eng dieser Vernunftbegriff historisch mit der Herausbildung einer spezifischen Wissenschaft des Ökonomischen verbunden ist, hat Vogl (2010) in seiner Studie zum Gespenst des Kapitals thematisiert. Der »interessegeleitete Mensch«, der im 17. und 18. Jahrhundert an der Schnittstelle von politischer Anthropologie, Moralphilosophie und Nationalökonomie theoretisch verfeinerte Gestalt annimmt, und später als homo oeconomicus kanonisch wird, »operiert wie eine kleine *Insel der Rationalität* in einer zufälligen und eigentlich unvernünftigen Welt« (38, herv. SW).

<sup>18</sup> Vgl. zur Diskussion der epistemologischen Dimension des soziologischen Funktionalismus in Anlehnung an den biologischen Funktionalismus weiterführend Wagner 2012, S. 60ff.

mulationsstudien zu evolutionären Algorithmen (Golbeck 2002, Haider 2005, Kristian Lindgren).<sup>19</sup> Ein dritter Weg sind agentenbasierte Simulationen, bei denen einzelne Agenten als algorithmisch definierte Entscheidungsträger konzipiert werden können. Deren Rationalität ist dann zunehmend eine Frage des Designs (vgl. Axelrod 1997).

### *Bio-algorithmische Epistemologie?*

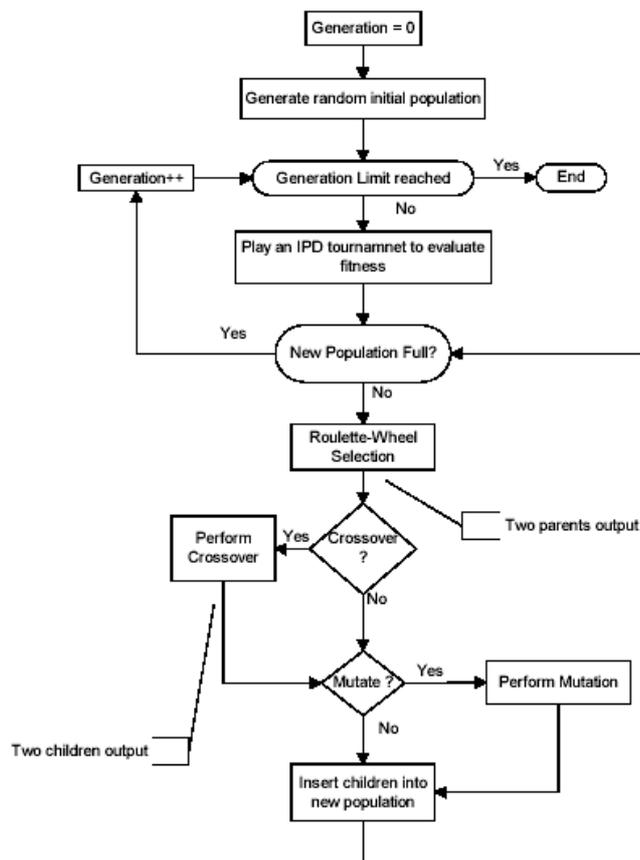
Zusammenfassend lässt sich festhalten, dass Axelrods ökologische und evolutionäre Perspektive Verschiebungen auf mehreren Ebenen impliziert. Die Matrix wird zum Konstruktionselement von Milieus, wenn es statt um Auszahlungswerte von Strategiepaarungen um den Reproduktionserfolg und die Häufigkeitsverteilung von Algorithmen innerhalb von Populationen geht. Die Wahl (choice) wird zur Auswahl (selection), wenn Rationalität nicht mehr an Individuen als bewusste Entscheidungssubjekte gebunden ist, sondern als quasirationaler Effekt evolutionärer Gesetzmäßigkeiten verstanden wird. In der evolutionären Perspektive treten nicht mehr einzelne Programme gegeneinander an, sondern Populationen, die in Konkurrenz zueinander ihre biologische Fitness beweisen. Statt um die Entscheidungen individueller konkreter oder abstrakter Akteure geht es um Populationsdynamiken von Vielheiten innerhalb eines evolvierenden Milieus.

