INSTITUTE OF MATHEMATICAL ECONOMICS

WORKING PAPERS

No. 301

Implementation of the Kalai-Smorodinski Bargaining Solution in Dominant Strategies

by

Claus-Jochen Haake

October 1998

University of Bielefeld

33501 Bielefeld, Germany

Implemention of the Kalai–Smorodinski Bargaining Solution in Dominant Strategies

Claus-Jochen Haake*

October 1998

Abstract

In the existing literature the Kalai–Smorodinsky bargaining solution is implemented either by using the Nash equilibrium or the subgame–perfect equilibrium concept. In this paper we provide a setup for implementing bargaining solutions and construct a strategic mechanism for n players that implements the Kalai–Smorodinsky bargaining solution in dominant strategies. Moreover we have uniqueness of dominant strategy equilibria in each of the induced games. From this mechanism we can derive an extensive game form such that the final outcome in the unique subgame–perfect equilibrium again coincides with the Kalai–Smorodinsky bargaining solution. So we get both from the original mechanism — dominant strategy equilibrium and also subgame–perfect implementation.

Keywords: Bargaining, Implementation, Kalai–Smorodinsky Solution

*ACKNOWLEDGEMENT: Helpful comments and encouragement from Prof. W. Trockel are gratefully acknowledged.

0 Introduction

One application of implementation theory is to give a non–cooperative foundation for cooperative solution concepts. In any possible situation (e.g. for any possible set of players) we want to obtain the cooperative solution by playing a non–cooperative game. The rules that determine these games are given by a mechansim or game form. The "type of implementation" then depends on the solution concept we apply to the induced non–cooperative games. So *implementation in dominant strategies* requires that all games that are induced by the mechanism possess an equilibrium in dominant strategies and the outcome in this equilibrium will be the given cooperative solution.

A special class of cooperative solutions is of course the class of bargaining solutions. In this paper we want to concentrate on one special bargaining solution, which was introduced by Kalai & Smorodinsky (1975). Their solution for the bargaining problem is axiomatized by some monotonicity axiom and incorporates fairness by allocating equal proportions of blisspoint payoffs.

In the early fifties efforts started to give a non–cooperative foundation of axiomatic bargaining solutions, known as the "Nash–program". Nash (1953) himself was the first to make a contribution to this topic by discussing his "simple demand game". He obtains his bargaining solution by a Nash equilibrium of this game. However, the problem is that this game has multiple equilibria and one could get any Pareto–efficient allocation of utility by an appropriate Nash equilibrium.

For the Kalai–Smorodinsky bargaining solution (or KS solution) we only have few references concerning a non–cooperative foundation. In the existing literature the KS solution is either implemented by a subgame–perfect equilibrium or by a Nash equilibrium.

Crawford (1978) suggests a procedure to achieve Pareto–efficient egalitarian–equivalent allocations in a two–player exchange economy by combining an auction with a sort of *Divide and Choose* mechanism. Those allocations yield the Kalai–Smorodinsky solution in the referring utility space. Crawford's procedure was refined and extended to an n–player setup by Demange (1984). Both authors derive games in extensive form, where the first stage consists of an auction that determines one divider and an order of the choosers. Stage 2 is left for proposing a division (by the divider) and consecutively agreeing or rejecting by the choosers. With this mechanism the KS solution is implemented by a subgame–perfect equilibrium (SPE), meaning that the outcome in the SPE of the game yields the Kalai–Smorodinsky solution. To observe this result, both authors have

