Game Theory

Joachim Rosenmüller and Walter Trockel

Institute of Mathematical Economics
IMW
University of Bielefeld
D-33615 Bielefeld
Germany

March 1, 2001

Keywords: *allocation, apportionment, Aumann economy, axiomatic approach, balanced games, bargaining solution, characters, coalitional form, convex games, cooperative game, coordination game, core, cost sharing, directed games, equivalence principle, exchange economy, extensive form, evolutionary stability, fictitious play, game form, homogeneous games, implementation, incentive compatible, Kalai–Smorodinsky solution, knowledge, LP–game, Maschler–Perles solution, mechanism, modiclus, Nash equilibrium, Nash solution, normal form, nucleolus, outcome function, payoff function, reduced games, repeated games, revelation principle, Shapley value, simple games, social choice rule, TU-game, vNM–Stable Set, voting games, Walrasian equilibrium,*

Contents

Glossary

Axiomatic Approach: The attempt to uniquely characterize a solution concept on a class of games by a set of mathematical axioms the interpretation of which is considered self–evident.

Bargaining Problem: An NTU-game where the only admissible coalitions are singletons and the grand coalition.

Bargaining Sets: A class of solution concepts for NTU-games containing only payoff vectors that are immune against admissible alternatives proposed by coalitions as objections.

Coalitional Function: A mapping that assigns utilities or outcomes to coalitions about which these coalitions can enter binding agreements.

Common Knowledge: The paradigm of a situation in which the players can fully observe their opponents knowledge, their observations about ones own knowledge etc. on arbitrary levels of hierarchy.

Cooperative Solution: A mapping that assigns utilities to players in view of their bargaining and coalitional power as described by the coalitional function.

Cost Sharing: In Game Theoretical terms the application of solution concepts of cooperative game theory towards the fair or equal distribution of disutilities among cost generating divisions or sections of a firm.

Equilibrium: A stable situation described in terms of strategies, payoff vectors, or consumption decisions from which to deviate is not profitable for a player or a group of players.

Economy: A specification of data thought to completely describe an economic system thereby representing the exogeneous restrictions for economic agents' interaction in markets.

Equivalence Principle: A class of results characterizing Walrasian (competitive) equilibria of large economies as (elements of) game theoretic solutions.

Evolutionary Stability: A property of strategies in a symmetric bimatrix game assuring that a strategy is a best response to itself that moreover beats other best responses as a response to them.

Extensive Form: The representation of a game describing a stochastic process susceptible to the players successive actions and resulting in intermediate and final payoffs to them.

Fictitious Play: The prototype of learning models in a social context: In a sequence of game repetitions each player chooses a strategy that is an optimal response to an estimated mixed strategy combination of the other players derived from observation of their past behavior.

Game: The paradigm of human competitive and cooperative interaction calling for decisions (under uncertainty) of the players involved and resulting in payoffs or outcomes for these players.

Game Form: A list of strategy sets for n players, interpreted as sets of admissible messages, together with an outcome function associating with any strategy profile an outcome interpreted as a social state.

Implementation: A simultaneous realization of a social choice rule by non-cooperative equilibria for a whole class of games generated by a given game form together with a family of preference profiles.

Incentive Compatibility: A game form (or mechanism) is incentive compatible if it is direct, i.e. strategy sets are the sets of possible preferences or types, and truth telling is the unique equilibrium.

Mechanism: A device processing signals of the players permitting to correlate their announcements and resulting in a cooperative decision, thereby possibly resulting in strategic behavior of players anticipating the structure of this device. Sometimes the term is used as a synonym for "game form".

Nash Program: A research agenda in game theory, initiated by John Nash, that tries to characterize certain payoff vectors for players alternatively via axiomatic cooperative solutions and via non-cooperative equilibria, thereby making both approaches more transparent.

Normal Form: The representation of a game describing the strategic alternatives available to the players and the payoff functions defined on these alternatives. No binding agreements available.

Nucleoli: A class of (nonlinear) solution concepts based on a fair or equalizing assessment of complaints to be lodged by coalitions against proposals for utility assigments.

Pareto Efficiency: The most popular axiomatic property of social outcomes that makes any individual improvement impossible unless at least some other persons are worse off.

Revelation Principle: A group of results asserting that any social choice rule that can be implemented in Nash or Bayesian Nash equilibrium allows for an alternative thruthful implementation in a direct game.

Social Choice Rule: A mapping associating with any admissible profile of preferences on a set of social states a subset of this set interpreted as socially desired.

Solution: A mapping associating with any member of a class of games some set of feasible payoff vectors for the players.

Strategy: A complete plan of decisions for a player to be implemented contingent to all possible states of nature, characteristics of the opponents, their knowledge and intentions, at the same time respecting the possible choice of strategies of the opponents.

TU–NTU Transferable versus nontransferable utility. Concepts attached to cooperative games according to whether a universal medium of transfer ("money") is thought to be available or not.

Type: A state of nature describing the characteristics of a player, randomly generated and in general (privately) observable to this player, thus allowing him to infer conditional probabilities concerning the other players types.

Values: In Game Theoretical context a class of (linear) solution concepts that reflect *a priory* assessments or expectations of gains, mainly axiomatically justified.

Welfarism: An ideological position in welfare economics and social choice theory claiming that social welfare depends in a society only on the utilities of their individual members. Axiomatic bargaining models in the tradition of Nash are welfaristic. In contrast, models with an underlying outcome space representing a specific economic or social context are not.

Summary

Game Theory describes human interaction involving conflict, cooperation and competition, the term Interpersonal Decision Theory is synonymous. The term reflects the fact that most essential features of this field are manifested in parlor games. This topic-level treatment covers large parts of the basic concepts and methods and sketches some fields of recent applications. The simultaneous occurrence of strategic, stochastic and dynamic phenomena, the fundamental role of epistemic aspects like knowledge and information and the impact of institutional and organizational structures make game theoretic analysis a highly complex task.

In order to deal with various facts of social interaction different forms of strategic or cooperative game models have been developed. The Normal (or Strategic) Form describes the strategic alternatives, the Extensive Form reflects the evolvement of games in time as governed by players' successive decisions during play. In particular, Repeated Games with Incomplete Information describe iterated plays of the same randomly influenced game about which the players receive asymmetric information. The Coalitional Form describes power of coalitions.

Equilibria and solutions represent various approaches to solve games or to describe stable, fair, expected or just likely payoffs of games.

In mechanism design an imperfectly informed planner with limited enforcement power creates rules of a game that ensure that any potential population of players by playing an equilibrium according to those rules ends up with a socially desired state.

The Equivalence Principle deals with an important application of game theory to large economies, where due to the dominating power of competition distinct solution concepts asymptotically coincide with the Walrasian equilibria.

Recent applications of game theory to evolutionary biology and evolutionary models of social systems and of learning are also briefly sketched.

Finally, results from game theoretic analysis based on perfectly rational players are contrasted with laboratory experiments that have been performed with real, hence at best boundedly rational, players.

A brief assessment of game theory as a part of Operations Research (or vice versa) concludes.

1 Introduction

Game Theory is a mathematical theory of socio–economic phenomena exhibiting human interaction, i.e., conflict and cooperation between decision making individuals (the players). The theory is based on the structural procedures of mathematics and directed towards problems in various fields of applications.

An appropriate synonym is "Multipersonal Decision Theory". The main paradigms are those of strategic behavior, incomplete information, mutual anticipation of actions, bargaining power, fairness and equity.

Game Theory approaches the problem of decisions for a group of individuals under uncertainty, it deals with lack of information about the state of the environment, the state of the interpersonal decision process and the state of the opponents incentives and abilities. Hence, a probabilistic context is inevitable. The states of nature as well as the strategic behavior of the players involved are generally thought to be randomly influenced.

In addition, the mutual anticipation of opponents strategic behavior, the mutual knowledge about the opponents knowledge and the recursive influence of such kind of consideration on the state of knowledge as well as the resulting strategic consequences are modeled; again they are thought to be randomly influenced. This way an idea of "common knowledge" enters the scene.

Also, Game Theory focuses on aspects of cooperation, enforced by legal contract or by long standing experience. It treats problems of fair distribution of resources, acceptable outcomes to joint operations, the representation of bargaining power and coalitional influence, the *a priori* expectation of gains to be achieved from cooperative decisions. The power of coalitions and the resulting influence of individuals, principles of bargaining and axiomatic treatment of solutions, complaints and threats, efficiency and effectiveness, reputation and learning are being discussed on a formal level.

The performance in strategic or cooperative situations (in "the game") requires an incentive. A version of utility theory is underlying most game theoretical models. This implies that the individuals involved (the players) are capable of expressing preferences with regard to the decisions at stake. Thus, it is required that, for each player, there is a preference ordering or a utility function defined on the set of decisions available to *all players*.

Given a player's incentives, he may have incomplete (and randomly influenced) information about the incentives, preferences, or utilities of his opponents. Indeed, Game Theory is capable of describing situations in which players are uncertain about the game they are playing and the opponents they are facing.

Game Theory is also concerned with clarifying the notion of rational behavior. It does not explicitly so, but the concept appears implicitly formulated in various attempts to find a "solution " of a game. Solutions more ore less imply that the players achieve benefits by acting rationally on the basis that everyone else behaves rationally as well.

Game Theory basically uses the language of Mathematics, it embraces the analysis of structural relations due to mathematical thinking. Models are formulated in precise definitions, theorems are stated and proved. The mathematical techniques vary through a great range, they involve linear algebra and analysis, measure theory, probability and statistics, stochastic processes and potential theory, partial differential equations, functional analysis, combinatorics, graph theory, optimization and more.

The main fields of application can be found in economics. However, sociology, political sciences, psychology, industrial organization, management science, biology, warfare etc. are all open to the formulation and formal treatment via games.

Within these various fields Game Theory is set to, the formal mathematical treatment contains various degrees of rigor; descriptions of games may be purely verbal and strategic behavior may be treated in a less rigorous framework. Model builders have a tendency to more or less incorporate the methods and the language of their respective field. In this context Game Theory changes its appearance. Economists tend to a version that resembles their way of thinking in the tradition or ideology of certain schools, biologist use the language of evolutionary theory etc. In such a context, mathematical rigor is sacrificed against greater adherence to the methods and dogmas of the particular field.

Historically, Game Theory developed along various different lines of thought, most of them rather disjoint. Mathematicians (in particular the french school Laplace, De Moivre, Pascal) considered the probabilistic aspects of the casino. Daniel Bernoulli (1738) (motivated by Jean and Nicolas Bernoulli) considered the St. Petersburg problem; he discussed not only the probabilistic intricacies but also came up with an early version of utility theory. This line was continued by Louis Bachelier (1901), who also created the first version of Brownian motion representing the stock markets fluctuation. Emile Borel contributed greatly to put probability theory into its present shape based on measure and integration. (1925-1939). But he was surpassed by John von Neumann when he unsuccessfully tried to solve the Min–Max Problem (1921).

The early economists Cournot (1838) and Bertrand (1889) discussed oligopoly and developed a notion of strategy. Bertrand also treated the game of *baccarat*. This line was continued by Edgeworth (1881), Zeuthen (1930) and Stackelberg (1952). Edgeworth in addition started a line of discussion leading to the cooperative approach.

At about 1713 (the same time that Jean and Nicolas Bernoulli report to him the St. Petersburg Problem) De Montmort was also in connection with J. (the Earl of) Waldegrave who analized a 2-person card game. Here probabilities occur rudimentary reflecting strategic behavior.

Warfare appears in context with strategic thinking. Clausewitz discusses the battle field coolly from the strategic viewpoint. At the beginning of the 20th century some english engineers developed simple evader-pursuer models which resemble differential games between airplanes.

The two decades between 1920 and 1940 reflect the final attempt to view the Theory of Games as a comprehensive field. Von Neumann's proof was based on fixpoint theorems which in the mid-thirties where particularly developed by Banach, Mazur, Ulam, Erdös, Steinhaus, Kuratowski . Ville was the first to provide a proof based on a separation theorem.

Oskar Morgenstern met von Neumann when both men had to leave Europe in the late thirties. They laid the the foundations of the field of Game Theory with their seminal volume *The Theory of Games and Economic Behavior*, where they stressed the similarity between strategic and cooperative behavior in the economic context as well as in parlor games. Random influence was considered to be inevitable.

Due to these authors three versions of "the game" emerge. Games appear in ***normal form*** (strategic form), in ***extensive form*** and in ***coalitional form***. The first two are close relatives, they constitute the basic paradigm of ***Noncooperative Game Theory*** . The coalitional form is the basic paradigm of ***Cooperative Game Theory*** .

The *normal form*, consists of a complete list of possible strategic alternatives for each player. This way each player is assigned a ***strategy space*** . In addition, a ***payoff function*** is specified for each player. Thus, any simultaneous and independent choice of strategies (one by each player) results in a payoff (a real number, a utility, a money term) to each player .

The normal form is also referred to as the ***strategic form***, in view of the fact that it provides an overview over the strategic options available to a player.

It is one of the main tasks of the model builder to recognize "The Game". Given the data of a multipersonal decision problem of a possibly foggy and unclear nature, one has to specify a normal form game which contains the essential features and is "close" to reality.

For the normal form the basic "solution concept" is ***equilibrium***. This is a strategic situation (an $n-$tuple of strategies) with dominant stability properties. An equilibrium may reflect versions of "rational behavior" and in some case may be identified with "optimal strategies". In most games, however (as in the real world), there is no "optimal behavior", equilibria may (or may not) exist in abundance and result in gains of greatly varying utility to the players.

The *extensive form* was originally conceived to explain the "rules of the game" (von Neumann–Morgenstern). Preferably one might think of a time-structured (and stochastically influenced) process that is subject to repeated actions of the players. Intermediate and final payoffs (or costs) are awarded to the players. Decisions at an early state should, therefore, be regarded with respect to the present reward *and* with respect to the future consequences. The process as well as its history may not be fully observable. Players receive private information concerning the state of the process and the choice of actions of the opponents. Strategic behavior is to be defined according to observations and the development of the process. This way, the extensive form results in a normal form and is subject to the analysis thereof. Then equilibrium can be recognized. The extensive form may provide the basic environment, time structure and "rules of the game" but the normal form provides the solution concept.

