

Universität Bielefeld/IMW

**Working Papers
Institute of Mathematical Economics**

**Arbeiten aus dem
Institut für Mathematische Wirtschaftsforschung**

Nr. 173

**Über Informationsprobleme
bei der Implementation
von Mechanismen**

von

Walter Trockel

April 1989



H. G. Bergenthal

**Institut für Mathematische Wirtschaftsforschung
an der**

Universität Bielefeld

Adresse / Address:

Universitätsstraße

4800 Bielefeld 1

Bundesrepublik Deutschland

Federal Republic of Germany

Abstract

The present paper provides a discussion of Nash-implementation under the specific aspect of necessary informational requirements for the agents. Recent literature is discussed and it is indicated how the present informational requirements might be related so as to become more realistic.

Key words: Mechanism, Game Form, Nash-equilibrium, Incomplete Information.

1. Das Implementationsproblem

Ausgangspunkt unserer Überlegungen ist eine Gesellschaft, deren Ziele durch eine soziale Auswahlregel beschrieben sind. Eine solche Regel bestimmt für jede mögliche Konfiguration individueller Charakteristika eine Menge realisierbarer Zustände der Gesellschaft. Man interpretiert die Elemente der ausgewählten Menge als "sozial erwünscht" oder "optimal".

Sind die relevanten Charakteristika aller Individuen allgemein bekannt, also insbesondere auch dem zentralen Planungsbüro (Planer), dann kann die soziale Auswahlregel implementiert werden, indem ein Element der zu dieser bekannten konkreten Konfiguration von Charakteristika gehörenden Menge von Optima realisiert wird.

Die Implementation wird erst dann zum Problem, wenn, wovon man im allgemeinen ausgehen muß, dem Planer diese Charakteristika nicht bekannt sind. Versucht er nämlich durch Befragung der Individuen deren Charakteristika kennenzulernen, so kann er im allgemeinen nicht erwarten, die Wahrheit zu erfahren. Da nämlich die soziale Auswahlregel beschreibt, wie der Planer die von den Individuen erhaltene Information verarbeitet, besteht für diese ein Anreiz, die Antwort strategisch und damit möglicherweise nicht wahrheitsgemäß zu wählen. Die Wahrheit zu sagen, ist für die Individuen nicht notwendig "anreizverträglich".

Aus diesen Überlegungen ergibt sich für den Planer das Implementationsproblem. Er muß versuchen, ein von jeder konkreten Konfiguration individueller Charakteristika unabhängiges System von Spielregeln, einen sogenannten Mechanismus, zu entwerfen, derart, daß die Gesellschaft, wenn alle Individuen im Rahmen dieser Spielregeln agieren, einen im Sinne der Auswahlregel optimalen sozialen Zustand realisiert. Dabei wird vorausgesetzt, daß dem Planer die Menge aller realisierbaren sozialen Zustände bekannt ist. Ein solcher Mechanismus ist so konzipiert, daß der Planer bei der Implementation der sozialen Auswahlregel auf Signale der Individuen und deren Auswertung angewiesen ist. Formalisiert wird ein solcher Mechanismus durch eine Vorschrift ("Ergebnisfunktion"), die jeder Liste individueller Signale einen realisierbaren sozialen Zustand zuordnet.

Es wird vorausgesetzt, daß der Mechanismus des Planers allgemein bekannt ist. Jedes Individuum weiß, daß und wie seine eigene strategische Wahl eines Signals zusammen mit den von den anderen Individuen gewählten Signalen den sozialen Zustand festlegt. Damit sind fast alle Ingredienzen eines nicht-kooperativen Spiels gegeben: Strategiemengen (die Mengen möglicher Signale) für die Individuen und eine Regel, die jeder Liste von Strategien ein Ergebnis (den sozialen Zustand) zuordnet. Ein solcher formaler Rahmen bildet eine Spielform. Während die Spielform von der konkreten Konfiguration individueller Charakteristika unabhängig ist, macht jede konkrete Gesellschaft eine solche Spielform zu einem von den Charakteristika ihrer Individuen abhängigen Spiel. Die Präferenzen der Individuen auf den sozialen Zuständen werden durch Nutzenfunktionen beschrieben. Zusammen mit dem Mechanismus wird (durch Hintereinanderausführung) somit für jedes Individuum eine "Auszahlungsfunktion" definiert.

Die Individuen der konkret realisierten Gesellschaft werden zu Spielern eines nicht-kooperativen Spiels. Unterstellt man ein von allen Spielern akzeptiertes Lösungskonzept für solche Spiele, dann ist es das Ziel des Planers, die Spielform so zu entwerfen, daß in allen möglichen resultierenden Spielen die Lösungen soziale Optima induzieren. Ein(e) solche(r) Spielform (Mechanismus) implementiert dann "die soziale Auswahlregel".

Je nach Spezifizierung der Information der Individuen, der Mengen möglicher Signale, des unterstellten spieltheoretischen Lösungskonzeptes und der Eigenschaften der unterstellten sozialen Auswahlregel ergeben sich verschiedene formale Rahmenbedingungen mit unterschiedlichen mathematischen Strukturen, welche für den jeweils betrachteten konkreten ökonomischen Anwendungsbereich mehr oder weniger geeignet sein können.

Zusätzlich zu der Annahme, daß der Mechanismus allgemein bekannt ist ("common knowledge" ist präzise definiert in Aumann (1976)), kennt jeder Spieler zumindest seine eigenen Charakteristika (Präferenz, Erstausrüstung, etc.). Bei vollständiger Information sind alle Charakteristika unter allen Spielern allgemein bekannt. Auch in diesem Falle gibt es durchaus Informationsprobleme, solange der Planer nicht ebenfalls vollständig informiert ist. Unabhängig von seinem Ziel, beim Entwurf des Mechanismus die richtigen Anreize zu setzen, besteht für ihn das Problem, wieviel an Information von den Spielern an das zentrale Planungsbüro gesendet werden muß, damit der Mechanismus wirksam sein kann. Dies ist das Problem einer effizienten Kodierung der Information, die zu Arbeiten über minimale Dimension der Signalmengen geführt hat (vgl. Reichelstein (1984), Williams (1986)). Das natürliche Lösungskonzept im Kontext