to impose strong assumptions on an individual's behaviour when he or she is indifferent. A similar result in a welfaristic bargaining approach is obtained by Moulin (1984). There the KS bargaining solution is implemented by a subgame–perfect equilibrium, using the "Game of Auctioning Fractions of Dictatorship". Stage 0 of this game is again a sealed bid auction that determines a ranking of the players. Then the first player has to propose a utility allocation that can be accepted or rejected by the other participants. The first one, who rejects this proposal (if it is rejected at all), is then allowed to behave like a dictator and ensure himself a utility that depends on the winning bid of the auction.
This game in extensive form possesses a subgame–perfect equilibrium, whose outcome is the Kalai–Smorodinsky bargaining solution. To obtain this result, one has to claim that a player has to accept a proposal, even if there is no incentive for him to do so. Without this behavioural restriction, one gets additional equilibria that ensure the KS utility to at most one of the participants.
Van Damme (1991) briefly discusses two mechanisms for two players, which Nash–implement the KS bargaining solution. The idea behind the derived one–shot games is to trade off own interest and probability that one's proposal will be enforced. Both games have a unique Nash equilibrium that yields KS utilities.

Trockel (1998) presents a mechanism that implements the Nash bargaining solution in dominant strategies. Since the concept of "equilibrium in dominant strategies" is stronger than that of a Nash equilibrium, he gets a stronger notion of implementation. For this he modifies Nash's original simple demand game. It turns out that he in fact suggests a method we can use to implement other bargaining solutions as well.

In the present paper we provide a setup (or implementation context), in which we will implement the KS bargaining solution in dominant strategies. For this we modify Nash's simple demand game in a different way. That means that our mechanism induces non–cooperative one–shot games in strategic form that have exactly one equilibrium in dominant strategies and the outcome in this equilibrium coincides with the KS bargaining solution. So individual payoff maximization will be the point that leads us to this result. With a slight modification, we extend the mechanism to the case, where n players are involved. Even in this case we have existence and uniqueness of a dominant strategy equilibrium. The rough idea behind is to trade off "modesty", which is connected with low gains from cooperation, and punishment of "unmodest behaviour".
This strategic mechanism also gives rise to the construction of games in extensive form, implementing the KS solution by a (unique) subgame–perfect equilibrium. In contrast to

the approaches discussed above, here we do not have to start with an auction. because a special order of the players is not necessary. Moreover, we need not impose any restrictions on players' "indifference behaviour" to achieve this result. And since we have uniqueness of equilibria, there is no coordination or choice problem for the players.

To summarize, this paper bears two main differences to the existing literature. First, we get a stronger notion of implementation than Nash implementation, namely implementation in dominant strategies. And second, we get subgame–perfect implementation without having to determine a special order of the players and without imposing restrictions on individuals' "indifference behaviour".

The organization of the paper is as follows: After a short introduction to bargaining and implementation theory, our mechanism is discussed in section 2. The extension to the n–player case is presented in Section 3. Subgame–perfect implementation is then derived in section 4. Section 5 concludes.

1 The Framework

This section roughly falls into two parts. The first part provides some basic definitions from (axiomatic) bargaining theory. In the desciption of a bargaining problem we focus on the welfaristic problem. That means that all analysis takes place in utility space. Having some von Neumann–Morgenstern utilities over lotteries in mind, we do not state explicitly where these utilities come from.

Part two deals with the concepts of implementation theory in general. This is mainly done by introducing the notion of a context. A context describes a setup, in which a social planner proposes the design for a mechanism, without knowing how the "real" situation looks like. Both worlds will then be linked in the next section by defining the so called *bargaining context.*

To introduce the basic framework we concentrate on the two–player case, but all definitions and concepts can easily be carried over to an n–player setup. This will be briefly discussed at the beginning of section three.

1.1 Bargaining Problems

A **2–person bargaining problem** is a pair (S, d) such that S is a closed, comprehensive and convex subset of $I\!R^2$, called **the set of feasible utility allocations**. The point $d \in S$ is called **disagreement point**. In addition we assume that the set of individual rational utility allocations $\{x \in S \,|\, x \geq d\}$ is bounded and that there exists an element $\bar{x} \in S$ satisfying $x >> d$. Let $\mathcal{B}^d$ denote the set of all 2–person bargaining problems with disagreement point d.

We think of S as the set of utility allocations arising from some **cardinal** utility functions over an abstract set of states of nature (e.g. allocations in an economy). Therefore we can assume w.l.o.g. that the disagreement point is zero.