Turning to the *cooperative* or *coalitional form*, we find that the notion of strategy is no more predominant. Rather it is the possibility of *contracts* and *cooperation* which is preeminent. Binding agreements are thought to be possible and enforceable. Thus, the power of coalitions and their influence on the results of a bargaining process is the central topic. A cooperative game is essentially a mapping assigning achievable utilities to coalitions. The task here is to make inference from the power of coalitions to the potential of the individuals. If we know the game, what will be the resulting possibilities, options, expectations, gains to the players?

The "solution concept" of Cooperative Game Theory is the idea of ***Stable Solution***. While adherent to some idea of equilibrium, the cooperative version of stability is much more static. Stability of the results of bargaining and cooperation, fairness and equity, the returns expected from cooperation, the consequences of an argumentative process, a final distribution of utility to the players achieved by agreement – these ideas are central to the coalitional form.

Thus, the balance of noncooperative versus cooperative theory is made precise by discussing strategic behavior and equilibrium strategies versus the power of coalitions and stable solutions.

In the detailed discussion, however, it turns out that the borderline is blurred. There is noncooperative imitation of cooperation: the stabilizing forces of reputation and punishment that appear in repeated games tend to exhibit elements of cooperation; the agency enforcing contracts can be replaced by the pressure of mutual punishment sustaining equilibrium. On the other hand, cooperative theory incorporates elements of strategic behavior. If uncertainty prevails about the opponents motivation, their preferences and the game one (thinks one) is involved, then ***mechanisms*** enter the scene. These are devices representing agreements dependent on private observations or knowledge of the players. As these observations cannot be verified independently, players may start to behave strategically with respect to the revelation of their information or their strategies. This sets the individuals involved in a noncooperative game after the contract has been agreed upon.

Some game theorists hold that cooperative theory is not an "independent" topic; in a sense all cooperation should be explained as resulting from strategic behavior. This view may be extended to a position opposed to cooperative theory at all. Another view, however, is that the idea of the "game" is something Platonic: the paradigm of human competitive and cooperative interaction in the presence of incentives and mutual dependence. Various shapes of this idea materialize, some of them in a precise and mathematically rigorous form.

2 Foundations of Noncooperative Game Theory

2.1 The Normal Form

The following formal definitions are meant to explain the basic and fundamental topics of Noncooperative Game Theory.

A ***noncooperative*** n***–person game*** in ***normal form*** is a $2n$-tuple

$$\Gamma \ = \ (\mathcal{S}^1, \ldots, \mathcal{S}^n; F^1, \ldots, F^n). \tag{1}$$

with the following ingredients. $\mathcal{S}^i$ $(i = 1, \ldots n)$ denotes the set of ***strategies*** of player i. This a complete list of decisions available to the player; at this stage the details of strategic behavior cannot be distinguished. Each F^i is a real valued function defined on the Cartesian product $\mathcal{S} := \mathcal{S}^1 \times \ldots \times \mathcal{S}^n$ of the strategy spaces. F^i denotes the ***payoff*** to player i, depending on the strategies chosen by *all* players. The choice of a strategy n-tuple is made simultaneously and independently. When preparing his choice each player is not aware of the opponents intentions. However, communication may take place in advance; a discussion of the merits and demerits of strategy n–tuples may well precede the actual choice of strategies.

A ***Nash equilibrium*** is an n-tuple $\bar{s} \in \mathcal{S}$ such that deviating is not profitable for a player provided his opponents stick to their choice. Formally:

$$F^i(\bar{s}) \geq F^i(\bar{s}^1, \ldots, s^i, \ldots, \bar{s}^n), \quad (s^i \in \mathcal{S}^i; \ i = 1, \ldots, n). \tag{2}$$

A priori nothing is said about the establishment of an equilibrium; however, the inherent stability of an equilibrium situation may prevent a player from leaving it. The existence of equilibria requires a basic set of mathematical assumptions, generally the strategy spaces should be (contained in) topological vector spaces and the payoff functions should be quasi–concave and continuous. The standard procedure is to construct the *best reply correspondence* which is a mapping assigning to each $n-1$-tuple of a players opponents the set of maximizers of his payoff. A fixed point theorem (KAKUTANI, KY FAN) provides a Nash equilibrium. The one first to establish the concept was JOHN F. NASH.

If these conditions fail to apply (e.g.,if the strategy sets are finite), then the game may be ***extended*** in various ways. The ***mixed extension*** randomizes the strategic choice of strategies. Assume that, on each strategy set $\mathcal{S}^i$, there is defined a σ-algebra $\underline{\underline{\mathbf{P}}}^i$ of measurable sets. The probabilities on $\underline{\underline{\mathbf{P}}}^i$ are called ***mixed strategies***. This way player i now chooses a random mechanism which generates his original "pure strategies" (the elements of $\mathcal{S}^i$).

Given an n-tuple of mixed strategies, the product measure, say $\sigma = \sigma^1 \otimes \ldots \otimes \sigma^n$ reflects the (stochastically) independent choice of strategies. The expectation $\bar{F}^i(\sigma) = \int F^i(s)\sigma(ds)$ is used to reflect the payoff to player i at this n-tuple of mixed strategies. Now, if $\mathcal{M}^i$ denotes the set of mixed strategies of player i, then we have defined a noncooperative n-person game in the sense of (1); this is

$$\bar{\Gamma} = (\mathcal{M}^1, \ldots, \mathcal{M}^n; \bar{F}^1, \ldots, \bar{F}^n), \tag{3}$$

the *mixed extension* of Γ.

With a suitable structure on the strategy spaces, there is a topology on the mixed strategy spaces (the *w^* -topology*) such that the functions $\bar{F}^i$ are continuous and (multi)-linear with respect to the mixed strategies. This way the above existence theorems can be employed to establish equilibrium in mixed strategies.

Nash equilibrium in mixed strategies can be reinterpreted as follows: for any player i, the $n-1$-tuple of mixed strategies of his opponents are regarded as his ***believes*** concerning the behavior of his opponents. A Nash equilibrium constitutes ***consistent beliefs*** of the players concerning their randomized choice of strategies.

The ***correlated extension*** is obtained by introducing a random experiment (a probability space) resulting in private information of the players (i.e., there are subfields of observable events for the players). A ***correlating strategy*** for a player is a random variable, measurable with respect to his observable events and resulting in strategies. If each player chooses such a correlating strategy, the expected payoff for all player (from the composition of the correlating strategies and the payoff functions of the original game) is well defined, hence we have a new normal form, the *correlated extension*. A ***correlated equilibrium*** is a Nash equilibrium of the correlated extension. Actually, the mixed extension can be embedded into the correlated one and, for many purposes, suffices to treat the relevant strategic aspects.

There is a host of applications of this model. It is used in oligopolistic competition and other descriptions of price setting mechanisms, statistics, market entry problems, evolutionary biology, for auctions, principle agent problems, inspection problems, insurance contracts, job assignment problems, traffic regulation, etc. In many cases the application of mixed strategies proves to be most successful.

2.2 The Extensive Form

The ***extensive form*** is a dynamic process admitting of control by n players. Stochastic influence is inherent in the system as well as in the strategic behavior of the players. Tree games in which the process moves along the edges of a graph are the favorite model in the literature. At each node a player chooses the subsequent edge. There may be the possibility of imperfect or incomplete information: While the process moves the players may not be aware of the state of the process and they may even not be aware of the utilities and strategic possibilities of their opponents, for short, their opponents' "types" (HARSANYI). We demonstrate the extensive form by a general Markovian dynamic game written

$$\Sigma \ := \ (\mathbf{X}, \mathbf{Y}; \mathbf{K}, \mathcal{K}; \boldsymbol{Q}, f, u, T) \tag{4}$$

Here, T is the horizon or duration of the process. $\mathbf{X}$ and $\mathbf{Y}$ are the state and action spaces respectively Each of them is time–structured, i.e. $\mathbf{X} = \mathbf{X}_0 \times \ldots \times \mathbf{X}_T$ etc. $\boldsymbol{Q}$ is a family of stochastic transition kernels governing the law of motion. $\mathcal{K}$ again is a family of stochastic kernels generating signals (in $\mathbf{K}$) which can be observed by the players while the process moves. f is a family of intermediate payoffs and u is a family of terminal payoffs. Assume that there is a path (x, y) of temporal development in the state space $\mathbf{X} \times \mathbf{Y}$. Then the ***evaluation*** for player i is written

$$C^i(x, y) \ = \ u^i(x_T) + \sum_{t=1}^{T} f_t^i(x_{t-1}, y_t) \tag{5}$$

Now, if (X, Y) is a stochastic process moving in the state space then $C^i \circ (X, Y)$ is a random variable the expectation of which evaluates the process. Now, the distribution of the process is governed by the strategic behavior of the players as follows.

Behavioral strategies are families of kernels, say $\boldsymbol{A}_t^i$, which reflect the random choice of an action by player i at each instant t, depending on the observable past (i.e. the stream of observable data of the process). The composition of the behavioral strategies $\boldsymbol{A}$ and the law of transition reflected by $\boldsymbol{Q}$ generates transition kernels (Markovian or with a memory) on the state space. Given an initial distribution μ on the initial states $\mathbf{X}_0$ there is, therefore, a measure (Markovian or reflecting memory) on the paths of the state space $\mathbf{X} \times \mathbf{Y}$, call it $\boldsymbol{m}_{\mu}^{\boldsymbol{A}}$. This distribution reflects the stochastic influence of behavioral strategies on the motion of the controlled process. Therefore, we consider the payoff to player i resulting from $\boldsymbol{A}$ to be

$$C_{\boldsymbol{A}}^{i\mu} = \mathbb{E}C^i \circ (X,Y) = \mathbb{E}_{\boldsymbol{m}_{\mu}^{\boldsymbol{A}}} C^i = \int C^i(x,y) d\boldsymbol{m}_{\mu}^{\boldsymbol{A}}(x,y). \tag{6}$$

Now, $C_{\bullet}^{i\mu}$ is a function on the product space of behavioral strategies, say $\mathcal{A} = \mathcal{A}^1 \times \ldots \times \mathcal{A}^n$. This way we have constructed the normal form generated by Σ (and μ) which is

$$\Gamma_{\Sigma,\boldsymbol{\mu}} = (\mathcal{A}^1, \ldots, \mathcal{A}^n; C_{\bullet}^{i\boldsymbol{\mu}}, \ldots, C_{\bullet}^{n\boldsymbol{\mu}}). \tag{7}$$

Now the whole apparatus of Nash equilibrium analysis may be employed.

2.3 Strategies, Equilibria, Refinements

Strategic behavior of players can also be modeled by ***pure strategies*** (the choice at each stage is deterministic) or ***mixed strategies*** (probabilities or "mixtures" over the pure ones). Then different normal forms occur. Behavioral strategies are appropriate in a wide range of games. This is due to KUHN'S THEOREM which states that (with ***perfect recall***, i.e., consistent memory) behavioral strategies generate the same distributions as mixed strategies.

Perfect recall may be violated if the formal structure of memory represented for a player is in some sense inconsistent. An inconsistent memory structure is the topic of a new branch of information based Game Theory (the *absent minded driver* (RUBINSTEIN)): What can one say about strategic behavior if players (or automata) forget systematically essential details of the past? Or if they are in a wider sense non–rational (governed by ***bounded rationality***)?

The temporal structure of stochastic games or tree games permits refinements of the equilibrium concept. Nash equilibria are ***subgame perfect*** (SELTEN), if the equilibrium property prevails in every truncated tail game (in the finite context they are obtained by backwards induction). However, tail games or subgames are well defined only with complete or perfect information. With imperfect information the construction of *a posteriori* probabilities for the state of the process may be conditioned on informations which, at equilibrium, have zero probability. This is solved by various versions of perturbing the game so as to enable Bayesian posteriors to be computed. Sequential equilibrium (KREPS–WILSON) therefore consists of pairs of strategies and "beliefs". Perfect equilibria (the first and basic concept due to SELTEN) is a close relative. Other versions (e.g., stable equilibria due to KOHLBERG–MERTENS) enhance the scene. There is a peculiar connection between this kind of stability concept, equilibrium selection, and the shape of the equilibrium correspondence manifold of the normal form generated by Σ.

The information structure of a game has intensively influenced research. HARSANYI pointed out that *incomplete information* (uncertainty about the type of other players) and *imperfect information* (uncertainty about the state of the process) are more or less the same. MERTENS–ZAMIR constructed the appropriate belief spaces such that an infinite number of hierarchies about knowledge of each others knowledge *etc.* can be formally constructed.

AUMANN established the idea of ***common knowledge*** which formalizes mutual knowledge about mutual knowledge again with infinite hierarchies. It is important that when a game is being played all players are informed about the game, all players are informed about the fact that all players are informed about the games, all players... etc. As this topic extends into epistemological questions (partially of philosophical nature) aspects of formal logic become more and more important for Game Theory. The construction of belief spaces, hierarchy of beliefs, learning and knowledge about learning are fascinating topics of the theory.

The mathematical intricacies of stochastic games in the most general sense are also enhanced by the information problem. Stochastic games in the proper sense were originated by SHAPLEY (complete information). The question of the existence and shape of equilibria in the general stochastic game with incomplete information is still an intriguing matter. Repeated games with incomplete information constitutes the topic treated most extensively. The details can be found in SECTION 6.

3 NTU-Games

3.1 The Coalitional Function

Within the framework of Cooperative Game Theory the notion of strategy loses importance. As the players are capable of binding agreements, they may commit themselves to certain actions which result in joint utilities for coalitions. These actions may well be strategies in an underlying noncooperative game (an idea basically favored by VON NEUMANN–MORGENSTERN). But as the agreement or contract can be enforced by some supervising agency, there is less room for strategic behavior.

A contract between members of a coalition includes a specific agreement concerning the distribution of utility which results from the joint venture. Usually there will be a great deal of alternatives a coalition can strive for. The model consists of comprehensive lists of utility vectors available to each feasible coalition. While bargaining the players will not only look to the possible achievements of a coalition which is presently being discussed. They will also look to the payoff vectors of other coalitions and they will argue with their outside options defined thereby. Therefore, the agreement finally reached, whether inside the grand coalition or some smaller subcoalition, eventually reflects all options and possibilities available to the various coalitions evaluated with respect to the players.