solcher alle Spieler betreffenden vollständigen Information ist das des Nash-Gleichgewichts. In Spielen mit eindeutigem Nash-Gleichgewicht hat kein rationaler Spieler einen Grund, eine andere Strategie zu wählen, da jeder aufgrund der vollständigen Information seine optimale Strategie ermitteln kann. Fehlende Information über die Präferenzen oder Strategiemengen der anderen Spieler bedeutet aber, daß der betreffende Spieler gar nicht weiß, welches Spiel er spielt. Insbesondere kann er kein Nash-Gleichgewicht ermitteln. Das führt auf natürliche Weise dazu, das Impelementationsproblem für andere Lösungskonzepte zu betrachten. Wenn alle Spieler dominante Strategien besitzen, die also optimal sind ganz gleich, was die jeweils anderen tun, dann ist die vollständige Information natürlich nicht erforderlich. Ein anderes Lösungskonzept im Kontext unvollständiger Information ist das des Bayes-Gleichgewichts. Hier hat jeder Spieler subjektive Wahrscheinlichkeiten über eine Menge möglicher Charakteristika jedes Mitspielers, der sogenannten Typen. Jede Typenkonstellation würde, wenn sie allgemein bekannt wäre, ein Spiel mit vollständiger Information repräsentieren. Da diese vollständige Information aber nicht vorliegt, werden die subjektiven Wahrscheinlichkeiten eines jeden Spielers über den Typenmengen der jeweils anderen zum Teil seiner Charakteristika. Auf dieser Basis lassen sich Spiele mit unvollständiger Information definieren. Das Nash-Gleichgewicht in diesem neuen Rahmen ist das Bayes-Gleichgewicht (vgl. d'Aspremont und Gerard-Varet (1979a,b), Myerson (1985), Repullo (1986)).

Ich werde nicht im folgenden im wesentlichen mit der Implementation in Nash-Gleichgewichten beschäftigen. Dabei werde ich die Probleme, insbesondere Informationsproblem, bei der Nash-Implementierung erörtern. Anschließend wird ein formaler Rahmen geliefert, in dem Maskins Charakterisierungstheorem für Nash-Implementation formuliert werden kann. Die Arbeit schließt mit der Präsentation und Diskussion neuerer Sätze, die auf der Basis des Maskin Theorems Nash-Implementation bei unvollständiger Information behandeln.

2. Nash-Gleichgewicht und Nash-Implementation

Es ist einleuchtend, daß ein Mechanismus, der eine soziale Auswahlregel implementieren soll, nicht besser oder akzeptabler ist, als das zugrunde liegende spieltheoretische Lösungskonzept. Dies gilt insbesondere auch für das Nash-Gleichgewicht. Das Nash-Gleichgewicht ist das natürliche, grundlegende Lösungskonzept für nicht-kooperative Spiele. Es hat dennoch gewisse Schwächen, die für die Beurteilung des Konzepts der Nash-Implementierung wichtig sind. Der schwerwiegendste Nachteil des Nash-Gleichgewichts ist der, daß im allgemeinen Spiele eine Vielfalt verschiedener Nash-Gleichgewichte besitzen. Selbst wenn rationale Spieler die zugrunde liegende Philosophie dieses Lösungskonzeptes akzeptieren, hindert diese Mehrdeutigkeit sie daran, ohne Koordination die "richtige" Gleichgewichtsstrategie zu wählen. Der Wunsch nach Einschränkung dieser Vielfalt verschiedener Gleichgewichte führte zu Verfeinerungen (d.h. speziellen Typen) von Nash-Gleichgewichten, was gleichzeitig ein Aussondern gewisser implausibler Gleichgewichte und Strategien bedeutete. Die Einschränkung auf perfekte Gleichgewichte (und Modifikationen) führt im allgemeinen jedoch nur zu einer Verkleinerung der Mengen betrachteter Gleichgewichte, nicht zu ihrer Reduzierung auf ein einziges Gleichgewicht. Diese Eindeutigkeit wäre jedoch idealerweise für eine wirksame Implementation notwendig. Die Auswahl eines eindeutigen Gleichgewichts erfolgt im Buch von Harsanyi und Selten (1988). Dort werden auch die Schwächen des Nash-Gleichgewichts ausführlich diskutiert. Das von Harsanyi und Selten eingeführte eindeutige Gleichgewicht ist noch zu neu, als daß es von der Implementationsliteratur bereits hätte aufgegriffen werden können. Aber auch dieses eindeutige Gleichgewicht und die verschiedenen Verfeinerungen sind natürlich spezielle Nash-Gleichgewichte. Diese Tatsache unterstreicht die fundamentale Bedeutung dieses Lösungskonzepts.

In der Literatur herrscht weitgehend Übereinstimmung darüber, daß die Verwendung des Nash-Gleichgewichts voraussetzt, daß die Struktur des Spiels und die Charakteristika aller Spieler allgemein bekannt sind (vgl. Maskin (1985)). Eine interessante alternative Rechtfertigung des Nash-Gleichgewichts, die nicht auf dieser "common knowledge" Annahme beruht, findet man bei Stinchcombe (1988). Ein Ausweg aus der Mehrdeutigkeit wird häufig darin gesehen, daß es oft einen "natürlichen" oder "ausgezeichneten" Kandidaten unter den Gleichgewichten gibt, auf den sich jeder Spieler in nicht bindenden Vorabgesprächen im eigenen Interesse festzulegen bereit ist (vgl. Postlewaite (1985)).

Die Fähigkeit, ein konkretes Nash-Gleichgewicht wirklich zu ermitteln, hängt abgesehen von Fragen algorithmischer Komplexität davon ab, welche Information über die Charakteristika der jeweils anderen Spieler jeder Spieler besitzt. Es ist durchaus möglich, daß zur Berechnung nicht die exakte Kenntnis der Charakteristika aller anderen notwendig ist. Diese Tatsache gewinnt mit zunehmender Spieleranzahl an Bedeutung. Selbst approximative Kenntnis aller Charakteristika ist im Kontext großer Ökonomien eine heroische Annahme. Es ist besonders im Kontext großer Ökonomien wichtig, wenn jeder Spieler in der Lage ist, seine Nash-Strategie auf der Basis gewisser "aggregierter" Information über die anderen berechnen zu können. Daß solche Aggregationsfunktionen über die jeweils anderen existieren, ist Teil der Axiome des Theorems in Dubey, Mas-Colell & Shubik (1980). Spiele, die diese Eigenschaft haben wurden auch in Experimenten verwendet (van Huyck, Battalio & Beil (1988a,b), vgl. auch Crawford (1989)).