Denote by $\mathcal{B}$ the set of all 2-person bargaining problems with disagreement point 0, i.e. $\mathcal{B} := \mathcal{B}^0$ and identify the bargaining problem $(S, 0) \in \mathcal{B}$ with the utility set S itself. No confusion will result.

A **bargaining solution** is a mapping $\varphi : \mathcal{B} \to I\!R^2$ satisfying the feasibility assumption $\varphi(S) \in S$.

We omit further "reasonable" properties of bargaining solutions like *indvidual rationality*, *efficiency* or *covariance under affine linear transformations*, because we do not want to characterize a bargaining solution by a set of axioms.

Next we define a special bargaining solution introduced by Kalai & Smorodinsky (1975). For a given bargaining problem $S \in \mathcal{B}$ we define

$$\begin{aligned} \bar{\beta}_i(S) &:= \max\left\{t \in I\!R \,|\, t\, e^i \in S\right\} \quad (i = 1, 2) \\ \bar{\beta}(S) &:= \bar{\beta}(S)_1\, e^1 + \bar{\beta}(S)_2\, e^2 \end{aligned}$$

$\bar{\beta}_i(S)$ is called **blisslevel of player** i. Due to comprehensiveness of each utility set S, $\bar{\beta}_i(S)$ gives the maximal utility player i can achieve, while guaranteeing at least a utility of zero to the other player. $\bar{\beta}(S)$ is called **blisspoint**. Note that (for notational convenience) $\bar{\beta}_i(S)$ is a real number whereas $\bar{\beta}(S)$ describes a point in $I\!R^2$.

Denote by $l(S)$ the largest fraction of the blisspoint that is feasible in S, i.e.

$$l(S) := \max\left\{t \in I\!R \,|\, t\, \bar{\beta}(S) \in S\right\}.$$

Then the **Kalai–Smorodinsky (KS) bargaining solution** is the mapping $\kappa : \mathcal{B} \to I\!R^2$, satisfying

$$\kappa(S) := l(S)\, \bar{\beta}(S).$$

The idea of fairness behind this bargaining solution is that both players should get the same proportion of their individual blisslevels. This idea of a proportional allocation ensures envy–freeness in the sense that all players get the same percentage of their "maximal utility". Moreover the solution is Pareto–efficient, since this percentage is chosen maximal.

One can identify absolute utilities arising from application of a bargaining solution with the referring fractions of blisslevels. In the sequel we consider bargaining solutions as functions assigning fractions rather than utilities. Therefore we define the **fractional KS bargaining solution** $\kappa^f : \mathcal{B} \to [0,1] \times [0,1]$ by

$$\kappa^f(S) := (l(S), l(S)).$$

This completes the introduction to two person bargaining problems.

1.2 Mechanisms and Implementation

The mechanism design problem treats the question how to achieve a socially desired outcome. This should be done by offering incentives to the individuals such that they behave accordingly. To explain what the term "socially desired" means, we can think of a social planner (or society itself), who selects for each possible state of nature (characterized by individual preferences) one or more outcomes. Formally this is reflected in a *social choice rule* (see formal definition below). *Implementation* of a social choice rule means that the planner "sets up rules for a (non–cooperative) game" such that application of a certain solution concept yields the desired outcome.

In order to formalize this problem, we have to clarify the following questions:

1. In which setup (or context) does the planner operate?
2. Which social choice rule should be implemented?
3. Which solution concept should yield the socially desired outcome?
4. How does the referring mechanism look like?

The remainder of this section treats these questions in general. Section two then describes a special context, linking bargaining problems with implementation theory.

A **context** describes an environment in which a mechanism designer can operate. This includes the number of possible players, the set of possible outcomes, a set of possible preference profiles over outcomes and finally a class of mechanisms from which the planner will choose one. Note that the description of a context and especially the description of a mechanism does **not** include any specific characteristics of players.