The formal description is provided as follows. A ***Cooperative Game Without Side Payments*** or for short an ***NTU-Game*** is a triple $(I, \underline{\underline{\mathbf{P}}}, \boldsymbol{V})$. Here, I is the set of ***players*** (frequently assumed to be finite, e.g., $I = \{1, \dots, n\}$). $\underline{\underline{\mathbf{P}}}$ is a system of subsets of I which is interpreted as the collection of ***feasible coalitions***. Finally $\boldsymbol{V} : \underline{\underline{\mathbf{P}}} \to \mathcal{P}(\mathbb{R}^n)$ is the ***coalitional function***. This function assigns to every coalition a set of utility vectors. Certain regularity assumptions are imposed upon the function $\boldsymbol{V}$ in order to render it feasible for a "game". For instance, as coalition S usually can assign utilities only to its members, it makes sense to assume $\boldsymbol{V}(S) \subseteq \mathbb{R}^n_S$. Also it is assumed that $\boldsymbol{V}(S)$ is *comprehensive*, that is utility can be freely disposed of (formally: every vector dominated by an element of $\boldsymbol{V}(S)$ belongs to $\boldsymbol{V}(S)$). In addition some version of *boundedness* from above ensures that utility is not unlimited available. *Convexity* assumptions also are quite common.

The economic context provides various examples. E.g. if

$$\mathcal{E} = (I, \mathcal{X}, (u^i)_{i \in I}, (a^i)_{i \in I}), \tag{1}$$

is an ***exchange economy*** represented by a set of agents, a commodity space, and a family of utility functions on commodities for each player, then we can construct the corresponding NTU-Game. For each coalition we collect all the utilities available by mutual exchange. Formally the ***market game*** $\boldsymbol{V} = \boldsymbol{V}^{\mathcal{E}}$ is given by

$$\boldsymbol{V}(S) = \left\{ \left(u^i(x^i)\right)_{i \in S} \;\middle|\; x^i \in \mathcal{X}\ (i \in I),\ \sum_{i \in S} x^i = \sum_{i \in S} a^i \right\} \quad (S \in \underline{\underline{\mathbf{P}}}) \tag{2}$$

or, more technically, the comprehensive hull of this set (i.e., we admit free disposal of utility). Thus, coalition S can attain all utilities for its members that can be obtained by reallocating the commodities within this coalition. Market games obviously establish a close connection between Game Theory and General Equilibrium Theory (see SECTION 4.)

Within the context of NTU-Games, the class of *bargaining problems* is obtained by admitting only the grand coalition I and the singleton coalitions $\{i\}, (i \in I)$. Thus, players can either join in the grand coalition or be on their own. Observe that it suffices to specify two data in order to define a bargaining problem: the set $\boldsymbol{U} := \boldsymbol{V}(I)$ of utilities available to the grand coalition and the maximal utility $\underline{u}_i$ that can be achieved by player $i \in I$. A ***bargaining problem*** is, therefore, defined by a pair $(\underline{u}, \boldsymbol{U})$, the ***status quo point*** and the ***feasible set***.

The most important class of games is generated by admitting ***side payments***. Imagine that, within each coalition, the players are entitled and capable of exchanging utility on a universal scale so that a unit of utility can be transferred from one player to another one without changing the nature of its value. More precisely, whenever $\boldsymbol{x} \in \boldsymbol{V}(S)$ holds true, i is a player in S and ε is a small quantity of utility, then we assume that the vector $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i - \varepsilon, \ldots, \boldsymbol{x}_j + \varepsilon, \ldots, \boldsymbol{x}_n)$ is an element of $\boldsymbol{V}(S)$ as well. It turns out immediately that every $\boldsymbol{V}(S)$ is necessarily of the form

$$\boldsymbol{V}(S) \;=\; \left\{ \boldsymbol{x} \;\middle|\; \sum_{i=1}^{n} \boldsymbol{x}_i \leq \boldsymbol{v}(S). \right\} \tag{3}$$

Here, $\boldsymbol{v}(S)$ is a real number, the utility assigned to coalition S. Obviously the function $\boldsymbol{V}$ is specified once the function $\boldsymbol{v} : \underline{\underline{\mathbf{P}}} \to \mathbb{R}$ is defined. Therefore, we call the triple $(I, \underline{\underline{\mathbf{P}}}, \boldsymbol{v})$ a cooperative game with side payments, with transferable utility, or for short a ***TU-Game***.

3.2 Solutions

We now turn to ***solution concepts***. A solution in the general sense describes outcomes of the bargaining process. This may involve varying vantage points. A solution may represent an evaluation of the bargaining power of players deduced from the game, it may respect fairness considerations or principles of equity, expected gains in some (vaguely defined) stochastic environment, or results of a specified procedure involving arguments, counter arguments, objections and counter objections. Solutions may also be defined as the result of a noncooperative game which is based on the data of the cooperative game and represents a bargaining process. A Nash equilibrium of such a game may result in a solution of the cooperative game. The interpretation of this noncooperative Nash equilibrium may furnish a justification of the cooperative bargaining solution resulting. All this can be formalized within the proper context.

The number of solution concepts is considerable, their esteem is greatly at variance among game theorists. There is, however, agreement that solutions have to prove their merits by the results they yield on a sufficiently rich class of games. Formally, a solution concept is a mapping, point valued or set valued, defined on some class $\mathcal{V}$ of NTU-Games. If the set of players I is fixed, then

$$\varphi \ : \ \mathcal{V} \to \mathbb{R}^I \quad or \quad \phi \ : \ \mathcal{V} \to \mathcal{P}(\mathbb{R}^I) \tag{4}$$

defines a ***solution***. That is, for every game there is an assignment (or a set of assignments) of utility to each player "resulting" from the game.

Solution concepts should exhibit certain appealing properties expressed by conditions or *axioms*. Ideally, they are uniquely defined by an appropriate set of axioms, this is the ***axiomatic approach***. Procedural approaches, definitions by extension of "natural" or "canonical" concepts, solutions based on the economic tradition or the more mathematical approach via invariance properties are also common.

Let us shortly mention solution concepts for *bargaining problems* which are suitably called ***bargaining solution***. A bargaining solution obeys standard axioms: as a mapping on (a subclass of) bargaining problems it commutes with permutations of the players (i.e. the names of the players are irrelevant). It commutes with positive affine transformation of $\mathbb{R}^I$, i.e., follows a transformation of the scale. Frequently it is *Pareto efficient*. That is, at the evaluation of a fixed game, eventually no player can strictly improve his outcome unless another player suffers. Finally, one requires *individual* *rationality*: no player accepts less than he can achieve by his own efforts.

In general it needs a specifically defined further axiom in order to generate a uniquely defined bargaining solution. The historically first and basic approach is provided by the ***Nash solution***. The decisive axiom is called "Independence of Irrelevant Alternatives". It determines an outcome on the Pareto surface of the feasible set maximizing the coordinate product (relative to the status quo point's coordinates). Further solutions are the ***Kalai–Smorodinsky solution*** (uniquely characterized by a weak monotonicity axiom) and the ***Maschler–Perles solution*** (defined by superadditivity).

Meta Bargaining Theory and Bargaining with Incomplete Information are modern extensions of the traditional Bargaining theory.

The first one strives to deal with the problem of choosing between various concepts of bargaining theory on the basis of axiomatic or procedural treatment. In the context of incomplete information, players are not fully informed about certain characteristics of their opponents. These characteristics may describe the preferences or endowments of other players. There may be a common prior concerning the distribution of such characteristics and a single player gets some private information (reflected by a chance move at the beginning of the bargaining process or the like) concerning his own data. For short, each player knows his own "type" and has a priori probabilities about the opponents' types. Agreements may be registered with respect to certain types of ***mechanisms***, i.e., mappings generating decisions in dependence of all players' announcements of their types. Now, the way to announce (and possibly misrepresent) one's type is dictated by strategic behavior. Hence, there appears a noncooperative n–person game, the equilibria of which may correspond to incentive compatible mechanisms. The study of such mechanisms, axiomatic treatment etc. is the relevant topic within this field (see also SECTION 5).

Solution concepts for general NTU–games are mainly discussed after the fashion of the TU–concepts. We therefore postpone the discussion of this topic.

4 TU-Games

4.1 Classification of Games

A ***cooperative game with side payments*** or, for short a ***TU-game*** is formally represented by a triple $(I, \underline{\underline{\mathbf{P}}}, \boldsymbol{v})$. Here, I is ***the set of players***, $\underline{\underline{\mathbf{P}}} \subseteq \{S \mid S \in I\}$ is the collection (usually a field) of ***feasible coalitions***, and $\boldsymbol{v} : \underline{\underline{\mathbf{P}}} \to \mathbb{R}$, $\boldsymbol{v}(\emptyset) = 0$, is the "characteristic" or ***coalitional function*** (frequently referred to as "the game" as well). For most of our discussion we assume the player set to be finite, we use $I = \{1, \ldots, n\}$. Also, $\underline{\underline{\mathbf{P}}}$ can be viewed to be the power set of N, i.e., all coalitions are admitted for cooperation. Intuitively, a coalition $S \in \underline{\underline{\mathbf{P}}}$ by agreeing to a contract about cooperation can achieve (or is awarded) a worth or value $\boldsymbol{v}(S)$. This worth is a monetarian or utilitarian quantity, all players see it on a universally accepted scale and arbitrary quantities of this medium can be transferred by some (unidentified) mechanism between the players. In view of these simplifying assumptions, the theory of TU-games is extensively developed.

In addition, TU–games have been applied in the contest of ***cost sharing***, which is rather a topic of industrial organization. Here, coalitions my be subsidiaries or divisions of a company or, more generally, groups of cost generating factors. The coalitional function describes the cost generating structure. Thus it assigns "disutilities" (costs, expenses) instead of utilities.

Concrete interpretations of the nature of a TU-game may depend on the context but also on the mathematical nature of the set function $\boldsymbol{v}$. For the sake of this discussion we assume that $\boldsymbol{v}$ is nonnegative.

An ***additive*** set function $\boldsymbol{m}$ on $\underline{\underline{\mathbf{P}}}$ is, for finite player set I tantamount to vector $m \in \mathbb{R}^n$ via

$$\boldsymbol{m}(S) := \sum_{i \in S} m_i \quad (S \in \underline{\underline{\mathbf{P}}})$$

Within the framework of Cooperative Game Theory additive set functions (or ***measures***) are meant to represent distributions of utility. As a game, they are "trivial" or "inessential" as cooperation does not improve a coalitions worth: for *disjoint* coalitions $S, T \in \underline{\underline{\mathbf{P}}}$ we have $\boldsymbol{m}(S) + \boldsymbol{m}(T) = \boldsymbol{m}(S \cup T)$.

The situation changes when we consider ***superadditive games*** characterized by the defining inequalities $\boldsymbol{m}(S) + \boldsymbol{m}(T) \leq \boldsymbol{v}(S \cup T)(S, T \in \underline{\underline{\mathbf{P}}}$, disjoint $)$. In such games the formation of coalitions is worthwhile as the total gains increase, thus players can expect to achieve a larger share of utility by cooperation.

A subclass of superadditive games is provided by the class of ***balanced games***. Call a system of coalitions $\underline{\underline{\mathbf{S}}} \subseteq \underline{\underline{\mathbf{P}}}$ ***balanced*** if there is a set of positive coefficients $(c_S)_{S\in\underline{\underline{\mathbf{S}}}}$ such that

$$\sum_{S\in\underline{\underline{\mathbf{S}}}} c_S 1_S = 1_I \tag{1}$$

holds true. We interpret the coefficients as an "intensity" to operate a coalition with. Hence, a balanced system is a collection in which the players are running the various coalitions with reduced intensity instead of joining in the grand coalition.

A game $\boldsymbol{v}$ is called **balanced** (SHAPLEY–BONDAREVA) if, for any balanced system $\underline{\underline{\mathbf{S}}} \subseteq \underline{\underline{\mathbf{P}}}$ and corresponding coefficients $(c_S)_{S\in\underline{\underline{\mathbf{S}}}}$ it follows that

$$\sum_{S\in\underline{\underline{\mathbf{S}}}} c_S \boldsymbol{v}(S) \leq \boldsymbol{v}(I) \tag{2}$$

holds true. Verbally: it pays off to join within the grand coalition as there is no better way to achieve the same utility by splitting into any balanced system and running the respective coalitions with a reduced intensity.

Next, a game is ***totally balanced*** if all restrictions $\boldsymbol{v}_{|S}$ to coalitions $S \in \underline{\underline{\mathbf{P}}}$ are balanced. The totally balanced games are formally equivalent with ***market games*** DEBREU, VIND, AUMANN, SHAPLEY–SHUBIK. Thus, they can be constructed as originating from a side payment or TU exchange economy or market. In such a market a coalition is permitted to transfer utility by reallocating its goods freely in order to achieve maximal joint utility. A coalitional function appears which is indeed totally balanced. Formally, if $\mathcal{E}$ is defined as in (1), then analogously to (2) the coalitional function $\boldsymbol{v} = \boldsymbol{v}^{\mathcal{E}}$ is given by

$$\boldsymbol{v}(S) = \max\left\{ \sum_{i\in S} u^i(x^i) \,\middle|\, x^i \in \mathcal{X}\ (i \in I),\ \sum_{i\in S} x^i = \sum_{i\in S} a^i \right\} \quad (S \in \underline{\underline{\mathbf{P}}})\ . \tag{3}$$

Furthermore, totally balanced games appear as ***LP–games*** (OWEN). If a (positive) linear programming setup $\mathcal{L} = (A, \boldsymbol{b}, c)$ is specified by an input–output matrix A, a vector valued measure of resources $\boldsymbol{b}$ and an objective function given via a vector c, the resulting LP–game is represented by a function $\boldsymbol{v} = \boldsymbol{v}^{\mathcal{E}}$ via

$$\boldsymbol{v}(S) = \boldsymbol{v}^{(A,\boldsymbol{b},c)}(S) = \max\left\{ cx \mid x \in \mathbb{R}^m_+, Ax \leq \boldsymbol{b}(S) \right\}, \quad (S \in \underline{\underline{\mathbf{P}}}). \tag{4}$$

That is, a coalition may pool its resources and use the production process represented by the matrix A optimally in order to obtain the worth of joint production.

A further representation of totally balanced games can be obtained by seeing them as ***MIN–games***. These are the minima of finitely many ($\sigma-$) additive functions, say $\boldsymbol{\lambda}^1, \ldots, \boldsymbol{\lambda}^r$ via

$$\boldsymbol{v}(S) \;=\; \bigwedge \left\{\boldsymbol{\lambda}^1, \ldots, \boldsymbol{\lambda}^r\right\}(S) \quad (S \in \underline{\underline{\mathbf{P}}}). \tag{5}$$

This way, totally balanced games appear to be ***glove games*** (SHAPLEY): a minimum of each resource (left hand gloves / right hand gloves) determines the amount of utility (pairs of gloves) a coalition can achieve. Remarkably, all these classes are technically equivalent. There are actually more representations generating totally balanced games, e.g., as games on graphs (*max flow – min cut* setups, KALAI–ZEMEL).