Diese Vorstellung, daß auch im Falle nicht vollständiger Information der Spieler über die Spieler das Nash-Gleichgewicht noch ein sinnvolles Lösungskonzept sein kann, findet sich auch in der jüngsten Literatur zur Nash-Implementation, wo sie zur Erweiterung des Charakterisierungssatzes von Maskin (Maskin (1985), Repullo (1986), Williams (1986)) geführt hat. Auf Maskins Theorem und diese neueren Beiträge werde ich im nächsten Abschnitt eingehen.

Zuvor sind jedoch noch einige Bemerkungen zur Nash-Implementierung angebracht.

Das Konzept eines Gleichgewichts in dominanten Strategien ist natürlich aus der Sicht des Planers grundsätzlich dem eines normalen Nash-Gleichgewichts vorzuziehen. Wenn jeder Spieler weiß, daß seine gewählte Strategie für ihn optimal ist, ganz gleich, was alle anderen tun, besteht intuitiv ein viel stärkerer Anreiz für ihn, nicht abzuweichen als im Nash-Gleichgewicht, indem er sich zwar durch Abweichen nicht verbessern kann, aber ggfs. von aus irgendwelchen Gründen erfolgenden Abweichungen anderer betroffen sein kann.

Der eigentliche Grund, warum man überhaupt Nash-Implementation statt der attraktiveren Implementation durch Gleichgewichte mit dominanten Strategien betrachtet, ist ein fundamentales Unmöglichkeitstheorem für Implementation in dominanten Strategien für bestimmte "direkte" Mechanismen. Direkte (subdirekte) Mechanismen sind solche, bei denen die Menge der möglichen Signale eines Individuums identisch ist mit

der Menge seiner Charakteristika (bzw. Präferenzen) (vgl. Repullo (1986)). Hurwicz hatte in seiner bahnbrechenden Arbeit (Hurwicz (1972)), in welcher er auch das Konzept der Anreizverträglichkeit einführte, bewiesen, daß es kein individuell rationalen anreizverträgliches die Effizienz implementierenden Mechanismus gibt. Dieses Unmöglichkeitsergebnis betrifft insbesondere den kompetitiven Mechanismus und schließt damit die Implementierbarkeit von Walras-Gleichgewichten in dominanten Strategien aus. Ein ähnlicher Unmöglichkeitssatz gilt jedoch nicht für Implementation in Nash-Gleichgewichten. Vielmehr existieren Sätze, die die Nash-Implementierbarkeit von Walras-Gleichgewichten oder im Kontext öffentlicher Güter von Lindahl-Gleichgewichten gewährleisten.

Die Hoffnung, das Unmöglichkeitstheorem von Hurwicz durch Ausweichen auf beliebige indirekte Mechanismen unter Beibehaltung des Gleichgewichts in dominanten Strategien zu vermeiden, wird durch das sogenannte "revelation-principle" zerstört. Dieses Prinzip, das in modifizierten Formen auch für Nash- und Bayes-Gleichgewichte und Maximin-Gleichgewichte gilt (Maskin (1985), Myerson (1985), Repullo (1986)) besagt, daß es zu jedem indirekten einen äquivalenten direkten Mechanismus gibt, daß man also "ohne Beschränkung der Allgemeinheit" sich auf direkte Mechanismen beschränken kann. Diese sehr vage Beschreibung führt jedoch häufig zu Fehlinterpretationen! (vgl. dazu Maskin (1985), Postlewaite (1985), Postlewaite und Schmeidler (1987)). Für Nash- und Bayes-Gleichgewichte gilt das revelation principle uneingeschränkt nur für direkte Mechanismen, für subdirekte Mechanismen nur unter einer restriktiven zusätzlichen Annahme. Außerdem gilt für diese Gleichgewichte nur eine schwache Form dieses Prinzips. Wenn eine soziale Auswahlregel durch einen indirekten Mechanismus Nash-implementiert werden kann, dann gibt es einen direkten Mechanismus, der diese Auswahlregel wahrheitsgemäß (truthfully) implementiert. Wahrheitsgemäße Implementation bedeutet aber nur, daß die Liste der Signale, welche die wahren Charakteristika angeben, ein Gleichgewicht bildet, welches in einem Element der Auswahlmenge resultiert. Das direkte Spiel kann aber durchaus Gleichgewichte besitzen, die außerhalb der Auswahlmenge liegen und zudem für alle Spieler besser sind (vgl. Repullo (1986), Postlewaite und Schmeidler (1987)). Eine solch schwache Implementierbarkeit ist natürlich nicht sehr überzeugend.

Natürlich ist der Bedarf an direkten Mechanismen nicht mehr so klar, wenn man die Wahrheit ohnehin nicht erzwingen kann.

Bei der Nash-Implementation gilt also das Hauptaugenmerk der Verwendung solcher Mechanismen und zugehörigen Signalmengen, die ein möglichst geringes Maß an Information bei den Spielern benötigen, einerseits um den Mechanismus effizient zu gestalten, andererseits um nicht zu unrealistische Anforderungen an den Informationsstand der Spieler zu stellen.

3. Das Charakterisierungs-Theorem von Maskin

Es sei $N = \{1, \dots, n\}$ die Menge der Individuen in einer Gesellschaft. Für $i \in N$ sei $E_i \neq \emptyset$ die Menge seiner möglichen Charakteristika. Genauer gesagt bezeichnet jedes i zunächst nur eine Rolle, die dann bei festem $e_i \in E_i$ durch das durch e_i repräsentierte Individuum ausgefüllt wird.

$E = \prod_{i \in N} E_i$ ist die Menge der möglichen Profile (Listen) $e = (e_1, \dots, e_n)$ individueller Charakteristika. Sei A mit $\# A \geq 2$ die Menge aller realisierbaren sozialen Zustände.

Eine soziale Auswahlregel wird durch eine Korrespondenz $F : E \rightarrow A$ beschrieben. Jedes e_i enthält (oder induziert) eine Präferenz \succsim_i (vollständige Präordnung) auf A .