Following basically the notation in Osborne & Rubinstein (1996) we formally describe a context by the tuple $(N, A, \mathcal{P}, \mathcal{G})$, where

- $N := \{1, 2\}$ denotes the set of *player positions* (for two players),
- A denotes the outcome space,
- $\mathcal{P}$ denotes a set of *preference profiles* over A,
- $\mathcal{G}$ denotes a set of (strategic) mechanisms with outcomes in A (see formal definition below).

Each preference profile represents a set of "real" players. And whenever "real" players enter the scene they have a certain set of socially desired outcomes in mind. This is formalized by a social choice rule.
A **social chioce rule** is a correspondence $\alpha : \mathcal{P} \Longrightarrow A$, assigning a set of outcomes to each preference profile.
Implementing a social choice rule means that we have to construct rules for a non–cooperative game without taking players' preferences into account. This leads us to the definition of a mechanism.
A **strategic mechanism** (or strategic game form) is a tuple $G = (N, \Sigma, g)$ consisting of the set N of player positions, a set of *strategy profiles* $\Sigma := \Sigma_1 \times \Sigma_2$ and an **outcome function** $g : \Sigma \to A$.

Note that a strategic mechanism together with a representable preference profile $\succeq = (\succeq_1, \succeq_2) \in \mathcal{P}$ defines a non–cooperative game by composing the outcome function g with (cardinal) utility functions u_i representing $\succeq_i$[1]. So when "real" players enter the scene, then a mechanism suggests to play a non–cooperative game depending on these players' preferences over A. Hence the notion of a *solution concept for a context* should refer to these induced games. Formally a **solution concept** for the context $(N, A, \mathcal{P}, \mathcal{G})$ is a correspondence $\Phi : \mathcal{G} \times \mathcal{P} \Longrightarrow \Sigma$, assigning each induced game a set of strategy profiles in Σ. For example the *Nash equilibrium concept* would consider the game induced by a pair $(G, \succeq) \in \mathcal{G} \times \mathcal{P}$ and assign the set of Nash equilibria of this game to $(G, \succeq)$.

[1]The term "strategic" refers to the fact that the derived games are games in strategic form. Although implementation does not only deal with games in strategic form, we only want to formalize this direction.

Now let Φ be a solution concept for the context $(N, A, \mathcal{P}, \mathcal{G})$. We say that the strategic mechanism $G = (N, S, g) \in \mathcal{G}$ **Φ–implements** the social choice rule $\alpha : \mathcal{P} \Longrightarrow A$, if for each preference profile $\succeq = (\succeq_i)_{i \in N} \in \mathcal{P}$ the implementation condition

$$g(\Phi(G, \succeq)) \subseteq \alpha(\succeq)$$

is satisfied. In this case we call α **Φ–implementable**.

We can think of an "implementation procedure" as follows: First a social planner has one social choice rule in mind, i.e. a rule that determines a set of (socially desired) outcomes for each possible preference profile. Then he designs a mechanism $G \in \mathcal{G}$. This mechanism together with any specific preference profile in $\mathcal{P}$ defines a (non–cooperative) game. For these induced games the planner thinks of a specific solution concept. This solution concept yields strategies for any of the induced games, i.e. for any preference profile that is combined with G. So for a specific preference profile, we get a set of strategies (in the induced game). Now applying the outcome function g of the mechanism to any of these strategies gives us an outcome that belongs to the set of outcomes determined by the social choice rule for this preference profile.
So actually a mechanism is a collection of rules for a game and again does **not** depend on any characteristics of "real" players.

2 Implementation in Dominant Strategies

After a short introduction to axiomatic bargaining and the implementation problem, we now want to link both fields by setting up the so–called *bargaining context.* For this context we then define the *Kalai–Smorodinsky choice rule* (or KS choice rule), which will of course be strongly related to the referring bargaining solution. After that we just state what is meant by implementation in dominant strategies (DSE–implementation).
The second part of this section contains some preliminary work for the construction of a mechanism that DSE–implements the KS choice rule. And finally the third part provides the construction of the mechanism and proves the implementation conditions.