Axelrods Studie zum Gefangenendilemma markiert einen Scheitelpunkt im Hinblick auf die Bedeutung, Zurechnung und Interpretation von ›Entscheidung‹ zwischen realen Subjekten und Algorithmen, zwischen Personalisierung und De-Personalisierung. In nachfolgenden Forschungen wurde das Problem des Gefangenendilemmas entsprechend zunehmend als ein Optimierungsproblem von Algorithmen begriffen.

---

<sup>19</sup> Eine umfassende Einführung in evolutionäre Algorithmen als Methode der Informatik bietet Weicker 2007. Bei evolutionären Algorithmen dient die natürliche Evolution als Vorbild. In der »simulierten Evolution« sollen durch das »Wechselspiel zwischen Modifikation und Auswahl« bessere Individuen erzeugt und identifiziert werden (ebd., 1). Insbesondere werden Konzepte wie »Population, Reproduktion durch Vererbung und Variation, das Prinzip der Selektion sowie die Kodierung von Informationen in einem Genotyp« der Biologie entlehnt (ebd., 39). Die Anpasstheit oder »Fitness« wird nicht notwendig durch die Anzahl von Nachkommen gemessen, sondern kann durch beliebig definierte »Bewertungsfunktionen« erfolgen, passend zu dem »Optimierungsproblem« für das eine algorithmische Lösung gesucht wird. Weiterhin wird typischerweise eine Selektionsoperator definiert, der aus einer Population eine bestimmte Anzahl von Individuen auswählt, die für die Erzeugung von Nachkommen verwendet werden. Der Begriff der Population hat eine zentrale Funktion, weil damit »eine Ansammlung von Lösungskandidaten« bezeichnet wird, die als Individuen betrachtet werden (ebd., 1).

Abb. 6: Flow-Chart Prisoner's Dilemma mit Genetischen Algorithmen aus Haider 2005, II.



Die evolutionäre Spieltheorie unterscheidet sich gegenüber der klassischen Spieltheorie nicht nur durch die Hinwendung zu Simulationstechnologie als Instrument für Modellbildungen und theoretische Experimente; wichtig ist die Beobachtung, dass die Umstellung auf Algorithmen und Computer als Medien der Entscheidung *gleichzeitig* mit einer theoretischen Umorientierung auf evolutionsbiologische Konzepte erfolgt. Ein Kreuzungspunkt dieser doppelten Verschiebung ist die prinzipielle Gleich-Gültigkeit der bio-algorithmischen Perspektive gegenüber Menschen, Tieren, Pflanzen, Bakterien und Computerprogrammen. Denn weder für die Algorithmen als formale Operationalisierungen von (Entscheidungs-)Handlungen noch im flexiblen Evolutionsnarrativ macht es einen grundlegenden Unterschied, ob es sich um die Modellierung der Populationsdynamiken von Bakterien, Menschen oder »egoistischen Genen« handelt.

Gerade im Vergleich mit der evolutionären Spieltheorie wird deutlich, wie sehr in der klassischen Spieltheorie der Begriff von Entscheidung mit einem ökonomisch konturierten Rationalitätskonzept von Rational Choice gekoppelt ist. In der evolutionären Modellierung des Gefangenendilemmas ist diese Verknüpfung aufgelöst. Wenn man nach der »Rationalität« bio-algorithmischer Entscheidungen fragen wollte, wäre diese, wenn überhaupt, vermutlich nur auf der Ebene von Populationen oder als Ergebnis evolutionärer Prozesse zu bestimmen. Im Rahmen post-souveräner Rationalitätskonzepte werden vormalige »Entscheidungssubjekte« und ihre tradierten Theoriefiguren vielleicht bald als evolutionäre »Bricolage von Algorithmen« (Mirowski 2002, 516) und zufällig aufgelesenen Subroutinen zu denken sein.<sup>20</sup>

<sup>20</sup> Mirowski verwendet den Ausdruck »bricolage of algorithms«, um die projizierte Auflösung des homo oeconomicus in der Wirtschaftstheorie angesichts der Umwandlung der Wirtschaftswissenschaften zur titelgebenden »Cyborg Science« zusammenzufassen. Der Zusammenhang bei Mirowski ist auf die Situation der ökonomischen Theorie zugespißt (Mirowski 2002, 516), ich verwende den Ausdruck hier in einem weiter gefassten Sinne für die möglichen Effekte der bio-algorithmischen Komposition von Entscheidung und Rationalität.