A nice (proper) subclass of totally balanced games is provided by ***convex games*** (SHAPLEY). Within this setup one discusses set functions with *increasing marginal worth*: the marginal contribution of a player i to a coalition S, say $\boldsymbol{v}(S \cup \{i\} - \boldsymbol{v}(S)$, increases with increasing coalitions. In such games the incentive to form large (the grand) coalitions is particularly compelling. In the cost sharing context (if we reverse the sign or consider the difference of an additive and a convex set function) we obtain *concave* cost structures which nicely concur with decreasing returns to scale.

A further subclass of cooperative games deserves a separate discussion because it is important in the political rather than in the economical context. The class of ***simple games*** consists of functions $\boldsymbol{v} : \underline{\underline{\mathbf{P}}} \to \{0, 1\}$. The interpretation is not so much that coalitions may win a unit or not. Rather the idea is that a coalition S with $\boldsymbol{v}(S) = 1$ is *winning* and all others are *losing* coalitions. Simple games are used to describe group decisions in political bodies, parlaments, committees. Special simple games are ***directed games*** characterized by a canonical procedure to impose an ordering on the bargaining power of the players. A player is stronger than another player if his marginal contribution to every coalition exceeds the one of his opponent.

Simple games that admit of a representation by voting strength, i.e. ***voting games***, are of particular interest. Such games are given my an additive set function $\boldsymbol{m} \geq 0$ representing the *distribution of votes* over the players (parties in a parliament) and a *majority level* α. A coalition is winning if its combined voting power exceeds α. The coalitional function $\boldsymbol{v} \;=\; \boldsymbol{v}^{\alpha}$ is thus given by

$$\boldsymbol{v}(S) \;=\; \begin{cases} 1 & \text{if} \quad \boldsymbol{m}(S) \geq \alpha \\ 0 & \text{if} \quad \boldsymbol{m}(S) < \alpha \end{cases} \qquad \left(S \in \underline{\underline{\mathbf{P}}}\right) \;. \tag{6}$$

It is easily seen that voting games are directed: a representation naturally induces an ordering of the voting power. (consistently, as players may be equivalent).

The representation of a voting game is by no means unique: the same coalitional function obviously may result from a host of pairs $(\boldsymbol{m}, \alpha)$. A traditional subclass which admits essentially of a *unique* representation is given by the ***homogeneous games*** (VON NEUMANN–MORGENSTERN). Call a representation $(\boldsymbol{m}, \alpha)$ *homogeneous* if every minimal winning coalition has the same voting power. Then the set of players disintegrates into three ***characters***: ***dummies***, ***sums***, and ***steps***. A dummy has no marginal contribution to offer to any coalition. A sum is a player who, in a minimal winning coalition, can be replaced by smaller players (hence, his weight is the sum of the weights of smaller players). Everyone else is a step. Now the unique ("minimal") representation (OSTMANN) is essentially obtained by assigning 0 to the dummies and 1 to the smallest non dummy (who is a step). Then, recursively, sums are awarded a canonical sum of the voting power of smaller players. With steps, this canonical sum is exceeded by 1.

Within the political context, the computation of a coalitional function resulting from the ballot taken from a (huge) population is called *assignment* or ***apportionment***. Essentially, the votes assigned to the parties by an election define a coalitional function, the procedure by which the distribution of votes in the parliament is defined (based on the election results) assigns another coalitional function to the parties. In the political context the procedures named *d'Hondt*, *Hare*, *Imperiali*, *Danish*, and others perform such a task. Frequently, the coalitional function is not preserved, a problem sometimes resulting in "paradoxical" assignment of votes. (The "Alabama paradox"). If the population game happens to be homogeneous, then the computation of the minimal representation can also be seen as apportionment – and of course this always preserves the coalitional function by definition.

4.2 Solutions

Having described various classes of games and a possible environment in which to apply such classes, we now turn to ***solution concepts***. The idea to "solve" a game in the context of Cooperative TU–Game Theory is tantamount to assign additive set functions to games. An additive set function yields an award, utility or monetarian value to each player. Thus, if $\mathcal{W}$ is a class of games and $\mathcal{A}$ denotes additive functions, then

$$\varphi \ : \ \mathcal{W} \to \mathcal{A} \ \ or \ \ \phi \ : \ \mathcal{W} \to \mathcal{P}(\mathcal{A}) \tag{7}$$

denotes a solution concept. This it not at variance with the definition provided by (4) as (for a finite set of players) additive set functions are essentially vectors of $\mathbb{R}^I$.

Again, point valued and set valued versions are being considered. The most universal set valued solution concept is the ***core***. . Given a coalitional function $\boldsymbol{v}$, the core of $\boldsymbol{v}$ is the set $\mathcal{C}(\boldsymbol{v})$ of all additive functions dominating $\boldsymbol{v}$ with the same total mass. Intuitively: all distributions of utility of the grand coalition (i.e. Pareto efficient distributions) which cannot be objected against by smaller coalitions on the grounds that they could do better by cooperating within their own ranks. Formally we have

$$\mathcal{C}(\boldsymbol{v}) \ := \ \left\{\boldsymbol{a} \in \mathcal{A} \,\middle|\, \boldsymbol{a}(S) \geq \boldsymbol{v}(S) \ \ (S \in \underline{\underline{\mathbf{P}}}), \ \boldsymbol{a}(I) = \boldsymbol{v}(I)\right\}. \tag{8}$$

The core is nonempty if and only if the game is balanced; hence a further characterization of balanced games occurs. As we have seen, a balanced game implies a certain pressure towards organization in the grand coalition and the existence of a nonempty core tells us that "organization" may be seen as to award players in an fashion not to be contested by smaller coalitions: the core is a concept of stability. Totally balanced games thus admit of a nonempty core for the subgame of every coalition. The various manifestations of totally balanced games may be reflected by appropriate properties of the core: in a TU market game the payoff obtained in a Walrasian Equilibrium (of the exchange economy used for representation) is a core element. In an LP–game, the shadow prices of the grand coalition applied to the vector valued measure of initial assignments yields a core element (a close relative of the Walrasian equilibrium). In a MIN–game the convex combination of those measures featuring minimal total mass (of the grand coalition) yields a core element (again, this version is related to the Walrasian equilibrium).

Convex games have a particularly nice core: the extreme points (geometrically speaking) are being obtained by "bandwagon processes": coalitions form successively by one player joining and receiving his marginal contribution until the grand coalition is achieved (SHAPLEY).

Mass phenomena are studied in the framework of *large games*. By these slogan we mean that either limiting theorems are stated with respect to increasing (markets–, LP–) games obtained by repeating the characteristics of the players. Or else, one considers games on a continuum of players (I is a measurable space, $\underline{\underline{\mathbf{P}}}$ is a $\sigma-$algebra of measurable sets, and $\boldsymbol{v}$ is a function defined on $\underline{\underline{\mathbf{P}}}$ (usually non atomic)). Equivalence theorems state the coincidence of solution concepts for "large games". E.g., to say that the core is "equivalent" the Walrasian equilibrium for large games means that the core shrinks toward the equilibrium payoff (the shadow price distribution) when the market (and the resulting game) is properly replicated. Or, in the non atomic context, the statement means that the equilibrium payoffs (the dual solutions) are the only ones surviving in the core.

By contrast, the core is usually the improper concept for the discussion of simple games. Here, the core is empty unless there are veto players and if those are present, the core is not very decisive.

Historically the first set valued solution concept is the ***vNM–Stable Set*** (von Neumann–Morgenstern solution) This is a (not necessarily unique) set of ***imputations*** (i.e.,measures $\boldsymbol{x}$ with $\boldsymbol{x}(I) = \boldsymbol{v}(I), \boldsymbol{x}(\{i\}) \geq v(\{i\}), i \in I$) with the property of *external stability* and *internal stability*. The construction is much more sophisticated, the class of games admitting of such solutions is as yet not specified. However, traditionally zero–sum homogeneous games (even with multiple levels) admit of such solutions (von Neumann–Morgenstern). For convex games, the core is the only vNM–Stable Set (Shapley) and, more recently, this has also been established for large (non atomic) exact MIN–games (i.e., all measures in the representation having the same total mass). (Einy, Holzman, Monderer, Shitovitz). For non-exact MIN–games (in particular large ones, the non atomic context) vNM–solutions can be established predicting cartelization in such games (convex combinations of normalized measures absolutely continuous to the representations), provided the representation is "orthogonal" (Rosenmüller–Shitovitz).

The first point valued concept to be mentioned is the ***Shapley value*** (Shapley). It cannot always be seen as a "solution" but rather represents an "expected payoff", "average marginal contribution", or "power index" for players facing a game. There is a closed formula which, for any given $\boldsymbol{v}$, assigns

$$\Phi_i(\boldsymbol{v}) \;=\; \frac{1}{n!} \sum_{\{S \,\mid\, i \in S\}} \frac{(s-1)!(n-s)!}{n!} \Big(\;\; \boldsymbol{v}(S) - \boldsymbol{v}(S - \{i\}) \; \Big) \tag{9}$$

to player $i \in I$ ($s := |S|, n := |I|$).) On the other hand, the Shapley value admits of a unique set of axioms defining it (Pareto efficiency, additivity, covariance with renaming the players by a permutation, and zero assignment for dummy players). It can also be seen as the unique extension of the uniform distribution defined as a solution for “unanimous” simple games. The Shapley value also obeys an “equivalence principle”: for large market games it converges towards the core.

There is an extensive version of the Shapley value for large or, more precisely, non atomic games, the ***Aumann–Shapley*** value. Basically, the player set is a measure space as above and one considers set functions of bounded variation as “games” . On certain subspaces of such kind of games there exists a mapping assigning to each game a ($\sigma-$) additive measure and obeying certain axioms which generalize the finite version. A value does not exist on all bounded variation games but typically on the subspace generated (with respect to the bounded variation norm) by polynomials of non atomic measures, on absolutely continuous games, etc. (AUMANN–SHAPLEY). The theory of the Shapley value is now a new field extending classical functional analysis in considering measure–valued functionals with certain invariance properties.

A value theory exists for countably many players as well, however, it is less developed as certain pathologies prevent the generalization.

A further family of solutions is provided by nucleolus type concepts (SCHMEIDLER). in order to obtain the ***nucleolus*** (very generally speaking) , one lists the ***excesses***

$$e(S, \boldsymbol{x}, \boldsymbol{v}) \;=\; \boldsymbol{v}(S) - \boldsymbol{x}(S)$$

(*“reasons to complain”*) for any imputation $\boldsymbol{x}$ in a (weakly) decreasing order, say

$$\theta(\boldsymbol{x}) \;:=\; (\ldots, e(S, \boldsymbol{x}, \boldsymbol{v}), \ldots). \tag{10}$$

Then the nucleolus $\boldsymbol{\nu}$ is the unique imputation such that $\theta(\bullet)$ is lexicographically minimal, i.e.

$$\theta(\boldsymbol{\nu}) \quad \preceq_{\text{lexic}} \quad \theta(\boldsymbol{x}) \quad \text{for all imputations} \quad \boldsymbol{x}. \tag{11}$$

Thus, the largest complaints against a proposed solution are minimized, then the second largest, etc.

The *modified nucleolus* or **modiclus** $\boldsymbol{\psi}$ lists *bi–excesses*

$$e(S, \boldsymbol{x}, \boldsymbol{v}) - e(T, \boldsymbol{x}, \boldsymbol{v})$$

and proceeds accordingly (SUDHÖLTER). The modified version involves the *dual game* which reflects the preventive power derived from the original game.

We are not in the position to describe the abundance of solution concepts available. The Kernel and Bargaining Set (with a great many variations) are defined in the spirit of the nucleolus but emphasize different versions of bargaining. Objections and counterobjections are raised by the players and a stability is achieved when arguments are in balance (Davis–Maschler, Peleg). Axiomatizing many solution concepts (like the core, the nucleolus and many others) has recently been successful on a large scale by applying the construction of ***reduced games*** (Peleg, Sobolev).

The discussion of solution concepts for NTU–games is based on the results in TU–territory. A general approach is to strive for a generalization of TU concepts. The core of an NTU game accordingly can be defined as the set of Pareto efficient payoffs of the grand coalition such that no smaller coalition can improve upon. As payoffs are individual (not to the coalition) in the NTU context, the term "improve" has to be clarified (all players are strictly better off or one player is strictly better off and all the others do not lose).

For the Shapley value, the way to extend the concept to NTU framework is not unique and involves fixed point procedures. (Shapley, Aumann, Hasanyi, Kalai–Samet, Hart–Mas-Colell) Nucleoli can be defined after a discussion of the version of excess one wants to apply. The Bargaining Set and Kernel allow for various modifications (Granot–Maschler, Zhou). Thus, with NTU games in general, solutions concepts are less canonically defined; they are extensions of the TU–concepts and differ in the various approaches.

4 The Equivalence Principle

4.1 Walrasian Equilibrium

The concept of a *competitive* or *Walrasian equilibrium* is deeply rooted in the history of game theory and its influence on the development of game theory cannot be overstated. Moreover there are several interesting structural relations between the Walrasian equilibrium and different solution and equilibrium notions of game theory. Nevertheless the Walrasian equilibrium is not a game theoretical concept, it rather describes equilibrium states of economic systems, briefly, *economies*. There are various models with different levels of complexity to describe an economy. An economy is defined as a list of date describing the set of agents in the society, the set of possible consumption bundles in a vector space of commodities and the agents' economic characteristics like initial endowments of commodities, sets of technologically feasible production plans, shares of agents in production possibilities and agents' preferences over commodities or, more generally, over allocations of commodities.

The case of an exchange economy $\mathcal{E}$ as defined before is particularly prominent because of its role in the development of economic theory and due to the fact that it represents many important features of more general economic systems. Preferences over commodity bundles are usually described by transitive, complete continuous binary relations. These are often assumed to have in addition some monotonicity or convexity property. Though such preferences are representable by utility functions u it is the purely ordinal aspect which underlines most of the economic analysis. Different utility representations generated from u by order preserving mappings from $\mathbb{R}$ to $\mathbb{R}$ do not influence most parts of standard economic analysis.