As we have seen in the first section, a mechanism together with "real players", who are represented by their preference profile, induces a non–cooperative one–shot game in strategic form. So, to give a first intuition how the mechanism works, we briefly describe these induced games. As we will see, a preference profile will be identified with a specific bargaining problem, say $S \in \mathcal{B}$. How does the game that is induced by S look like? Both players are asked to announce a number in the unit interval. The payoff function now makes use of the underlying bargaining problem. Let $x_1 \in [0,1]$ be the announcement from player 1 and consider all points in S that give player 2 a utility of **at least** x_1 multiplyed with his blisslevel $\bar{\beta}_2(S)$. Among all these points in S, we determine the largest possible utility for player 1. And, roughly speaking, the minimum of this and $x_1 \bar{\beta}_1(S)$ (1's announced fraction) will be player 1's payoff in the game[2].

The idea behind this is to punish "bold behaviour". That means, for a "high announcement x_1" the set of allocations giving this fraction to player 2 will be small and as a consequence player 1's final payoff will be low. So player 1 may improve by choosing a lower announcement. Conversely for a "low announcement x_1" this fraction will essentially be his payoff. So he may improve by increasing x_1. And this will exactly be the point that leads us to the Kalai–Smorodinsky bargaining solution.

The incentive for choosing "KS strategies" is that they constitute an equilibrium in dominant strategies in every induced game.

2.1 The bargaining context

Again we only consider the two–player case but the extension to an n–player context is straightforward. This is briefly discussed in section 3.

Let $N := \{1,2\}$ be the set of player positions and let the outcome space A be the set of all mappings on $\mathcal{B}$ with values in $[0,1]\times[0,1]$, i.e. $A := ([0,1]\times[0,1])^{\mathcal{B}}$. For a given $L \in A$ and $x = (x_1, x_2) := L(S)$ we think of x_i as the fraction of player i's blisslevel $\bar{\beta}_i(S)$ that he will receive. Then for a given bargaining problem S, the outcome $L \in A$ will assign a total utility of $(L(S))_i \bar{\beta}_i(S)$ to player i.

To avoid some impossibility result (cf. Gibbard–Satterthwaite theorem) we want to restrict the set of all preference profiles over A to those that are appropriate for this situation. Since we assume that the bargaining problem two "real" players face is common knowledge, only those preferences that are defined exclusively via this bargaining problem seem

[2]In the formal description we will rather deal with fractions of blisslevels than with absolute utilities. So his payoff will be the referring fraction of $\bar{\beta}_1(S)$

to matter. Let us therefore define for $S \in \mathcal{B}$ a preference $\succeq_i^S$ on A for player i by

$$L \succeq_i^S L' \iff (L(S))_i \geq (L'(S))_i \quad (L, L' \in A, \quad i = 1, 2),$$

and set $\mathcal{P} := \left\{\succeq^S = (\succeq_1^S, \succeq_2^S) \mid S \in \mathcal{B}\right\}$.

Note that we can (and will) represent a preference $\succeq_i^S$ by the utility function $u_i^S : A \to I\!R$ satisfying $u_i^S(L) := (L(S))_i$. So each player is only interested in his own coordinate. Because of the special structure of $\mathcal{P}$, we can identify $\mathcal{P}$ with the set of bargaining problems $\mathcal{B}$ itself.

The set of strategic mechanisms in this context is the set of tuples $(N, \Sigma_1 \times \Sigma_2, g)$ with strategy set $\Sigma_1 \times \Sigma_2$ and outcome function $g : \Sigma_1 \times \Sigma_2 \to A$. Here we take the unit interval as strategy set for both players, i.e. $\Sigma_i := [0,1] \quad (i = 1, 2)$. Again combining a mechanism G with a prefernce profile (u_1^S, u_2^S) (i.e. choosing an element of $\mathcal{G} \times \mathcal{B}$) we get a non–cooperative game in strategic form $\Gamma^S := (\Sigma_1, \Sigma_2, v_1^S, v_2^S)$ with strategy sets $\Sigma_i = [0,1]$ and payoff functions $v_i := u_i^S \circ g$.