$L(a, e_i) := \{b \in A \mid b \succsim_i a\}$. Für beste Zustände $e \in A$ gilt somit $A = L(a, e_i)$.

Sei $M = \emptyset$ die Menge der möglichen Signale (message space).

$\mu_i : E_i \rightarrow M, i = 1, \dots, n$, sind individuelle Signalkorrespondenzen.

$\mu : E \rightarrow M$, definiert durch $\mu(e) := \prod_{i \in N} \mu_i(e_i)$, ist die Signalkorrespondenz.

μ wird als surjektiv vorausgesetzt. Das ist keine Beschränkung, da man sonst $M = \bigcup_{e \in E} \mu(e)$ definieren würde.

Eine Ergebnisfunktion $h : M \rightarrow A$ ordnet jeder Liste von Signalen einen sozialen Zustand zu.

(μ, h) nennen wir den Signalprozess.

Def.: Der Signalprozess (μ, h) realisiert die soziale Auswahlregel $F : \Leftrightarrow h \circ \mu = F$

Diese Realisierung ignoriert die individuellen Anreizprobleme und betrachtet nur die kommunikativen Aspekte der Implementierung (vgl. Williams (1986)).

Wir kommen nun zur eigentlichen Implementation.

Seien $S_i \neq \emptyset, i=1, \dots, n$ Strategiemengen der Individuen $i \in N$.

$S := \prod_{i \in N} S_i$ ist der Strategienraum. Die Elemente s_i der S_i sind individuelle Strategien, die Elemente $s = (s_1, \dots, s_n) \in S$ sind Strategien-Profile.

Eine Abbildung $g : S \rightarrow A$ heißt Spielform (auf A).

Da die e_i durch Nutzenfunktionen u_i repräsentierbare Präferenzen induzieren, definiert ein Paar (g, e) ein Spiel $(S, (u_i \circ g)_i \in N)$, das wir auch kurz mit (g, e) bezeichnen. Die Menge E ist also Parametermenge für eine Familie von Spielen.

Ordnet man jedem $e \in E$ die Menge der Nash-Gleichgewichte im durch e und g definierten Spiel zu, so erhält man die Nash-Gleichgewichtskorrespondenz

$$N_g : E \rightarrow S \text{ mit } N_g(e) = \bigcap_{j \in N} \{ \bar{s} \in S \mid \forall s_j \in S_j : (s_j, \bar{s}_{-j}) \in L(g(\bar{s}), e_j) \}$$

Dabei ist $(s_j, \bar{s}_{-j}) := (\bar{s}_1, \dots, \bar{s}_{j-1}, s_j, \bar{s}_{j+1}, \dots, \bar{s}_n)$.

Def.: Die Spielform g Nash-implementiert die soziale Auswahlregel $F : \Leftrightarrow g \circ N_g = F$.

$NE(g)$ bezeichne die Menge aller Elemente von S , die für mindestens ein $e \in E$ ein Nash-Gleichgewicht bilden, d.h.

$$NE(g) := \bigcup_{e \in E} N_g(e)$$

Sei $g^* := g|_{NE(g)}$ die Einschränkung von g auf $NE(g)$. Dann können wir feststellen:

Wenn F durch g Nash-implementiert wird, dann realisiert der Signalprozess (N_g, g^*) die Auswahlregel F .

Das Problem der Nash-Implementation läßt sich auf folgende Weise formulieren:

Gegeben ein Signalprozess (μ, h) , der F realisiert, wann kann $M = \mu(E)$ als Menge von Nash-Gleichgewichten und μ als Nash-Gleichgewichts-Korrespondenz eines geeigneten "größeren" Spiels implementiert werden?

Diese Fragestellung führt zu folgender Definition:

Def.: Sei (μ, h) ein Signalprozess der die soziale Auswahlregel F realisiert. Sei g eine Spielform, die F Nash-implementiert.

g "enthält" (μ, h) , wenn eine injektive Abbildung ("Einbettung")

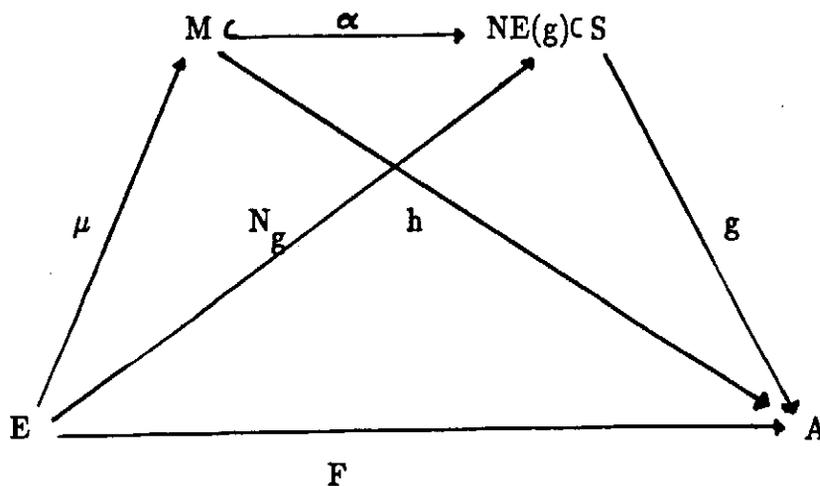
$\alpha: M \rightarrow S$ existiert, derart daß

i) $\forall e \in E: \mu(e) = N_g(e) \cap \alpha(M)$

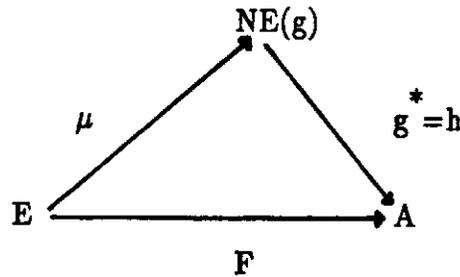
ii) $g|_M = h$

g stimmt mit (μ, h) überein, wenn sogar gilt

i') $\forall e \in E: \mu(e) = N_g(e)$



Dieses von Williams (1986) entlehnte Diagramm illustriert die obige Definition von "enthält". Bei "Übereinstimmung" erhält man $\alpha(M) = NE(g)$. Wegen $N_g = \mu$ vereinfacht sich das Diagramm dann wie folgt



Bei Übereinstimmung fallen also Realisierbarkeit und Nash-Implementierbarkeit zusammen. F ist Nash-implementierbar, wenn es zerlegbar ist in Nash-Gleichgewichts-Korrespondenz und Spielform.