Like in Game Theory also in Economic Theory one is interested in the results of interaction of agents or markets. A fundamental difference lies in the fact that in contrast to Game Theory neoclassical competitive Equilibrium Theory assumes agents to be *price takers* and thereby explicitly precludes strategic interaction. Rather than regarding other agents' decision processes price takers just take prices as given and maximize their preferences on the budget sets which they determine by evaluating their initial endowments with the given price system.

To make the idea of a competitive market equilibrium precise consider a finite set I of agents with initial endowments $a^i \in \mathcal{X} = \mathbb{R}^\ell_+$ and preferences represented by utility functions $u^i : \mathcal{X} \longrightarrow \mathbb{R}$ for all $i \in I$. Abbreviate $(u^i)_{i\in I}$ and $(a^i)_{i\in I}$ by u and a respectively. A mapping $x : I \longrightarrow \mathcal{X}, a \longmapsto x^i$ is called an ***allocation*** for $\mathcal{E}$. The allocation x is called ***attainable for the coalition*** $S \subset I$ if $\sum_{i\in S} x^i = \sum_{i\in S} a^i$, it is ***attainable*** if it is attainable for I. A *large economy* with many agents is defined in an analogue way as an ***Aumann economy*** i.e. a mapping $\mathcal{E} : (\mathbf{I}, \mathcal{B}(\mathbf{I}), \lambda) \longrightarrow \mathbb{R}^\ell_+ \times U : i \longmapsto (a^i, u^i) \equiv (a(i), u^i)$ where U is a space of continuous utility functions endowed with some suitable topology, $\mathbf{I}$ is the closed unit interval with its Borel-σ-algebra of admissible coalitions and the Lebegue (probability) measure λ. An *attainable allocation* is in this context an integrable map $x : \mathbf{I} \longrightarrow \mathbb{R}^\ell_+$ such that $\int_{\mathbf{I}} x(i) d\lambda = \int_{\mathbf{I}} a(i) d\lambda$. Rather than counting agents now sets of agents are measured. And instead of adding up their consumptions to *total* market consumption now integration determines *mean* or *average* market consumption.

A ***Walrasian equilibrium*** for an *Aumann* economy $\mathcal{E}$ is a pair (p^*, x^*), consisting of a linear price functional p^* on $\mathbb{R}^\ell$ and an attainable allocation x^* for $\mathcal{E}$ such that for λ-almost all $i \in \mathbf{I}$ (which is the technical version of "for all i") x^{*i} maximizes $u^i(x^i)$ on the *budget set* $B^i(a^i, p) := \{x^i \in \mathbb{R}^\ell_+ | p^* x^i \leq p^* a^i\}$. The definition of Walrasian equilibria for a finite exchange economy is analogous.

The first existence theorem for a finite economy due to Arrow and Debreu made use of the construction of an *abstract economy* or *generalized game* and of the existence proof of a social equilibrium for such a generalized game due to Debreu. A generalized game is an extension of a normal form game in such a way that every player in the game cannot freely choose his strategy in his strategy space but is restricted to some proper subset of it which varies in dependence of all other players' strategy choices. There may exist, therefore, strategy profiles which cannot be actually played because some coordinates violate restrictions implied by other coordinates. Playability already requires some consistency of choices. Among the consistent strategy profiles social equilibria are defined exactly as Nash equilibria. A Walrasian equilibrium can be seen as such a social equilibrium in a generalized (non-cooperative) game. It was shown later by Schmeidler that Walrasian equilibria of an economy may be even represented as Nash equilibria of some suitably designed normal form game. This result constitutes a non-cooperative example of the *Equivalence Principle*, which requires that two different solution concepts single out always the same set of allocations on a whole class of economies.

4.2 Walrasian Equilibria and Cooperative Solutions

The first instance of an equivalence theorem was Aumann's ***Core Equivalence Theorem*** stating the coincidence of Walrasian allocations and of core allocations for Aumann economies. To state this result formally one needs a definition of the *core* for an exchange economy. One way to proceed consists in extending to NTU-games the solution concept of the core which has been defined in SECTION 3.2 for TU-games. Then core allocations for the economy are defined as those allocations that induce utility profiles in the core of the NTU-game $V^{\mathcal{E}}$ induced by the economy $\mathcal{E}$. Although order preserving transformations of utilities induce transformed NTU-games also these games' core utility payoffs are exactly those induced by the core of the economy. An alternative approach followed in large parts of the literature in economic theory is to argue directly in the economy rather than its induced utility space. A very strong equivalence theorem is the *Bargaining Set Equivalence Theorem* due to MAS-COLELL from which Aumann's theorem follows immediately.

Consider an Aumann economy $\mathcal{E}$. Let x be an attainable allocation for $\mathcal{E}$. An ***objection*** of a coalition $S \in \mathcal{B}(\mathbf{I})$ to the allocation x is an allocation y that is attainable for S and satisfies:$u^i(y^i) \geq u^i(x^i)$ for λ-almost every $i \in S$ and $\lambda(\{i \in S | u^i(y^i) > u^i(x^i)\}) > 0$. A ***counter objection*** of a coalition $T \in \mathcal{B}(\mathbf{I})$ to the allocation y is an allocation z that is attainable for T and satisfies $u^i(z^i) > u^i(y^i)$ for λ-almost every $i \in T \cap S$ and $u^i(z^i) \geq u^i(x^i)$ for λ-almost every $i \in T|S$ and $\lambda(\{i \in T|S|u^i(z^i) > u^i(x^i)\}) > 0$. An objection to which no coalition has a counterobjection is called ***justified***.

The ***Core*** of the economy $\mathcal{E}, \mathcal{C}(\mathcal{E})$, is the set of attainable allocations to which there is no objection. The ***Mas-Colell Bargaining Set*** $\mathcal{MCB}(\mathcal{E})$ is the set of attainable allocations to which there is no justified objection. Moreover, denote by $\mathcal{W}(\mathcal{E})$ the set of Walrasian equilibrium allocations. Obviously by definition one has $\mathcal{C}(\mathcal{E}) \subseteq \mathcal{MCB}(\mathcal{E})$. Also, for finite exchange economies as well as for Aumann economies $\mathcal{W}(\mathcal{E}) \subseteq \mathcal{C}(\mathcal{E})$ holds true. Hence $\mathcal{W}(\mathcal{E}) \subseteq \mathcal{C}(\mathcal{E}) \subseteq \mathcal{MCB}(\mathcal{E})$. Now ***Mas-Colell's equivalence theorem*** states for Aumann economies with monotonic preferences that $\mathcal{W}(\mathcal{E}) = \mathcal{MCB}(\mathcal{E})$. An immediate consequence is Aumann's Equivalence Theorem, which asserts $\mathcal{W}(\mathcal{E}) = \mathcal{C}(\mathcal{E})$.

While the set valued concept of the core reflects some stability of the resulting outcomes the value represents more the intuition of compromise or average. Despite their differences, formally and regarding their intentions, both concepts satisfy the Equivalence Principle. Value allocations may be distinguished according to whether exchange economies with (ordinal) preferences or with (cardinal) concave utility functions are considered and whether an induced TU-game or NTU-game is analyzed. A first weak equivalence result for the Shapley TU-value and the utilities derived from Walrasian competitive allocations in TU-markets was given by AUMANN and SHAPLEY. After AUMANN distinguished between ordinal and cardinal *value* allocations he derived an equivalence theorem for Aumann economies between ordinal value allocations, cardinal value allocations and Walrasian allocations. Being based on differentiability assumptions this equivalence does not hold in the same generality as core or bargaining set equivalence. Without differentiability value allocations in Aumann economies are Walrasian but not vice versa (HART).

It is possible to axiomatize the set valued map from the class of Aumann economies with uniformly smooth preferences which exactly singles out the Walrasian allocations (DUBEY and NEYMANN). Clearly the core, the value and the Mas-Colell Bargaining set satisfy these axioms. The value equivalence may be destroyed, however, when the Shapley value is replaced by the Harsanyi value (HART and MAS-COLELL).

4.3 Approximate and Weak Equivalence

The Equivalence Principle represented by the unique characterization of Walrasian allocations in atomless economies by various important solution concepts from game theory contains a remarkable insight. The strategic and cooperative options of agents represented by the different solution concepts totally lose their impact on the allocation process under the dominating power of perfect competition of many. Clearly, an atomless economy as well as atomless games are only abstractions and the immediate question arises as to what extent this equivalence remains approximately true when one goes from a continuum model to a large finite model. Historically the analysis of large finite economies preceeded even that of the continuum case. Already 1838 EDGEWORTH had argued that if a finite exchange economy is replaced by its n-fold replication, i.e. where identical copies of each agent appear, then the set of allocations, now called core, shrinks to the set of Walrasian allocations if n tends to infinity. This insight had been put into a rigorous theorem by DEBREU and SCARF in 1963. So Aumann's Core Equivalence Theorem came as an elegant continuum version transmitting the same economic insight. The question remained whether also for large economies more general than the quite artificial replica economies an approximate equivalence would hold true. The first very general core limit theorem due to BEWLEY was followed by several others in 1970. The final answer is a result due to ANDERSON and to DIERKER which confirms the approximate equivalence and even allows to estimate the competitive gap between Core and Walrasian allocations for any given finite economy.

The Core Equivalence Theorem for general Aumann economies correctly reflects the fact that the gap between core allocations and Walrasian allocations converges to zero if the number of agents tends to infinity. It does not guarantee, however, that the core or the Walrasian allocations in such a sequence of increasing economies converge to the core or Walrasian allocation of the limiting Aumann economy. $\mathcal{C}(\mathcal{E})$ and $W(\mathcal{E})$ may be much larger sets containing properly the sets $\mathcal{C}(\mathcal{E}^n)$ and $\mathcal{W}(\mathcal{E}^n)$ for the n-th economy in the sequence. Only for regular economies with smooth preferences that have only finitely many locally constant Walrasian allocations, even the convergence of the core and of the Walrasian allocations holds true. But neither this strong nor the above weaker convergence property of the core turns out to be true for the Mas-Colell bargaining set. Even in perfectly nice replica economies an analogue to the core shrinking is not true. More specifically the set of all individually rational Pareto optimal equal treatment allocations not in the Mas-Colell bargaining set converges to the empty set in the Hausdorff distance (ANDERSON, TROCKEL AND ZHOU). This result considerably weakens the explanatory power of Mas-Colell's Equivalence Theorem.

The Core Equivalence Theorem has counterparts for several different economic scenarios. It holds true even for an infinite dimensional function space of commodities as was first shown by BEWLEY. If also the number of agents is infinite, a situation called "large square" by OSTROY, then depending on the sizes of the commodity space and of the space of agents a core equivalence can be proved if markets for all commodities are *thick*, i.e. on each commodity market enough agents are active to suppress via perfect competition effective strategic interaction.

It is even possible to get (weak) core equivalence theorems for finite economies. This requires a special structure which allows to represent them by specific TU market games. Techniques and arguments akin to those leading to core equivalence in finite linear production games (BILLERA and RAANAN, OWEN) then allow to prove finite convergence of the core of the game to its equilibrium utility payoff profile (ROSENMÜLLER).

4.4 The Nash Program

A weak approximate equivalence result between Walrasian equilibria and Nash equilibria of derived noncooperative games in normal form different from Schmeidler's game is based on earlier work on noncooperative exchange due to SHAPLEY and SHUBIK. In these games quantities of commodities build the strategy sets and the Nash equilibria are called COURNOT equilibria. For a sequence of increasing economies converging to an Aumann like continuum economy it can be stated that in many situations Cournot equilibria of the finite economies converge to a Walrasian equilibrium of the limiting continuum economy. The converse, namely that Walrasian equilibria of the continuum economy are limits of Cournot equilibria holds true whenever the Walrasian equilibria are regular.

One may interpret the Equivalence Principle in a broader sense. Then it relates any two solution concepts on certain classes of economies or of games to each other. An example for this is provided by what is called today the *Nash program*. This term refers to Nash's contributions in which he tried to support his cooperative bargaining solution as a strategic equilibrium of some normal form game. It reflects the intention to give axiomatic solution concepts of cooperative game theory some solid basis by deriving it from suitably modelled noncooperative strategic interaction. Several cooperative solution concepts can be supported in a noncooperative way either by normal form games or by games in extensive form. The most popular example is the alternating bargaining game due to RUBINSTEIN, where some discount rate devaluates the payoff during the course of the extensive form game. If this discount rate is nearly negligable than the unique *subgame perfect equilibrium* of this game approximates closely the Nash solution (BINMORE, RUBINSTEIN AND WOLINSKI). A direct support of the Nash solution as the unique strict Nash equilibrium can be established via the interpretation of an n-person bargaining game as a special economy whose unique Walrasian allocation coincides with the Nash solution (TROCKEL). Most of the results guaranteeing the support of a cooperative solution concept by equilibria of noncooperative games can even be interpreted as mechanism theoretical *implementations*. To achieve this goal an *outcome space* has to be available and the cooperative solution concept must be representable by some *social choice rule*. These concepts play a central role in the next section.

5 Mechanism Theory

5.1 Historical Background

Although mechanism theory is now established as a part of game theory its roots are laying in the theory of markets.

The idea of a mechanism originated from discussions in the 1930s concerning the economic feasibility of socialism. Market versus planning mechanisms were advocated in a by now famous debate between the economists Friedrich von Hayek, Oskar Lange, Abba Lerner and Ludwig Mises. Informational aspects of decentralization built the starting point for a search for suitable mechanisms to organize an economy. But a class of objects called "mechanisms" which could be compared as to their effectivity in organizing markets in a decentralized way was lacking. And despite Koopmans' use of the term "allocation game" for the description of the tatonnement in competitive markets or centrally planned economies neither games nor mechanisms as technical terms in their present understanding were used to formalize the problem. The first formal version of a mechanism was that of an *adjustment process*. At each point of (discrete) time any of n agents in the economy is supposed to send a message from a given message set. This message is publicly observable as is the state of the economic environment $e = (e^1, ..., e^n)$ that represents the exogeneously given economic data like preferences, endowments or production possibilities. A set of response functions $f^i, i = 1, ..., n$ determines the agents' messages at the next period of time via $m^i(t+1) = f^i(m(t), e), i = 1, ..., n$, where $m(t) = (m^1(t), ..., m^n(t))$ denotes the agents' messages at time t.

The first version of a mechanism is a simple dynamic one which still lacks any strategic features. Subsequently a static version was considered where the f_t^i where the same for all t and a stationary message profile m^* satisfying $m^*(e) = (f^1(m^{1*}(e^1)), ..., (f^n(m^{n*}(e^n)) =: f(m^*(e))$ represented a privacy preserving equilibrium where for each agent knowledge of his own characteristics and the message profile provides sufficient information for the choice of his message. Still this verification procedure had no strategic ingredients.