Hence a solution concept for this context is a mapping $\Phi : \mathcal{G} \times \mathcal{B} \to [0,1] \times [0,1]$.

The social choice rule that we want to implement is the **Kalai–Smorodinsky choice rule**. It says that whenever two players face a bargaining problem it is (socially) desired to solve the conflict applying the KS bargaining solution κ to this problem. Note that we do not want to say what is desired in bargaining problems that do not occur. Hence, we define a correspondence

$$\begin{aligned} \mathcal{K} &: \mathcal{B} \equiv \mathcal{P} \to A = ([0,1] \times [0,1])^{\mathcal{B}}, \\ \mathcal{K}(S) &:= \left\{L \in A \mid L(S) = \kappa^f(S) = (l(S), l(S))\right\}, \end{aligned}$$

called **KS choice rule**. Remember that $l(S)$ denotes the "KS fraction" occuring in the definition of the KS bargaining solution.

2.2 The concession function

We can describe a mechanism by describing all games that occur when players appear. Since players are characterized by their preferences and preference profiles are given by bargaining problems, we consider the game that is induced by the mechanism and a specific bargaining problem. Actually this will result in a pointwise definition of the outcome function.

Now, given a bargaining problem $S \in \mathcal{B}$, both agents are asked to announce simultaneously a number in the unit interval. Facing the announcement $x_1 \in [0,1]$ by player 1, there are two possible cases, depending on whether the point $(x_1\bar{\beta}_1(S), x_1\bar{\beta}_2(S))$ is feasible or not, i.e. if there is a utility allocation in S that gives to **both** players a fraction of x_1 of their blisslevels. If the announcement is "humble enough" and this is possible, then player 1's payoff should be his announced fraction x_1 of his blisslevel. In the other case the payoff function will assign player 1 the largest possible utility considering all utility allocations in S that give player 2 **at least** a utility of $x_1\bar{\beta}_2(S)$.

Of course the same should hold for the payoffs for player 2.

We will say that in the case, where an announcement x_i is "too bold", player i has to *make a concession*, although it is in fact the payoff function that enforces a lower payoff than $x_i\bar{\beta}_i(S)$.

Formally we describe this idea by introducing functions $z_i^S : [0,1] \to [0,1]$ called *minimal concession functions*. If the announcement x_1 is "humble enough" (in the sense described above) then z_1^S returns this fraction but otherwise it returns the largest possible fraction while ensuring x_1 to player 2. In that case player 1 has to make the lowest possible concession, which is given by $x_1 - z_1^S(x_1)$.

Figure 1 illustrates the construction of player 1's minimal concession function in the case $\bar{\beta}_1(S) = \bar{\beta}_2(S) = 1$. Here we can identify fractions of blisslevels with absolute utility.

This leads us to the formal definition of the minimal concession functions. For given $S \in \mathcal{B}$ we define functions $z_i^S : [0,1] \to [0,1]$ $(i = 1,2)$ by

$$z_1^S(t) := \begin{cases} \max\left\{y \in [0,1] \mid (y\,\bar{\beta}_1(S), t\,\bar{\beta}_2(S)) \in S\right\} & , \quad (t\,\bar{\beta}_1(S), t\,\bar{\beta}_2(S)) \notin S \\ t & , \quad (t\,\bar{\beta}_1(S), t\,\bar{\beta}_2(S)) \in S \end{cases}$$

$$z_2^S(t) := \begin{cases} \max\left\{y \in [0,1] \mid (t\,\bar{\beta}_1(S), y\,\bar{\beta}_2(S)) \in S\right\} & , \quad (t\,\bar{\beta}_1(S), t\,\bar{\beta}_2(S)) \notin S \\ t & , \quad (t\,\bar{\beta}_1(S), t\,\bar{\beta}_2(S)) \in S. \end{cases}$$

Observe that a fixed fraction $x_i \in [0,1]$ induces a mapping $z_i^{(\cdot)}(x_i) : \mathcal{B} \to [0,1]$.