Für den Charakterisierungssatz benötigt man zwei weitere Eigenschaften.

Def.: Die soziale Auswahlregel F ist monoton : \Leftrightarrow

$$\forall e, e' \in E \forall a \in F(e) : [\forall i \in N : L(a, e_i) \subset L(a, e'_i)] \Rightarrow a \in F(e')$$

Def.: Die soziale Auswahlregel schließt Veto-Rechte aus: \Leftrightarrow

$$\forall e \in E \forall a \in A : [\#\{i \in N | L(a, e_i) = A\} \geq n - 1] \Rightarrow a \in F(e).$$

Charakterisierungs-Theorem für Nash-Implementierbarkeit

(Maskin (1985), Repullo (1987), Williams (1986)):

Sei $F : E = \prod_{i \in N} E_i \rightarrow A$, $n \in \mathbb{N}$ eine surjektive Auswahlregel. Wenn F Nash-implementierbar ist, ist F monoton. Wenn $n \geq 3$ und $2 \leq \# A < \infty$ gilt und wenn F monoton ist und Veto-Rechte ausschließt, dann ist F Nash-implementierbar.

Die von Maskin ursprünglich konstruierte Spielform kann allerdings für einen vollständigen und korrekten Beweis nicht verwendet werden. Durch Wahl eines größeren Strategienraums konnte Williams (1986) Maskins Theorem beweisen. Er erlaubt zudem durch eine zusätzliche Annahme an A die Ausdehnung auf unendliche Mengen sozialer Zustände. In Maskin (1985) ist die für Spieler i gewählte Strategiemenge $S_i = E \times A$. Jeder Spieler gibt als Signal ein Profil von Charakteristika an, die in diesem Falle nur aus Präferenzen bestehen ($e_i = \xi_i$) und einem nicht notwendig erwünschten Zustand a (d.h. $a \notin F(e)$ ist möglich!). Wie Williams zeigt reicht aber dieser Strategienraum nicht aus. Williams verwendet mit Erfolg die Menge $S_i := E \times A \times \{0, \dots, n-1\}$. In jedem Fall beruhen die Konstruktionen auf der Verwendung von Signalen, die alle Präferenzen betreffen. Es wird unterstellt, daß alle Präferenzen allgemein bekannt sind. Die Verwendung ganzer Zahlen als Teile der Signale beim Design von Mechanismen findet man auch in Postlewaite und Schmeidler (1986). Sie dienen dazu, von Abweichungen abzuschrecken, da jeder Spieler bei jeder gegebenen Konstellation von Abweichungen aller anderen Spieler vom Nash-Gleichgewicht durch Wahl einer geeigneten Zahl seinen favorisierten Zustand durchsetzen kann.

4. Nash-Implementation nach Maskin

Eine wesentliche Reduzierung der für die Spieler bei Nash-Implementation bisher vorausgesetzten Information gelang Saijo (1988). Er reduziert die von Williams verwendete Strategiemenge wesentlich. Saijo verwendet $S_i := E_i \times E_{i+1} \times A \times N$. Durch Identifizierung von $n + 1$ mit 1 erhält man eine zyklische Anordnung, in der jeder Spieler genau einen direkten Vorgänger und einen direkten Nachfolger besitzt. Eine Strategie (bzw. ein Signal) besteht nun aus einer Mitteilung über die eigene Präferenz, über die des direkten Nachfolgers, eines sozialen Zustandes und einer Zahl $i \in N$. Wiederum spielen die Zahlen von 1 bis n die Rolle, daß in gewissen Situationen inkonsistenter Signale jeder Spieler bei gegebenen Signalen aller anderen Spieler die Möglichkeit hat, den sozialen Zustand zu bestimmen. Das macht solche Situationen für alle Spieler unattraktiv und setzt den Anreiz, nicht abzuweichen.

Eine Schlüsselrolle in Saijos Konstruktion spielt die Tatsache, daß für jeden Spieler zwei Präferenzsignale an den Planer übermittelt werden. Wenn $n - 1$ Signale konsistent sind, betrachtet der Planer den verbleibenden Spieler als Abweichler. Dessen Präferenzsignal wird von der Spielform ignoriert. Saijo (1988) beweist folgenden

Satz: Sei $n \geq 3$. Wenn die soziale Auswahlregel F monoton ist und Veto-Rechte ausschließt, dann gibt es eine Spielform mit den Strategiemengen

$$S_i := E_i \times E_{i+1} \times A \times N, \text{ welche } F \text{ Nash-implemertiert.}$$

Eine Strategie hat also die Form $s_i = (e_i^i, e_i^{i+1}, a_i, m_i)$.

Mit e_i^j wird das von i verkündete j betreffende Signal bezeichnet.

Die Spielform g wird wie folgt konstruiert:

Regel 1: Wenn für alle $i \in N$ gilt $e_i^i = e_{i-1}^i$ und $a_i = a$ und wenn $a \in F(e_1^1, \dots, e_n^n)$ dann gilt $g(s) := a$.

Regel 2: Wenn für alle $i \in N \setminus \{j, j + 1\}$ gilt $e_i^i = e_{i-1}^i$, für alle $i \in N \setminus \{j\}$ gilt $a_i = a$ und $a \in F(e_1^1, \dots, e_{j-1}^{j-1}, e_{j-1}^j, e_{j+1}^{j+1}, \dots, e_n^n)$, dann

$$\text{gilt } g(s_j, s_{-j}) = \begin{cases} a_j & \text{falls } a_j \in L(a, e_{j-1}^j) \\ a & \text{andernfalls} \end{cases}$$

Regel 3: Wenn weder Regel 1 noch Regel 2 anwendbar ist, dann gilt

$$g(S) = a_i \text{ mit } i - 1 = \sum_{h \in N} m_h \pmod{n}.$$

Es ist relativ leicht, nachzuweisen, daß für alle $e \in E$ gilt $F(e) \subset g \circ N_g(e)$.

Der Beweis der umgekehrten Inklusion ist ziemlich komplex und beruht teilweise auf Modifikationen von Maskins Konstruktion.