Yet, the availability of a concept of a mechanism led to several interesting questions starting from the basic theorems of welfare economies. These theorems state that perfectly competitive prices induce a Pareto efficient allocation and that any Pareto efficient allocation may be induced by suitable competitive prices in neoclassical economic environments. Now one could ask whether there are other decentralized mechanisms, more effective in terms of information processing than the Walrasian tatonnement; or others defined on larger economic environments than the classical one, that would lead to Pareto efficiency. Also other desirable properties like for instance fairness or envy-freeness could replace Pareto efficiency as a social goal to be implemented by suitable mechanisms. Also an extension of interest from the economic system to general social systems could be considered. A next step towards our present day notion of a mechanism was, accordingly, the idea of realizing a given social goal represented by some correspondence F through an adjustment process and some outcome function h. Then $h(m^*(e)) \in F(e)$ determines the allocation on the basis of the stationary message profiles.

It is the lack of incentive aspects that distinguishes this realization of socially desired goals from implementation by a mechanism formalized as a game form. The agents' behavior is prescribed here by the response function rather than freely chosen on the basis of preferences and strategic decisions. The inclusion of the important aspect of strategic choice of messages was caused by Samuelson's sceptical attitude toward the Lindahl equilibrium, the analogue of the Walrasian competitive equilibrium for economies with public goods. He pointed to likely misrepresentation of preferences in the special case that agents' messages are reports on their own preferences.

In today's terminology Samuelson claimed the incentive incompatibility of the Lindahl equilibrium. Obviously, an adjustment process is not the adequate concept for dealing with incentive problems. A natural approach appears to be one via non-cooperative games: The sets of strategies for the n agents (now players) are sets of preference relations on allocations including their true preferences. This specific kind of game called ***direct revelation game*** or just ***direct game*** allows to interpret a player's message as his lying or telling the truth.

It turned out later that not only Samuelson was right, but that whenever truth telling is a Nash equilibrium in a direct game it is even an equilibrium in dominant strategies. This means that truth telling, if it is consistent with the Nash equilibrium, it is even optimal for each player independent of whether the others tell the truth or lie. There are no truth telling equilibria in direct games with public goods and, moreover, the same happens in classical private good exchange economies - provided the rules are designed in such a way that truth telling is individually rational as well as Pareto efficient. Pareto efficiency, which one has automatically when using dominance equilibrium, has to be added as a desirable criterion when one is dealing with the Nash equilibrium. The impossibility results in the framework of neoclassical private goods and public goods environments allowed for several possible reactions. Would the main picture change for large economies? Would similar impossibility hold true in a social choice framework where the continuum of allocations would be replaced as an outcome space by some finite set?

Would a change of the employed equilibrium concept, i.e. Nash equilibrium rather than dominance equilibrium lead to an improvement?

For neoclassical private goods economies it turned out that the incentive to misrepresent preferences tends to zero when the number of agents in the economy becomes very large. In public goods economies, on the contrary, the incentives to misrepresent preferences increase with the number of agents.

5.2 Implementation of Social Choice Rules

In this section the focus is on ***social choice rules*** and their implementation by equilibrium outcomes of suitable normal form games.

Let $I = \{1, ..., n\}$ now be the set of players' positions, A be a non-empty set, called ***outcome space*** and M^i sets of possible messages m^i among which a player in position i may choose, $i \in I$.

The outcome space A represents all possible social states for a n-person society. A social choice rule associates with any profile $u = (u^1, ..., u^n)$ of utility functions $u^i : A \longrightarrow \mathbb{R}$ a set of states considered as socially desirable for a population represented by u. To formalize this idea, let U^i be non-empty sets of utility functions on A representing all admissible utility functions that players in position i may have, $i \in I$.

Consider $U \subseteq U^1 \times ... \times U^n$. A (point valued) mapping $f : U \longrightarrow A$ is called ***social choice function***. A set valued mapping $\mathbf{F} : U \longrightarrow \mathcal{P}(A)$ is called ***social choice rule***.

The planner's task is it to make sure that any population of rational agents represented by some $u \in U$ that obeys the rules designed by him automatically realizes some social state in $\mathbf{F}(u)$. To formalize this idea one needs the concept of a game form. For this one considers a map $g : M^1 \times ... \times M^n \rightarrow A$ $m := (m^1, ..., m^n) \longmapsto g(m)$. Such a map is called an ***outcome function***. The list $(M^1, ..., M^n, g)$ is called ***game form*** or ***mechanism***. Because of the bijective association between $(M^1, ..., M^n)$ and $M := \Pi_{i \in I} M^i$ one denotes a mechanism alternatively also by (M, g). The following observation is crucial for mechanism theory. For each admissible profile of utility functions, i.e. for each $u \in U$, the mechanism (M, g) induces a game $\Gamma_{g,u}$ in normal form defined by $\Gamma_{g,u} := (M^1, ..., M^n; u^1 \circ g, ..., u^n \circ g)$. For obvious reasons $\Gamma_{g,u}$ is also denoted $(M, u \circ g)$.

Denote by $NE(\Gamma_{g,u})$ the set of Nash equilibria of $\Gamma_{g,u}$ and the Nash equilibrium outcomes and payoffs for $\Gamma_{g,u}$ respectively, $NO(\Gamma_{g,u}) := g(NE(\Gamma_{g,u}))$ and $NP(\Gamma_{g,u}) = u \circ g(NE(\Gamma_{g,u}))$. Note, that in general the sets $NO(\Gamma_{g,u})$ and $NP(\Gamma_{g,u})$ vary with u as does $\mathbf{F}(u)$.

The planner's design problem is it to find a g such that equilibrium behavior in any game $\Gamma_{g,u}$ with $u \in U$ induces a socially desirable outcome. The notion of Nash-implementation of a Social Choice Rule makes this idea precise.

A mechanism (M, g) ***Nash-implements*** a social choice rule $\mathbf{F}$ on the domain U if $NO(\Gamma_{g,u}) \subseteq \mathbf{F}(u)$ for all $u \in U$.

There is some disagreement in the literature about whether this (weak) implementation is the right concept or rather full implementation requiring equality instead of only inclusion.

Notice, however, that $NO(\Gamma_{g,u})$ is not a satisfying non-cooperative solution concept by its own. It rather collects all singleton-valued equilibria, each of which is a solution concept and if played excludes simultaneous play of the other equilibria. Therefore, consistent strategic behavior in the framework of non-cooperative games can only result in some point in $\mathbf{F}(u)$ and never cover all of $\mathbf{F}(u)$. However, in the case of weak but not full implementability of a social choice rule one might want to understand why some points of $\mathbf{F}(u)$ result from equilibrium behavior while others do not. In case of full implementation this problem does not arise. For social choice functions both notions of implementation coincide. Several contributions in the literature are concerned with providing sufficient and necessary conditions for the Nash-implementability of a social choice rule (see ***Mechanism Design***).

Replacing the concept of the Nash equilibrium by some of its refinements like *dominance equilibrium, undominated Nash equilibrium, strong* or *strict Nash equilibrium* leads to the analogous notions of weak and full implementation.

5.3 The Revelation Principle

One of the most well known and most applied (group of) result(s) of mechanism theory is the *Revelation Principle*.

It is represented by some formal theorems ascerting for various notions of equilibrium under varying assumptions versions of replacement of implementing mechanisms by direct mechanisms. The general value of this principle lies in the fact the planner when looking for a suitable mechanism may restrict his search to the much smaller family of direct mechanisms. The problem with the Revelation Principle lies in the fact that in the class of situations which allows its strongest and most satisfying version this leads to general impossibility results, while in cases where it would be most helpful only such versions hold true which have considerable drawbacks. This dichotomy concerns dominance equilibria versus Nash or Bayesian Nash equilibria. Once the equilibrium concept is fixed literally the same version of the revelation principle can be stated. But clearly it's truth depends on the employed equilibrium concept. While the dominance equilibrium by definition is unique a game may have several Nash equilibria. *Bayesian* or *Bayesian Nash* equilibria are Nash equilibria of normal form representations of so called *Bayesian games*. In contrast to a standard normal form game where there is common knowledge among the players about the data of the game including all players' payoff functions one may formalize strategic interaction in situations of incomplete information about the characteristics of the fellow players. This may be expressed by probability measures the players have on their coplayers' characteristics. Now in this framework not only payoff functions but also these probability distributions are part of players' characteristics. The characteristics of a player are called his ***type***.

Harsanyi's idea to model strategic scenarios under incomplete information as Bayesian games with *consistent beliefs*, i.e. player's probability distributions on all players' type space as marginals of the same *a priori distribution* was consistently carried out by Mertens and Zamir and later by Brandenburger. A sketchy description of a simple special case follows. A *Bayesian game* $\mathbf{B}$ is given by a list $(M^1, ..., M^n, \theta^1, ..., \theta^n, p^1, ..., p^n, C^1, ..., C^n)$ with the following interpretation:

M^i and θ^i are player i's finite *action set* and finite *type space*, respectively. There is some $\vartheta^i \in \theta^i$ determining (p^i, C^i) uniquely, i.e. $(p^i, C^i) \equiv (p^i(\vartheta^i), C^i(\vartheta^i))$. Here p^i is a probability measure on $\Pi_{j=1,\ j\neq i}^{n} \theta^j$ and $C^i : M^1 \times ... \times M^n \longrightarrow \mathbb{R}$ is player i's payoff function. A ***strategy*** $s^i \in \mathcal{S}^i$ is a map from $\theta^1 \times ... \times \theta^n$ into M^i, $i \in I$. A ***Bayesian Nash equilibrium*** is a strategy profile, which for every type of every player optimizes his expected payoff given the other players' equilibrium strategies. In fact this is a Nash equilibrium of a suitably defined normal form game $\Gamma_{M,\theta,p,C}$.

Define the functions $\bar{C}^i : \mathcal{S}^1 \times ... \times \mathcal{S}^n \longrightarrow \mathbb{R}$ by $\bar{C}^i(s^1, ..., s^n)$
$= \sum_{\vartheta^i \in \theta^i} \int_{\theta^{-i}} u^i(\vartheta^i)(s^1(\vartheta^{-1}), s^2(\vartheta^{-2}), ..., s^n(\vartheta^{-n}))dp^i$.

Then a Bayesian Nash equilibrium of the Bayesian game defined above is a

Nash equilibrium of the normal form game $\Gamma = (\mathcal{S}^1, ..., \mathcal{S}^n, \bar{C}^1, ..., \bar{C}^n)$. As any Bayesian game has a normal form representation a Bayesian equilibrium may be seen as a special Nash equilibrium. On the other hand, if there is common knowledge among the players about all players' types then $p^1 = ... = p^n$ degenerates to a Dirac measure with total mass one on the true type profile. Expectation building as in $\bar{C}^i$ becomes trivial. Also every normal form game allows representation as a Bayesian game with degenerate type spaces and beliefs. In this sense any Nash equilibrium is a Bayesian Nash equilibrium. It is this fact which has been employed by REPULLO to establish versions of the Revelation Principle for Bayesian as well as for Nash equilibrium.

The basic idea of the *Revelation Principle* is it to replace some game form that implements a social choice rule, i.e. which has only socially desired equilibria, by a direct game form, where truthtelling constitutes an equilibrium. Depending on the employed equilibrium concept, weak or full implementation, the number of equilibria and whether the social choice rule is even a social choice function, one may derive results of different generality under the heading Revelation Principle. The power of this principle is very often overestimated as its different versions are often not distinguished. For a better understanding a second, weaker notion of implementation for direct games is needed. Let $\mathbf{F} : U \longrightarrow \mathcal{P}(A)$ be some social choice rule and E some equilibrium concept. The direct mechanism (U, g) ***truthfully*** E ***- implements*** $\mathbf{F}$, if for each utility profile $u \in U$ the strategy profile $u' \equiv u$, i.e. the truth is an E-equilibrium satisfying $g(u') \equiv g(u) \in \mathbf{F}(u)$. Note that truthful implementation in case of multiple equilibria is consistent with other non-truthful equilibria having outcomes outside $\mathbf{F}(u)$. Moreover such a non-truthful equilibrium may even Pareto-dominate the truthful one.

If in a direct mechanism implementing a social choice rule the E-equilibrium is always unique then truthful implementation coincides with weak implementation. In this case the revelation principle is quite strong. In fact, the E-equilibrium is then even in dominant strategies and the mechanism is called ***incentive compatible***. In the general situation, however, truthful E-implementation is much weaker than E-implementation and it is more than euphemistic to claim that restriction to a direct mechanism is possible without loss of generality. Nevertheless the revelation principle finds frequent use in many applications of mechanism design like the theories of auction design, contract design or voting.

6 Repeated Games

6.1 Evaluations

The general stochastic game with infinite horizon offers a host of new problems. Convergence of the payoffs can be established by either discounting or averaging and both versions exhibit fascinating aspects.

The version with ***discounted evaluation***

$$C^i(x,y) = (1-\rho)\sum_{t=1}^{\infty} \rho^{t-1} f^i(x_{t-1}, y_t) \tag{1}$$

is thought to be much closer to the finite game: after waiting sufficiently long the payoffs in the tail game are so heavily discounted that they do not bear relevant influence. This observation can be made precise and is the basis to the first approach due to SHAPLEY who established the value of a stochastic zero sum game. Here every player observes a finite set of matrix games repeatedly played, the transition being Markovian.

The payoff indicated by the ***averaging evaluation***

$$C^i(x,y) \;=\; \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} f^i(x_{t-1}, y_t) \tag{2}$$

(with various definitions concerning the limiting procedure) offers the modeling of a much more unsecure future. No essential payoff is contributed to every finite round of initial games. The essential payoffs will be achieved in the far away future and the nearby present may be used for signaling, agreements about contracts, punishments for violation of such contracts, indication of regret and the like.

The transition between discounted versions and average versions technically is manipulated by a "Tauberian Theorem", which links the behavior of the coefficients of a power series with the limiting behavior of the corresponding holomorphic function. This way GILETTE established the first connection between the values of the two types of stochastic games. The asymptotic behavior of the value of the discounted game when the discount factor approaches 1 turns out to be the key to treating the average situation. The value is algebraic in a neighborhood of the limit (BEWLEY–KOHLBERG) and, as a consequence, it turns out that the general stochastic zero sum game with averaging evaluation has a value (NEYMAN–MERTENS).