## Souveränität und Konflikt

Die hier vorgestellte vergleichende Analyse ging von der These aus, dass es zum Kennzeichen der Spieltheorie gehört, Spiele als Medien der Entscheidung zu konzipieren. Die formalisierte Behandlung von Entscheidungen ist innerhalb der Spieltheorie voraussetzungsreich. Wie unter anderem deutlich wurde, ist die besondere Verwendung, die von Spielen gemacht wird, mindestens eine doppelte. Durch die Formalisierung wird der Spielbegriff an Konzepte von Berechenbarkeit angelehnt, Spiele werden im Sinne der Spieltheorie algorithmisierbar. Damit kann der Computer als Medium ein wichtiger Bestandteil spieltheoretischer Argumentations- und Beweisführung werden. Spiele sind darüber hinaus als Szenarien zu verstehen, die konflikthafte Situationen beschreiben, erzählen und modellieren, in denen Entscheidungen den Fortgang und das Ergebnis bestimmen. Am Beispiel des Standardnarrativs des Gefangenendilemmas als Kriminalgeschichte wird dies besonders anschaulich. Die Auszahlungsmatrix als Normalform der Formalisierung und das Konfliktnarrativ der Verhör- und Geständnissituation wirken zusammen, um das Gefangenendilemma spieltheoretisch zu qualifizieren. Nachvollziehbar wird hierüber wie Spiele sowohl narrativ, inszenatorisch, als auch formalisiert, berechenbar, als Medien der Entscheidung verwendet werden.

In der Gegenüberstellung von evolutionärer und klassischer Spieltheorie zeigt sich entsprechend eine zweifache Verschiebung oder Transformation. Sowohl das Narrativ wird abgewandelt – aus der Kriminalgeschichte zweier Gefangener wird eine Erzählung über die evolutionär modellierter Konkurrenz zwischen Populationen – als auch die Art der Formalisierung und Berechnung ändert sich. Zwar bleibt die Matrix als Funktion der Mathematisierung erhalten, aber die Strategien der Entscheidung werden algorithmisch verkörpert und mit Verfahren untersucht, die nur noch mit Computern handhabbar sind. Diese Transformation weist in eine Richtung, in der Entscheidung und Entscheidungsrationalität zunehmend abgelöst werden von der Vorstellung individueller Entscheidungsobjekte, um auf Populationen, Vielheiten übertragen zu werden, die algorithmisch definiert und interpretiert werden.

Wenn man den Versuch wagen wollte, aus der Spezifik der Theoriegeschichte des Gefangenendilemmas allgemeinere Thesen zu Medien der Entscheidung abzuleiten, dann ließe sich das Gefangenendilemma möglicherweise als exemplarischen Beleg dafür verstehen, dass Medien der Entscheidung auf vielfältige Weise mit Vorstellungen von Rationalität, Wahl, Konflikt, Vernunft, Souveränität sowie mit Erkenntnis- und Wissensproduktion verbunden sind. Im Vergleich zwischen Medien der Entscheidung oder zwischen unterschiedlichen geschichtlichen Varianten derselben kann erkennbar werden, wie jeweils mit Konzepten von Vernunft, Wahl, Rationalität etc. umgegangen wird und welche Politiken und Konjunkturen sie durchlaufen.

Der Blick auf *Medien* der Entscheidung kann darüber hinaus dazu beitragen, nicht nur unterschiedliche mediale Funktionen und Formen (wie Vermittlung, Erzählung, Modellierung, Berechnung, Prozessierung) zu benennen, sondern darüber hinaus mediale Konstellationen als Ausdruck historisch-gesellschaftlicher Reflexion und damit aus analytischer Perspektive als Teil der Selbstbeschreibung einer Gesellschaft zu verstehen.