Aus Regel 1 sieht man unmittelbar, daß die Angabe der wahren Präferenzen e für alle Spieler gekoppelt mit einem beliebigen für alle identischen $a \in F(e)$ bei beliebigem $m = (m_1, \dots, m_n)$ ein Nash-Gleichgewicht für (g, e) bildet. Wie aber soll ein Spieler i , der zwar F und g kennt, selbst wenn alle die wahren Präferenzen signalisieren, wissen, welches a Element von $F(e)$ ist, wenn er e nicht kennt?

Das Ergebnis von Saijo ist sicherlich nicht geeignet, Nash-Implementation auch im Falle unvollständiger Information der Spieler zu proklamieren. Unterstellen wir aber den alten Maskinschen restriktiven Rahmen vollständiger Information, dann ist Saijos Ergebnis sicherlich nützlich. Es erlaubt die dezentrale Implementation über eine Spielform mit wesentlich kleinerem Strategienraum. Das ist besonders bei großer Spieler-Anzahl wesentlich. Die Dezentralisierung der sozialen Auswahl durch Nash-Implementation setzt zwar vollständige Information aller Spieler über alle Spieler voraus. Es muß aber nur ein Bruchteil dieser Information signalisiert werden. Der Rest der Information wird benötigt für die Entscheidung, welches Signal gesendet wird. Natürlich bleibt immer noch das Problem der Mehrdeutigkeit. Das aber gibt es bereits im Maskinschen Ansatz und unabhängig von Implementationsproblematik überhaupt im rein spieltheoretischen Rahmen.

Man kann zeigen (Trockel (1989)), daß in Saijos Theorem der Signalteil, der die Präferenz des Nachfolgers betrifft, durch eine aggregierte Information über alle anderen Präferenzen ersetzt werden kann.

Das wäre in der Tat sehr nützlich, wenn die für das Entscheidungsproblem benötigte Information eines Spielers mit der für das Kommunikationsproblem benötigten übereinstimmen würde. Es ist nämlich wesentlich realistischer, gewisse vage Informationen über alle anderen Spieler vorauszusetzen als völlig exakte Information über bestimmte Personen (etwa den Nachfolger).

Angesichts der für das Entscheidungsproblem notwendigen vollständigen Information ist diese Modifikation jedoch nicht mehr als eine technische Variante des Ergebnisses von Saijo. Allerdings führt die Möglichkeit, aggregierte Information zu signalisieren, zu einer interessanten Erweiterung der Ergebnisse von Matsushima (1987), auf die ich jetzt eingehen möchte.

4.2. Probabilistische soziale Auswahlfunktionen

Bei Matsushima sind die sozialen Auswahlregeln nicht Korrespondenzen sondern Funktionen, die zudem probabilistisch sind: Es wird eine Lotterie über sozialen Zuständen ausgewählt. Hier kommen auf natürliche Weise von Neumann-Morgenstern-Nutzenfunktionen ins Spiel. In diesem kardinalen Kontext erhält Matsushima einen Nash-Implementations-Satz. Sein Hauptergebnis ist starke Nash-Implementation in einem entsprechend modifizierten ordinalen Rahmen.

4.2.1 Der kardinale Rahmen

Wir behalten die bisherige Notation und Terminologie weitgehend bei. Für $i \in N$ sei $e_i = u_i$, wobei u_i eine von Neumann-Morgenstern Nutzenfunktion auf A bzw. der Menge Δ der Wahrscheinlichkeitsmaße ("Lotterien") auf A bezeichnet. A wird als endlich vorausgesetzt. $a \in A$ und $\rho_a \in \Delta$ mit $\rho_a(a) = 1$ werden identifiziert. Für alle $i \in N$ und alle $u_i \in E_i$ wird vorausgesetzt, daß gilt:

$$a, b \in A, a \neq b \Rightarrow u(a) \neq u(b).$$

Die von u_i induzierte Präferenz $R(u_i)$ auf A ist also eine vollständige Ordnung! Zudem sei $E' := E_i = E_j \quad \forall i, j \in N$, also $E = E'^N$.

Def.: Eine (probabilistische) soziale Auswahlfunktion ist eine Funktion $\varphi: E \rightarrow \Delta$

Für $e \in E$ wird $\varphi(e)$ als "sozial erwünscht" oder "optimal" interpretiert.

Eine Spielform ist nun eine Abbildung $g: S \rightarrow \Delta$

Eine strategische Vorschrift für Spieler i ist eine Abbildung $\nu_i: E \rightarrow S_i$.

$\nu: E \rightarrow S$, definiert durch $\nu(e') = (\nu_i(e'))_{i \in N}$ heißt strategische Vorschrift.

Das Paar (g, ν) bezeichnet man als Schema.

ν heißt Nash-Gleichgewichts-Selektion, wenn gilt: $\forall e \in E: \nu(e) \in N_g(e)$.

Def.: Ein Schema (g, ν) Nash-implementiert die soziale Auswahlfunktion φ wenn gilt:
 $\varphi = g \circ \nu$ und ν ist Nash-Gleichgewichts-Selektion. Eine soziale Auswahlregel φ ist Nash-implementierbar, wenn ein Schema (g, ν) existiert, welches φ Nash-implementiert

Matsushima beweist folgenden

Satz: Jede (probabilistische) soziale Auswahlfunktion ist Nash-implementierbar.

Im Beweis konstruiert Matsushima eine Spielform wie folgt:

$s_i := e'^i := (u'_{i-1}, u'_i, u'_{i+1}) \forall i \in N$ und alle $e' \in E$ mit der Konvention: $n + 1 = 1$.

Also $S_i = E'^3$ für alle $i \in N$. $\nu(e') := (e'^i)_{i \in N}$.

$g(s) := \varphi(e')$, falls für mindestens $n - 1$ Spieler $i \in N$ gilt $s_i = e'^i$.

Also folgt für alle $e' \in E: \varphi(e') = g(s) = g(\nu(e'))$.

Der Planer hat also bei jeder konsistenten Liste aus $n - 1$ Strategien die Möglichkeit, die eindeutige n -elementige Liste e' zu bestimmen.