6.2 Folk Theorems

For the non zero sum case the situation is more complicated. A first approach is established by what is known as a "Folk Theorem" because the concrete copyright is not easily established. According to this theorem one considers a bimatrix game infinitely often repeated with average evaluation (***the supergame***). It is not hard to see that the payoffs constitute the convex hull of the bimatrix game's payoffs, meaning that the repeated game yields the same payoffs as correlated strategies in the one shot game. The payoffs resulting in the latter game can be interpreted as a bargaining problem (cf. SECTION 3). Now it turns out that all the individually rational payoffs of this bargaining situation can be obtained by Nash equilibrium payoffs of the supergame.

To some extent this result is disappointing, as Nash equilibria do not even yield Pareto efficient payoffs, not to speak of a "solution". The attempt to improve upon this situation (RUBINSTEIN) by introducing subgame perfect strategies however, yields the same result.

6.3 Repeated Games with Incomplete Information

With incomplete information stochastic games are even more difficult to approach. The field which is surveyed most extensively is the one of ***repeated games*** with incomplete information. This technical term denotes a type of game in which the players are facing an information structure which is established by chance at the beginning of the game. We may think of a finite set of matrices (bimatrices), the games played in the various states of nature. One of the states is chosen by chance and the corresponding game is then being played repeatedly. In addition, there is an information matrix which reveals certain data about the true matrix (the true game) according to the actions chosen by both players. Both players observe these signals and they may add certain messages of their own by playing certain "natural" sequences of actions hinting towards their observations.

Given the chance moves the players may form priors concerning the true game which can now be updated in view of the signals they receive from the opponents.

For the zero sum case an essential result is presented by the ***vex–cav Theorem*** (AUMANN–MASCHLER, MERTENS–ZAMIR). Consider the value of the expected game presented by the mixture of the states of nature. Now, when the distribution varies, the value of the expected game appears as a function on the probability simplex. We may define the lower convex envelope of this function to be the largest convex function dominated by the value function.

The *vex–cav* Theorem states that the value of the repeated game exists if the successive formation of lower convex and upper concave envelope of the value function as described above commutes.

Again, for the non zero–sum game the situation is more difficult to analyze. As in the Folk Theorem, the idea of cooperation in the canonical bargaining situation (with incomplete information) defined by the information structure and the prior about the true game can be formulated. Again one strives to establish connections between Nash equilibria of the repeated game and individually rational payoffs of the cooperative game.

Cooperation is easily described: A ***mechanism*** in this context is given by a set of correlated strategies each one conditioned on the announcement of the players with regard to their observations. The corresponding payoff is the expectation generated by the original distribution of the chance move choosing a game and the observations resulting thereof. However, players may choose to misrepresent their type, so we have to consider *incentive compatible* mechanisms in the context of a bargaining situation with incomplete information (cf. SECTION 5) .

On the other hand certain types of Nash equilibria (***joint plan equilibria***) can be formulated in the repeated game. A joint plan is a triple consisting of the following data. There is a set of finite sequences of actions of each player serving as signals. Next, there is a response kernel which stochastically yields such signals depending on the players' observations. And finally, there is a contract which, depending on the joint signals, yields actions in the repeated game.

It is rather obvious that a joint plan induces a mechanism: composing or mixing the choice of signals with the contract which bases (correlated) actions on signals obviously constructs a correlated choice of actions based on observations of signals, i.e., a mechanism in the above sense.

On the other hand, a joint plan can be implemented by a joint plan Nash equilibrium of the repeated game in the following way. The equilibrium yields the same payoffs as the mechanism which corresponds to the joint plan. Also it generates the signals of the joint plan in the first stages of the game (when played in equilibrium). Moreover, the distribution of the equilibrium path of the process yields certain sequences of actions "agreed upon" the frequencies of which are constructed in a way to imitate the correlations prescribed by the contract. Hence, the cooperative payoff resulting from the plan via the induced mechanism is also achieved in the non cooperative equilibrium implementing the joint plan.

An early result is due to AUMANN–MASCHLER–STEARNS and has been generalized by SORIN and later by SIMON. Accordingly, in a 2-person game with incomplete information on one side, there exists an incentive compatible joint plan which is also individually rational and admits of a corresponding Nash equilibrium.

7 Evolution and Learning in Games

7.1 Introduction

One important strength of noncooperative game models lies in the fact that, although equilibrium behavior of players is defined and of central concern in the analysis, also non-equilibrium behavior is possible. The very definition of a Nash equilibrium does by no means imply that it is realized in an actual play of the game nor that persons playing that game have to act as rational players. It is exactly this feature which allows it to analyze non-equilibrium behavior and, in a dynamic context, convergence or divergence properties of chosen strategy profiles. Still in these cases the persons playing the game choose, even if not necessarily in a rational, consistent or only purposeful way, their strategies from given strategy sets.

Evolutionary biology in contrast uses the formal concept of a game, in static as well as in dynamic models, in quite a different way. Imagine some large population of individuals (of some species) each of which is labeled with some number $i \in \{1, ..., n\}$, which is interpreted to represent some type. Now consider some $n \times n$-matrix with entries $a_{ij}, i, j = 1, ..., n$. From the point of view of standard game theory one can imagine a random device by which two individuals, labeled i and j are selected from the population to play the normal form bimatrix game, where i chooses one of n rows and j chooses one of n columns as strategies. If i chooses $h(i)$ and j chooses $k(j)$ the resulting payoff vector is $(a_{h(i),k(j)}, a_{k(j),h(i)})$. Obviously, this game is direct as the sets of strategies and of types coincide.

Now evolutionary biology deviates in two respects. First, the two chosen individuals do not have any choice, they just are programmed to pick $h(i) \equiv i$ and $k(j) \equiv j$, i.e. to tell the truth about their labels. So they do not *play* the game although this terminology is used in large parts of the literature. Clearly, no outside observer seeing i and j picking $h(i) = i$ and $k(j) = j$ could tell whether they play the game or follow some deterministic device. Secondly, the interpretation of the payoffs is now different. Rather than money or utility now *fitness* is considered to be the medium of payment.

Fitness is a stylized index representing reproductive success. Frequently the expected number of offsprings of an individual is taken to define fitness, sometimes however more refined and detailed definitions are used. In a dynamic context fitness payed out to some strategy increases the probability that an individuum labeled by that strategy will be chosen to be thrown as a strategy into the next round of the game.

A probability measure on the population determines the distribution of labels, from which the two "players" are drawn as a sample of size two. If this measure is concentrated on one label evolutionary biologists are talking of a *monomorphic state* of the population. Otherwise the state is *polymorphic*. Selten and Hammerstein in their chapter *Game Theory and Evolutionary Biology* of the Handbook of Game Theory stress the fact that a mixed strategy admits a monomorphic and a polymorphic interpretation. In the context described so far a mixed strategy is just a distribution over labels which are strategies. So it can be seen as a polymorphic state of the population. But one can obviously extend the framework sketched above by giving each individual a label which is a mixed strategy. The interpretation is that the program which rules the individuals' behavior is not anymore deterministic but stochastic. Now an individual *is* (characterized by) a mixed strategy. In this framework a measure on the population defines a distribution over labels, i.e. mixed strategies. Then a distribution concentrated on one mixed strategy represents a monomorphic state.

The first formalization of *evolutionary stability* due to John Maynard Smith and George R. Price has become the predominant concept of evolutionary game theory. If the label of some individuum in a population is meant to indicate some innate type of behavior then stability of this behavior type is defined as stability of the strategy that is representing it in the bilateral game. Stability of a monomorphic population, i.e. of the only type or label in the population is meant as immunity against the invasion of mutants. In a polymorphic population this idea extends to immunity against the invasion of mutants and perturbed versions of the incumbent type. Evolutionary strategies can be characterized in a static normal form context as well as in a dynamic context.

7.2 Evolutionary Stable Strategies

Although evolutionary stability is a dynamic concept it can be represented in a normal form game. Let σ, σ' be mixed strategies of the bimatrix game described above, i.e. distributions on $\{1, ..., n\}$. Let $A = (a_{ij}), i, j = 1, ..., n$ denote the payoff matrix for the row player. His expected payoff from σ if the column player is labeled σ' is $\sigma A \sigma'$.

Now a mixed strategy σ^* is an ***evolutionary stable strategy (ESS)*** if for all mixed strategies σ there is some $\bar{\varepsilon} > 0$ such that for all $\varepsilon \in (0, \bar{\varepsilon}]$

$$\sigma^* A[(1-\varepsilon)\sigma^* + \varepsilon\sigma] > \sigma A[(1-\varepsilon)\sigma^* + \varepsilon\sigma]. \tag{1}$$

This condition expresses the idea that after a monomorphic population represented by the mixed strategy σ^* has been changed by invasion of a small fraction of mutants programmed to play σ, still σ^* is more effective than σ. It is not hard to see that for an ESS σ^* the pair (σ^*, σ^*) is a Nash equilibrium of the symmetric game represented by A. Moreover, this Nash equilibrium has an additional property. Every strategy $\sigma' \neq \sigma^*$ that is an equally good response to σ^*, i.e. that satisfies $\sigma^* A\ \sigma^* = \sigma' A\ \sigma^*$, is necessarily a worse response to σ' than σ^*, i.e. $\sigma^*\ A\ \sigma' > \sigma'\ A\ \sigma'$.

In fact, for any Nash equilibrium (σ^*, σ^*) with this additional property σ^* is evolutionary stable. As not every equilibrium strategy needs to be evolutionary stable the concept of an ESS defines a refinement of the Nash equilibrium. Unfortunately, in a non-neglible class of non-pathological games ESS do not exist. A further drawback lies in the fact that the notion of ESS is defined only for monomorphic states of the population. These insights led evolutionary biologists to an explicit dynamic analysis. In a dynamic discrete time model the game described above is used repeatedly to determine the payoffs of two individuals based on their strategy labels. The distribution over labels, however, changes during time depending on the payoffs of the previous round.

High payoffs to individuals turn into higher probabilities of their label to be selected for the next round. If this idea is modelled rigorously it leads to a dynamical system which is known as *Replicator Dynamics*. A corresponding model can be built for continuous time. It can be shown that in the framework sketched above any mixed strategy that is an ESS represents a population state that is asymptotically stable with respect to the replicator dynamics. This surprising result shows that the ESS defined in a static monomorphic mixed strategy setting when interpreted as a population state in a pure strategy population displays a polymorphic dynamic local stability.

7.3 Learning in Social Contexts

There are essentially two problems in social systems to which methods and insights from evolutionary biology are applied. What evolutionary forces lead to behavior and social states that are compatible with perfectly rational interaction of players? And, if there are multiple equilibria, which ones are limit points of evolutionary processes? The first problem may be seen as inducing search for a specific *equivalence principle* or, alternatively, as a problem of evolutionary *implementation* of rational solutions. The second problem falls into the realm of *equilibrium refinement*.

Obviously, the extreme position of evolutionary biology that treats individuals as programmed automata is not very useful and needs to be abandoned in the context of social systems. Also the selection process can hardly be based on payoffs reflecting the number of offsprings. Rather behavioral pattern like *imitation, adaption, experimenting* combined with *learning* should determine the evolution. The prototype of learning models in a social context is ***fictitious play***. It was independently introduced 1951 by GEORGE W. BROWN and by JULIA ROBINSON as an algorithm for computing Nash equilibria in certain classes of games. Each player in a certain round of a game uses his observations of the frequencies of the other players' strategy choices in the past as the basis for estimating their present mixed strategies. Then he chooses a best reply to these. A game has the *fictitious play property* if every limit point of a sequence of strategy profiles generated by fictitious play is a Nash equilibrium of that game. Convergence behavior of fictitious play has been analyzed in increasing generality by ROBINSON, MIYASAWA, SHAPLEY AND ROSENMÜLLER. ROBINSON proved in 1951 that any finite zero-sum two-person game has the fictitious play property. The same was shown to hold true for every non-degenerate 2×2-bimatrix game in 1961 by MIYASAWA. In 1971 ROSENMÜLLER derived this result, as well as SHAPLEY'S non-convergence result from 1964 for a class of 3×3-bimatrix games, from a more fundamental analysis of convergence behavior of fictitious play based on eigenvalue considerations. More recent results are due to JORDAN and to MONDERER AND SHAPLEY in 1991 and 1996, respectively. JORDAN proved that for sufficiently dispersed priors fictitious play does converge to a Nash equilibrium. MONDERER AND SHAPLEY established the fictitious play property for *weighted potential games*.

There are other classes of games in which Nash equilibrium appears to be the natural solution, where, however, fictitious play does not single it out. FOSTER and YOUNG provided a special coordination game, which they called *Merry-Go-Round Game* where players follow fictitious play in cyclic patterns and never coordinate.

Why do players not learn that they are trapped in cycles? This question leads to the idea of *adaptive learning* where in fact players are less rational and less informed than in fictitious play. PEYTON YOUNG modelled a context of short-sighted players with limited information and reasoning ability and a probabilistic error rate independent across the players as a *Markov process* which he termed *adaptive play*. The basic ingredients are bounded by rational reactions to predictions of other players' past behavior, that are estimated from limited data and stochastically distorted. In their model of adaptive learning Foster and Young formulated an alternative stronger solution concept termed *stochastic stability*. In contrast to the ESS a stochastically stable equilibrium is even robust against persistent random perturbations. In his book *Individual Strategy and Social Structure* Young develops an extensive theory of adaptive learning and relates stochastic stability to other well known concepts of game theory like, for instance, *risk-dominance, focal points* or *maximin contracts*. Among several interesting results relating adaptive learning and stochastic stability to the theories of bargaining and of contract his characterizations of the Nash and of the Kalai-Smorodinsky solutions as limits of stochastically stable payoff vectors are particularly remarkable. They demonstrate that adaptive behavior of people can come arbitrarily close to results which are predicted by axiomatic solutions from standard game theory based on rationality assumptions on the players.

Another example for boundedly rational behavior leading to an established rationality based outcome is due to VEGA-REDONDO. He shows that in oligopolistic markets with n identical firms involved in Cournot type quantity setting competition a learning process combining elements of experimenting and of imitating has as its long-run outcome the unique symmetric Walrasian equilibrium.

These insights that repeated interaction of boundedly rational individuals may produce outcomes compatible with standard rationality based game theory support well the optimistic conclusions VERNON SMITH was drawing from various experiments as carried out in section 8.