Denkt man in dieser Richtung über die Transformationen des Entscheidungsbegriffs nach, wie er in der Kontrastierung von klassischer und evolutionärer Spieltheorie deutlich wird, rückt das gespannte Verhältnis zwischen Algorithmen und Souveränität in den Blick. Wenn Entscheidungen nicht mehr bewussten Individuen, sondern Algorithmen zugewiesen werden, die über Selektion Autorität erlangen, kann dies zu Beunruhigung Anlass geben. Zu bedenken ist dabei, dass die hier zur Diskussion gestellten Veränderungen von Rationalitäts- und Entscheidungskonzepten selbst bereits historisch sind und seit Ende der 1990er Jahre vollständig ausgearbeitet vorliegen. Insofern sie anschlussfähig erscheinen an zeitgenössische Analysen zur Macht von Algorithmen, etwa im Zusammenhang mit Überwachung, sozialen Medien, Affective Computing und Big Data, deutet dies darauf hin, dass unsere technologisierte Gesellschaft möglicherweise zunehmend wird lernen müssen, ihre eigenen Zukunftsentscheidungen im Schatten bio-algorithmischer Logik und post-souveräner Rationalität zu (er)finden.

## Bibliografie

- Axelrod, Robert (1997): *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton, N.J.: Princeton University Press.
- Axelrod, Robert [1984] (2005): *Die Evolution der Kooperation*. (6. Aufl.). München: Oldenbourg.
- Berninghaus, Siegfried / Ehrhart, Karl-Martin / Güth, Werner (2010): *Strategische Spiele. Eine Einführung in die Spieltheorie*. (3., verb. Aufl.). Berlin; Heidelberg: Springer.
- Berry, David M. (2011): *The Philosophy of Software. Code and Mediation in the Digital Age*. New York: Palgrave Macmillan.
- Colman, Andrew M. / Pulford, Briony D. (2012): Problems and Pseudo-Problems in Understanding Cooperation in Social Dilemmas. In: *Psychological Inquiry* 23/1, S. 39–47.
- Dawkins, Richard (2007): *Das egoistische Gen*. (Jubiläumsausg.). München; Heidelberg: Elsevier, Spektrum, Akad. Verl.
- Ensmenger, N. (2011): Is chess the drosophila of artificial intelligence? A social history of an algorithm. In: *Social Studies of Science* 42/1, S. 5–30.
- Flood, Merrill (1952): *Some experimental Games*. RM789-  
[http://www.rand.org/content/dam/rand/pubs/research\\_memoranda/2008/RM789-1.pdf](http://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM789-1.pdf)  
(abgerufen am 09.04.2015).
- Gillespie, Tarleton (2016): Algorithm (preprint). In: Ben Peters (Hg.): *Digital Keywords*. N.J.: Princeton University Press. Online: <http://culturedigitally.org/2014/06/algorithm-draft-digitalkeyword/> (abgerufen am 09.01.2016).
- Golbeck, Jennifer (2002): Evolving Strategies for The Prisoner's Dilemma. In: *Advances in Intelligent Systems, Fuzzy Systems, and Evolutionary Computation 2002*, S. 299–306.
- Gramelsberger, Gabriele (2010): *Computereperimente. Zum Wandel der Wissenschaft im Zeitalter des Computers*, Bielefeld: Transcript.
- Haider, Adnan (2005): *Using Genetic Algorithms to Develop Strategies for The Prisoners Dilemma*, <http://mpira.uni-muenchen.de/28574/> (abgerufen am 17.11.2014).
- Holler, Manfred J. / Illing Gerhard: *Einführung in die Spieltheorie*. (7. Aufl.). Berlin; Heidelberg: Springer.
- Krajewski, Markus (2007): In Formation. Aufstieg und Fall der Tabelle als Paradigma der Datenverarbeitung. In: Michael Hagner et al. (Hg.): *Nach Feierabend. Zürcher Jahrbuch für Wissenschaftsgeschichte*. Bd. 3. Zürich; Berlin: diaphanes, S. 37–55.
- Lindgren, Kristian (1991): Evolutionary Phenomena in Simple Dynamics. In: Christopher Langton / J.D. Taylor / S. Rasmussen (Hg.): *Artificial Life II. SFI Studies in the Sciences of Complexity*, Bd. X, Reading, Mass. u.a.: Addison-Wesley, S. 295–312.
- Luce, Robert Duncan / Raiffa, Howard (1957): *Games and Decisions: Introduction and Critical Survey*. New York: Dover Publications.
- Mérő, László (1998): *Moral Calculations: Game Theory, Logic, and Human Frailty*. (übers. von A. C. Gösi-Greguss). New York, NY: Copernicus.
- Mirowski, Philip (2002): *Machine Dreams: Economics Becomes a Cyborg Science*, Cambridge [u.a.]: Cambridge University Press.
- Mueller, Ulrich (Hrsg.) (1990) *Evolution und Spieltheorie*. München: Oldenbourg Wissenschaftsverlag.
- Muhle, Maria (2013): Vom Vitalen zum Sozialen. Überlegungen zu einem politischen Wissen im Anschluss an Canguilhem. In: Astrid Deuber-Mankowsky / Christoph F. E. Holzhey (Hg.): *Situiertes Wissen und regionale Epistemologie: zur Aktualität Georges Canguilhems und Donna J. Haraways*. Wien: Verlag Turia + Kant (Cultural inquiry 7), S. 141–159.
- Neumann, John von (1928): Zur Theorie der Gesellschaftsspiele. In: *Mathematische Annalen*, Band 100, 1928, S. 295–320.
- Neumann, John von / Morgenstern, Oskar [1944] (2007): *Theory of Games and Economic Behavior*. Princeton, New Jersey: Princeton University Press.
- Pias, Claus (2000): *Computer-Spiel-Welten*, Weimar: Universität, <ftp://ftp.uni-weimar.de/pub/publications/diss/Pias/pias.pdf>.
- Rapoport, Anatol (1976): *Kämpfe, Spiele und Debatten: 3 Konfliktmodelle*. Darmstadt: Verlag Darmstädter Blätter.