Die Nutzenfunktion jedes Spielers wird dabei aus den konsistenten Signalen der jeweils anderen Spieler ermittelt. Man erhält für alle $i \in N$ und alle $s_i \in S_i$ und alle $e' \in E$

$$u_i(g(e'^j)_{j \in N}) = u_i(g(s_i, (e'^j)_{j \neq i})).$$

Abweichen von der konsistenten Signalliste aller anderen Spieler bringt also keine Verbesserung. Jede solche konsistente Signalliste ist also ein Nash-Gleichgewicht. Dies gilt insbesondere für das Strategienprofil, in dem alle die Wahrheit sagen.

Natürlich ist ohne Kommunikation und Koordination ein konsistentes Lügen aller sehr unwahrscheinlich. Die Wahrheit bildet einen sogenannten "focal point". Der Anreiz, die Wahrheit zu sagen ist daher in der Tat ziemlich groß.

Anders als im Kontext der Arbeit von Saijo (1988) stimmt hier die zur Kommunikation mit dem zentralen Planungsbüro notwendige individuelle Information mit der für das individuelle Entscheidungsproblem notwendigen Information überein.

Auch dieser Satz beruht auf der Voraussetzung der genauen Kenntnis nicht nur der eigenen sondern auch der Präferenzen zweier anderer Spieler. Zudem muß sich diese Kenntnis über andere auf eine ganz spezifische Weise gleichmäßig über alle Spieler verteilen.

4.2.2 Der ordinale Rahmen

Wir gehen von der Situation des kardinalen Rahmens in Abschnitt 4.2.1. aus. $R(u_i)$ sei die von u_i induzierte vollständige Ordnung auf A . Mit Ω bezeichnen wir die Menge aller solcher vollständigen Ordnungen.

Für $\rho, \rho' \in \Delta$ definieren wir ausgehend von $R \in \Omega$:

$$\rho \bar{R} \rho' : \Leftrightarrow \rho = \rho' \text{ oder } \forall u \text{ mit } R(u) = R: u(\rho) > u(\rho')$$

Da \bar{R} bei Identifizierung von ρ_a mit $\rho_a(a) = 1$ und a auf A gerade R induziert, bezeichne ich \bar{R} fortan durch R .

Eine ordinale soziale Auswahlfunktion ist eine Funktion $f : \Omega^n \rightarrow \Delta$. Wir setzen $E_i = \Omega$ für alle $i \in N$, also $E = \Omega^n$.

$$e := (R_i)_{i \in N} \in \Omega^n = E.$$

Wenn wir eine soziale Auswahlregel φ wie in Abschnitt 4.2.1 betrachten, für die gilt

$$\varphi(u_1, \dots, u_n) = \varphi(u'_1, \dots, u'_n) \text{ falls } \forall i \in N : R(u_i) = R(u'_i),$$

dann definiert φ eine ordinale soziale Auswahlfunktion durch $f(R(u_i)_{i \in N}) := \varphi(u_i)_{i \in N}$.

Ein Implementationssatz für den ordinalen Rahmen, analog zu dem im kardinalen Rahmen von Abschnitt 4.2.1. läßt sich beweisen, wenn die dortige Kenntnis der u_i durch die der R_i ersetzt wird. Auch die in 4.2.1. diskutierten Verallgemeinerungen sind möglich (Trockel (1989)).

Dabei Nash-implementiert ein Schema (g, ν) eine ordinale soziale Auswahlfunktion f wenn gilt $g(\nu(u_1, \dots, u_n)) = f(R_1, \dots, R_n)$ für alle $e \in \Omega^n$ und alle R_i repräsentierenden Nutzenfunktionen u_i , $i \in N$.

Matsushimas Hauptergebnis bezieht sich auf starke Nash-Implementierbarkeit.

Def.: Ein Schema (g, ν) Nash-implementiert eine ordinale soziale Auswahlfunktion f stark, wenn gilt:

- i) (g, ν) Nash-implementiert f
- ii) $\forall e \in E : \# g(N_g(e)) = 1$

Eine soziale Auswahlregel heißt stark Nash-implementierbar, wenn es ein Schema (g, ν) gibt, daß f stark Nash-implementiert.

Die starke Nash-Implementation gewährleistet aber nur eindeutige Nash-Gleichgewichts-Auszahlungsvektoren, keine Eindeutigkeit des Nash-Gleichgewichts. Obwohl der Implementationsbegriff in der Tat wesentlich stärker ist, wissen die Spieler dennoch nicht, welche der "gleich guten" Nash-Strategien die anderen Spieler wählen.

Unter einer abgewandelten Monotoniebedingung sowie dem Ausschluß von Veto-Rechten erhält man in diesem Rahmen einen zum Satz von Maskin analogen Implementationssatz.

Def.: Eine ordinale soziale Auswahlfunktion ist monoton im weiteren Sinne

$$: \Leftrightarrow \forall e, e' \in E : [\forall \rho \in \Delta : f(e) R_{\rho} f(e') \Rightarrow f(e') = f(e)]$$

Def.: Eine ordinale soziale Auswahlregel f schließt Veto-Rechte aus

$$: \Leftrightarrow \forall e \in E \forall a \in A : [C \subset N, \#C \geq n-1, a \text{ ist } R_{\rho}\text{-optimal } \forall i \in C] \Rightarrow [f(e)(a) = 1]$$

Satz (Matsushima (1987)): Eine stark Nash-implementierbare ordinale soziale Auswahlfunktion f ist monoton im weiteren Sinne. Ist f monoton im weiteren Sinne und schließt Veto-Rechte aus, dann ist f stark Nash-implementierbar.

Man kann diesen Satz jedoch auch beweisen, wenn man für jeden Spieler die Kenntnis nur einer Präferenz außer der eigenen unterstellt. Weit wichtiger erscheint mir jedoch die Tatsache, daß auch hier ein Beweis möglich ist, bei dem exakte Kenntnis der Präferenzen anderer durch aggregierte Information über alle ersetzt werden kann. Man kann, da A endlich ist, die Nutzenfunktion auf A mit einer endlichen Menge ganzer Zahlen indizieren. Die Summe der Indices der Nutzenfunktionen aller anderen reicht als Signal neben dem Index der eigenen Nutzenfunktion aus, um Nash-Implementierbarkeit in diesem Rahmen zu beweisen (Trockel (1989)).