8 Experimental Games

8.1 Introduction

It should have become quite clear by now that game theory is not *the* unified general theory of human interaction in groups or society. It is more a collection of concepts, methods and modelling devices whose use and power vary and depend on the specific problem and context. Common to all these game theoretic models except evolutionary game theory is the fundamental assumption of perfectly rational players. These have well defined goals, an unlimited power of learning, reasoning, understanding and computing and are assumed to use them in order to determine their decisions. Moreover in many situations heroic assumptions on information processing abilities and knowledge of players are made up to the extreme of *common knowledge*, that was discussed earlier. Experimental games roughly are concerned with the question to what extent game theory really describes and predicts human behavior. Starting with the degenerate but nevertheless highly complex and challenging theory of one person games or decision theory, where axiomatized models of consistent individual behavior like *expected utility theory* are tested, almost all aspects of game theoretic modelling have been put on trial in experiments. The design of experiments is an extremely sensible and ambitious business. The experimenter has to make sure that the players following precisely his rules actually play the game he has in mind. Moreover he must be able on the basis of the results to cleanly separate different possible causes and explanations. However ingeneous several experimental devices may be, the results cannot question game theory as a normative interpersonal decision theory. They only can possibly cast doubt on or even discredit it as a descriptive theory with predictive power. In the following sections some experimental results are briefly presented and discussed.

8.2 Repeated Prisoners' Dilemma

Tradition has it that the first attempt to test game theory was a hundred fold repetition of the Prisoner's Dilemma game played between the economists Alchian and Williams in 1950. In this game two players have two strategies each, namely *observing* or *breaking* a contract. If both break or both observe they receive ten units or one unit of payoff each, respectively. If they choose different strategies the player who breaks receives twenty units, the other one zero. To break is for both players a dominant strategy and constitutes the unique Nash equilibrium. Hence both players are condemned by rationally playing the equilibrium strategy to forgo the Pareto dominating and even Pareto efficient payoff vector $(10, 10)$

Although mutual defection represents the unique Nash equilibrium of the one shot normal form game the series of hundred plays showed mutual cooperation in 60 instances as opposed to only 14 mutual defections. Similar behavior was experimentally established by McKelvey and Palfrey about forty years later. According to the *Folk Theorem* this behavior would be consistent with equilibrium behavior in an infinitely repeated *Prisoners' Dilemma* game. In the finitely repeated game, however, it contradicts rational behavior because backward induction singles out consecutive mutual defection as the unique subgame perfect equilibrium. A large experimental literature including work by Rapoport and Chammah and by Selten and Stöcker has confirmed this basic result of deviating from game theory's rational equilibrium behavior as dictated by backward induction. More specifically, an *end effect* has been systematically observed where players after many stages of cooperation defected towards the end in accordance with the one shot game equilibrium. Evidence here seems to lead to a falsification of the rationality assumptions implicit in the game theoretical model.

8.3 Coordination Games

An essentially unsolved problem of non-cooperative game theory is it how to deal with multiple equilibria. A vast literature on refinements has been devoted to the attempt by excluding "implausible" equilibria to restore uniqueness. As it turned out all established refined equilibrium concepts, apart from the extreme dominance equilibrium, are open to non-uniqueness. This fact requires some kind of coordination between players and on the other hand opens the door for building up reputation in repetitions of the game. Experiments due to Schelling led to the insight that players have a surprising ability to coordinate their behavior and to base their selection of an equilibrium on shared senses of *salience* or *prominence* once they are in some joint social context. Driving on the right hand side of the street is the *focal point* equilibrium in most parts of the world, yet not everywhere.

Adding in a symmetric 2×2-bimatrix game with two equilibria in the upper left and the lower right corners for both players a third dominated strategy should have no effect on the frequency by which one or the other equilibrium is played. Nevertheless it has been shown in experiments that the specification of the opponents' payoffs for the added third dominated strategies may have dramatic consequences for the relative frequency of the choice of the first equilibrium. High amounts which could be gained there seem to have an impact.

Consider the purely symmetric game where the players' strategy sets are $\{R^1, R^2\}$ for the row player 1 and $\{C^1, C^2\}$ for the column player 2 and the payoff functions are given by $F^i(R^j, C^k) = \delta_{jk}, i, j, k = 1, 2$.

Despite game theory's inability to discriminate between the two equilibria (R^1, C^1) and (R^2, C^2) experiments showed differences depending on the cultural background of the players.

While a majority played (R^1, C^1), a group consisting of Chinese, Japanese and Koreans was more attracted by (R^2, C^2). The training to read rows from left to right might have created a salience for the majority when confronted with the game in bi-matrix form, while for the minority group (R^2, C^2) might have been more prominent. Experimenters' explanation here would be that some characteristics of the players not caught by a rational-choice based game model could be influencial for the outcome.

Another interesting insight from experiments comes from the observation that in games with several Pareto ordered Nash equilibria frequently not the Pareto optimal one was played. This hints to a serious coordination failure. A remarkable observation in this context has been made in series of experiments by Van Huyck, Batallio and Beil at Texas A & M University. In finite repetitions of some coordination game with multiple Pareto ranked equilibria observed behavior never came close to the payoff dominant, i.e. Pareto optimal equilibrium. But when the rights to participate in the game were auctioned before and the successful bids were commonly known among the players behavior always closely approximated the efficient equilibrium.

8.4 Bargaining Games

The most popular approaches to bargaining games are the axiomatic cooperative one due to Nash and the non-cooperative alternate offer model due to Rubinstein. Both approaches have been exposed to experimental tests. One of the fundamental difficulties in testing Nash's purely welfaristic model, in which only players' utilities determine the solution, lies in the fact that knowledge of the players' utilities is required. But these are not observable. Roth and Malouf devised a clever method to overcome this problem by arranging bargaining over distributions of lotteries. In the experiment the distribution of lottery tickets determines with which probabilities a random mechanism attributes high or low prizes to players.

When the row player 1 gets 80 tickets and the column player 2 gets 20 the consequence is that 1 has a 80 % chance of winning his high prize and a 20 % chance of winning his low prize. Important is the fact that the high and low prizes for the two players may be different. As in the Nash model the solution is independent of players' specific cardinal utility representations, the experimenter may set each players' high prize equal to one unit and its low prize equal to zero. Now, as players bargain over probabilities of receiving rather than over utilities, the experimenters do not need to know how much that is valued by the different players. Formally, players negotiate over how to divide the expected gain of one unit. The Nash solution selects the allocation that maximizes the product of the expected gains, the so called *Nash product*. This would require an equal division of lottery tickets.

According to ROTH and MALOUF the results in experiments centered around two distributions when players knew their opponents' prizes.

One is the Nash solution, i.e. equal number of lottery tickets for both players, the other one is the distribution that generated equal expected gains. So experiments provide partial support for the Nash solution.

The experiments testing Rubinstein's subgame perfect equilibrium predominate in the more recent literature. Truncated versions of Rubinstein's game were played by GÜTH, SCHMITTBERGER and SCHWARZ where, if no agreement was reached in round k, for $k = 1$ the game became an *ultimatum game*, in which both players lose unless an offer is accepted. Although subgame perfect equilibria prescribe a solution where the offering player gets close to everything, the average offer in experiments was 33 %! These results induced the authors to doubt that the subgame perfect equilibrium has predictive power in bargaining games.

Further experiments confirmed the impression that experimental results were in conflict with the predictions that money is a good proxy for utility payoff and that bargainers simply want to maximize their own incomes. They tend to indicate falsification of the joint hypothesis of expected payoff maximization and of backward induction.

Several experiments conducted by Albers in the *Institute of Mathematical Economics* in Bielefeld led him to the conclusion that *prominence* of numbers used as payoffs systematically influence players' decisions and outcomes in games.

8.5 Optimistic Conclusion

Vernon Smith who is running experiments in one of the world's largest experimental laboratories at the University of Arizona in Tucson comes after an uncounted number of experiments in bilateral bargaining games, oligopoly games, various sealed bid auctions and continuous double auctions to a surprisingly optimistic and puzzling result.

In situations of complete and common information about other players' preferences, which is the standard framework in which Nash equilibrium is considered the adequate solution concept for non-cooperative games, it has only scarce empirical confirmation.

However, in the more realistic context of repeated games with private incomplete information non-cooperative equilibria (and similarly Walrasian equilibrium) have according to experiments a high predictive power. Similarly, the experiments of Roth and Malouf on cooperative bargaining give a strong support to the Nash solution in bargaining with private information. Smith's irritating message is it that rational equilibrium and axiomatic bargaining outcomes lack experimental support in situations where they can be theoretically justified and have high predictive power where this is not the case. According to Vernon Smith "the theoretical problem that an equilibrium of a model might be approximated without agent knowledge or understanding of the model has important implications for the concept of common knowledge that allegedly underlies contemporary game theory."

9 Concluding Remarks

In his book *Choice and Consequence* Thomas Schelling has a chapter termed *What is Game Theory*. There he finds that in contrast to economic, statistical and decision theory *"there is no accepted name for whatever the field is of which "game theory" refers to the theoretical fontiers"* like economics, statistics or decisions. Despite several earlier developments described in the Introduction the appearance of von Neumann's and Morgenstern's book can be seen as the birth of Game Theory as a discipline. The authors, a mathematician and an economist, reflect perfectly the parent fields of the new discipline. 56 years later, when the recently founded *Game Theory Society* organized its first world congress "Game Theory 2000" in the Spanish Bilbao more than 700 researchers from several fields were a living proof of the enormous impact and fast growth of game theory as a field that had penetrated social sciences and found applications in biology, engineering and systems design. The treatment of game theory in the EOLSS in eight articles and this accordingly structured topic level survey does provide the fundamentals of the field and stresses some specific developments that were particularly important for the development of the discipline. It also hints to interesting new developments in the analysis of and experiments on evolutionary processes and learning in social systems that may turn out of value for future global sustainable development.

In the EOLSS Game Theory justifiably has been represented as one of the topics of the theme *Operations Research*.

It should not be hided that many researchers meanwhile would insist on a reverse relation, namely operations research as part of game theory. This point of view is consistent not only because of the variety of mathematical techniques employed but in particular under the aspect that any optimization, decision or planning problem can either be seen as an instance of a one-person game or, because of involved interpersonal interests and conflicts, falls in a natural way into the realm of interpersonal decision theory.

For instance a problem of *multicriteria decision making*, a field quite popular in Operations Research, can be modelled as a game where the various criteria represent interests of different players.

There are numerous subdivisions and fields of applications in game theory that are not even mentioned in the present treatment. The three volumes of the *Handbook of Game Theory* edited by Robert Aumann and Sergiu Hart carry the subtitle "with Economic Applications". Despite this focus the Handbook contains some sixty articles trying to cover game theory.

Chapters on Psychology, Inspection Games, Differential Games, Power and Stability in Politics, Game-Theoretic Aspects of Computing, Moral Hazard, Patent Licensing, Strategic Analysis of Auctions and Search reflect the enormous variety of a vastly developing discipline.

References

[1] R. J. Aumann, *A survey of cooperative games without side-payments* , Essays in Mathematical Economics in Honor of Oskar Morgenstern (M.Shubik, ed.), Princeton University Press, 1967, pp. 3 – 27, *Describes the state of the art at the time but actually constitutes the embarkation point for two dekades of research in NTU–games* .

[2] ______, *On the non-transferable utility value* , RM 55, Center for Research in Mathematical Economics and Game Theory, The Hebrew University, Jerusalem, 1983, *This paper completes the axiomatization of the Shapley value to NTU games* .

[3] R. J. Aumann and S. Hart (eds.), *Handbook of game theory with economic applications* , vol. 1 – 3, Elsevier Science Publishers, 1992, 1994, 2001, *These three volumes provide a thorough treatment of Game Theory and its applications and represent the state of the art* .

[4] R.J. Aumann and L.S. Shapley , *Values of nonatomic games* , Princeton University Press, Princeton, 1974, *Establishes Game Theory for a continuum or "ocean" of players* .

[5] M. A. Dimand and R. W. Dimand , *The history of game theory*, Routledge, London and New York, 1996, *This book descibes the historical development of Game Theory* .

[6] D. Fudenberg and D. Levine , *The theory of learning in games* , MIT Press, Cambridge, 1998, *A comprehensive treatment of evolutionary game theory and learning in games* .

[7] R. H. Gertner and R. C. Picker , *Game theory and the law*, Harvard University Press, Cambridge, 1994, *A thorough discussion of the application of game theory to law and legal problems* .

[8] D. Hurwicz, L. Schmeidler and H. Sonnenschein (eds.), *Social goals and social organization* , Cambridge University Press, Cambridge, 1985, *This volume contains extensive treatments of incentive problems and different areas of implementation and mechanism theory* .

[9] J. Kagel and A. Roth (eds.), *The handbook of experimental economics*, Princeton University Press, Princeton, 1995, *The Handbook contains several chapters covering experiments on games and decisions* .

[10] J.F. Mertens and S. Sorin (eds.), *Game-theoretic methods in general equilibrium analysis*, Kluwer Academic Publishers, Dordrecht, 1994, *This volume contains a variety of articles surveying the role of game theoretic solutions for competitive markets and discussing the equivalence principle.*

[11] M. Osborne and A. Rubinstein, *A course in game theory*, The MIT Press, Cambridge, Mass. - London, England, 1994, *A most comprehensive textbook.*

[12] J. Rosenmüller, *Game theory: Stochastics, information, strategies and cooperation*, Theory and Decision Library, C, vol. 25, Kluwer Academic Publishers Boston, Dordrecht, London, 2000.

[13] L. Samuelson, *Evolutionary games and equilibrium selection* , MIT Press, Cambridge, 1997, *A careful analysis of evolutionary game theory and its use for equilibrium selection in games* .

[14] L.S. Shapley, *Utility comparisons and the theory of games* , La Decision: Aggregation et Dynamique des Ordres de Preference (Paris), Editions du Centre National de la Recherche Scientifique, 1969, pp. 251 – 263, *The basic paper extending the Shapley value to NTU–games* .

[15] W. Vickrey, *Auctions and bidding games*, Recent Advances in Game Theory, Princeton Conference Series, vol. 29, Princeton University, 1962, pp. 15 – 27, *This is the basic article establishing the impact of Game Theory on the Theory of Auctions and Bidding* .

[16] J. von Neumann and O. Morgenstern , *Theory of games and economic behavior*, Princeton University Press, Princeton, 1944,1947,1953, *The basic volume containing the foundation of Game Theory* .