- Rieck, Christian (2006): *Spieltheorie. Eine Einführung*. (5. Aufl.). Eschborn: Rieck.
- Simon, Herbert Alexander (1957): *Models of Man. Social and Rational. Mathematical Essays on Rational Human Behavior in Society Setting*. New York: Wiley.
- Smith, John Maynard/ Price, George R. [1973] (1996): Die Logik des Konflikts. In: Ulrich Mueller (Hg.): *Evolution und Spieltheorie*. München: Oldenbourg, S. 15-23.
- Turing, A. M. (1950): Computing Machinery and Intelligence. In: *Mind* 59/236, S. 433–460.
- Vogl, Joseph (2010): *Das Gespenst des Kapitals*. Zürich; Berlin: diaphanes.
- Wagner, Gerhard (2012): *Die Wissenschaftstheorie der Soziologie. Ein Grundriss*. München: Oldenbourg.
- Weicker, Karsten: *Evolutionäre Algorithmen*. (2., überarb. und erw. Aufl.). Wiesbaden: Teubner.
- Wiemer, Serjoscha (2014): Strategiespiele und das Medium-Werden des Computers. Computerschach als Faktor der Rekonzeptionalisierung des Computers als programmierbare Maschine zwischen 1945 und 1960. In: Stefan Böhme / Rolf F. Nohr / Serjoscha Wiemer (Hg.): *Diskurse des strategischen Spiels. Medialität, Gouvernementalität, Topografie*. Münster: LIT, S. 83–112.
- Ziegenbalg, Jochen / Ziegenbalg, Oliver / Ziegenbalg, Bernd: *Algorithmen von Hammurapi bis Gödel. Mit Beispielen aus den Computeralgebra-Systemen Mathematica und Maxima*. (3., überarb. Aufl.). Frankfurt a. M.: Deutsch.