Auch hier beruht der Beweis auf der Konstruktion einer Spielform in der jeder Spieler Signale (Strategien) verwendet, welche seine eigene Präferenz und die des Vorgängers und des Nachfolgers betreffen. Auch hier läßt sich die Aussage beweisen, wenn man Signale, die präzise Aussagen über bestimmte andere Spieler enthalten, ersetzt durch präzise Signale der Spieler über sich selbst und aggregierte Information über die jeweils anderen (Trockel (1989)).

Vermutlich kann man eine allgemeine maßtheoretische Bedingung angeben, so daß ein entsprechendes Ergebnis auch für unendliche Mengen A sozialer Zustände hergeleitet werden kann.

Es läßt sich zeigen (Matsushima (1987), Proposition 2), daß für eine generische Menge ordinaler sozialer Auswahlfunktionen f eine Spielform konstruiert werden kann, die f eindeutig stark Nash-implementiert, wenn immer das Präferenz-Profil e hinreichend "konfliktreich" ist. Letzteres ist der Fall, wenn der Ausschluß von Veto-Rechten eine leere Bedingung ist.

Der probabilistische Ansatz von Matsushima führt zu Ergebnissen, die in starkem Kontrast zu einigen bekannten Unmöglichkeitssätzen im deterministischen Kontext stehen (Gibbard (1977), Muller & Satterthwaite (1977)). Ein Analogon bekommt man, wenn man f in unserem Rahmen "deterministisch" macht, d.h. wenn gilt

$$\forall e \in E \exists a \in A : f(e)(a) = 1.$$

Neben dem probabilistischen Aspekt ist insbesondere die Beschränkung auf Funktionen statt Korrespondenzen für die positiven Ergebnisse verantwortlich.

Wir haben in der Tat einen Rahmen gefunden, in dem sehr unvollständige Information über das zu spielende Spiel ausreicht, soziale Ziele dennoch durch Nash-Gleichgewichte zu implementieren.

Literaturverzeichnis

- Aumann (1976): Agreeing to Disagree, Annals of Statistics, 4, 1236-1239
- Crawford (1989): An "Evolutionary" Explanation of van Huyck, Battalio & Beil's "Experimental Results on Coordination", manuscript, University of California, San Diego
- Dasgupta, Hammond und Maskin (1979): The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility, Review of Economic Studies, 46, 181-216
- d'Aspremont und Gerard-Varet (1979a): On Bayesian Incentive Compatible Mechanisms, in: Laffont (ed.): Aggregation and Revelation of Preferences, North-Holland, 269-288
- d'Aspremont und Gerard-Varet (1979b): Incentives and Incomplete Information, Journal of Public Economics, 12, 25-45
- Dubey, Mas-Colell & Shubik (1980): "Efficiency Properties of Strategic Market Games: An Axiomatic Approach", Journal of Economic Theory, 22, 339-362
- Gibbard (1977): Manipulation of Schemes that mix voting with Chance, Econometrica, 45, 665-681
- Groves und Ledyard (1987): Incentive Compatibility since 1972, in: Groves, Radner & Reiter (eds.): Information, Incentives & Economic Mechanisms, University of Minnesota Press, 48-111
- Harsanyi und Selten (1988): A General Theory of Equilibrium Selection in Games, MIT Press
- Homann (1989): Incomplete Contracts and Nash-Implementation in the Case of Two Agents, IMW-Working Paper No. 170, Universität Bielefeld
- Hurwicz (1972): On Informationally Decentralized Systems, in: Radner & McGuire (eds.): Decision and Organization, North-Holland, 297-336
- Maskin (1985): The Theory of Implementation in Nash Equilibrium: A Survey, in: Hurwicz, Schmeidler & Sonnenschein (eds.): Social Goals and Social Organization, Cambridge University Press, 173-204
- Matsushima (1988): A new approach to the Implementation Problem, Journal of Economic Theory, 45, 128-144
- McKelvey (1985): Game Forms for Nash Implementation of General Social Choice Correspondence, Social Science Working Paper No. 579, California Institute of Technology
- Muller und Satterthwaite (1977): The Equivalence of Strong Positive Association and Strategy-Proofness, Journal of Economic Theory, 14, 412-418

- Myerson (1985): Bayesian Equilibrium and Incentive-Compatibility: An Introduction, in: Hurwicz et al. (eds.) (vgl. Maskin (1985))
- Postlewaite (1985): Implementation via Nash Equilibria in Economic Environments, in: Hurwicz et al. (eds.) (vgl. Maskin (1985))
- Postlewaite und Schmeidler (1986): Implementation in Differential Information Economics, Journal of Economic Theory, 39, 14-33
- Postlewaite und Schmeidler (1987): Differential Information and Strategic Behavior in Economic Environments: A General Equilibrium Approach, in: Groves et al. (eds.) (vgl. Groves & Ledyard (1987))
- Reichelstein (1984): Information and Incentives in Economic Organizations, Ph.D. dissertation, Northwestern University
- Repullo (1986): On the Revelation Principle under Complete and Incomplete Information, in: Binmore & Dasgupta (eds.): Economic Organizations as Games, Basil Blackwell
- Repullo (1987): A Simple Proof of Maskin's Theorem on Nash Implementation, Social Choice and Welfare, 4, 39-41
- Schmeidler (1980): Walrasian Analysis via Strategic Outcome Functions, Econometrica, 48, 1585-1593
- Saijo (1987): On Constant Maskin Monotonic Social Choice Functions, Journal of Economic Theory, 42, 382-386
- Saijo (1988): Strategy Space Reduction in Maskin's Theorem: Sufficient Conditions for Nash Implementation, Econometrica, 56, 693-700
- Stinchcombe (1988): Game Theory without Common Knowledge Assumptions: A Rational Expectations Approach to Nash Equilibria, Discussion Paper No. 88-56, Dept. of Economics, University of California, San Diego
- Trockel (1989): Informational Requirements for Nash Implementation (in Vorbereitung)
- van Huyck, Battalio & Beil (1988a): "Strategic Uncertainty, Equilibrium Selection Principles and Coordination Failure in Average Opinion Games", Working Paper No. 88-20, Texas A & M University
- van Huyck, Battalio & Beil (1988b): Tacit Coordination Games, Strategic Uncertainty and Coordination Failure, Working Paper No. 88-23, Texas A & M University
- Williams (1986): Realization and Nash Implementation: Two Aspects of Mechanism Design, Econometrica, 54, 139-151