Figure 2 shows an example of a concession function. But since utility allocation sets are supposed to be convex every concession function has a similar shape, i.e. it is concave, there exists a unique maximizer $\bar{x}$ and it coincides with the identity on the interval $[0, \bar{x}]$.

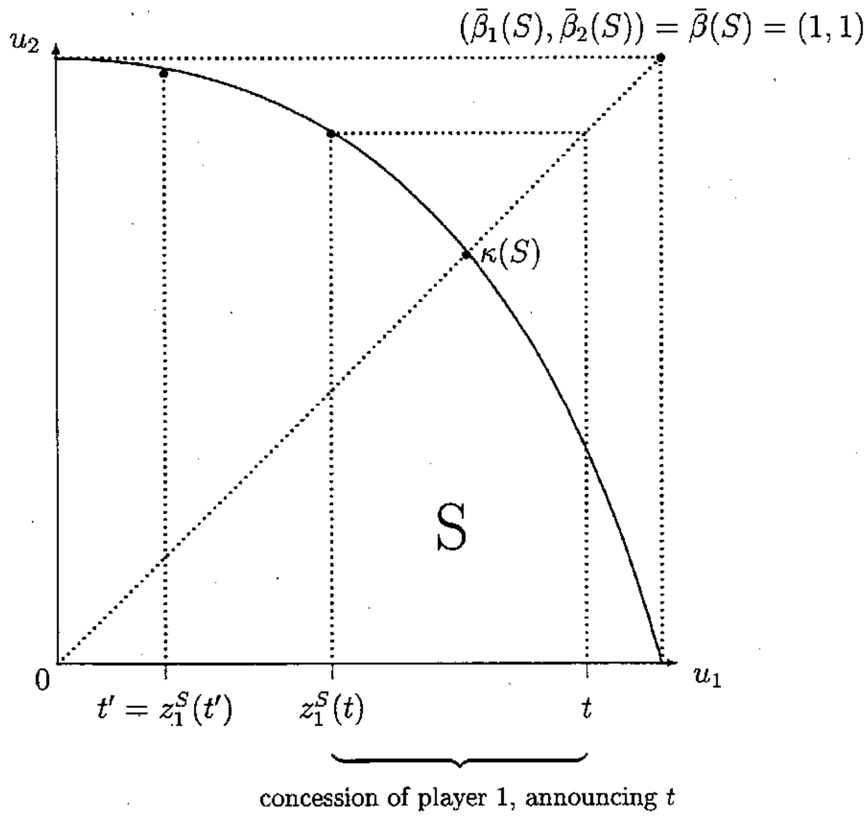

Figure 1: Constructiuon of z_1^S in the two player case

2.3 The mechanism for two players

With the aid of minimal concession functions we are now able to construct a strategic mechanism $G^* = (N, \Sigma_1 \times \Sigma_2, g^*) \in \mathcal{G}$ with outcome function $g^* : \Sigma_1 \times \Sigma_2 \to A$.

We take $\Sigma_i := [0,1]$ as the **strategy set** for player i. Note that the strategy sets are independent of an underlying preference profile, i.e. a specific bargaining problem.

Define the **outcome function** $g^* : [0,1] \times [0,1] \to A$ by

$$\begin{array}{rcl} (x_1, x_2) & \mapsto & g^*(x_1, x_2) \quad \in \quad A \\ g^*(x_1, x_2)(\cdot) & := & \left(z_1^{(\cdot)}(x_1), z_2^{(\cdot)}(x_2) \right). \end{array}$$

Because of the construction of the concession functions $g^*(x_1, x_2)$ is a function on $\mathcal{B}$ that yields *feasible* utility allocations for any bargaining problem $S \in \mathcal{B}$.

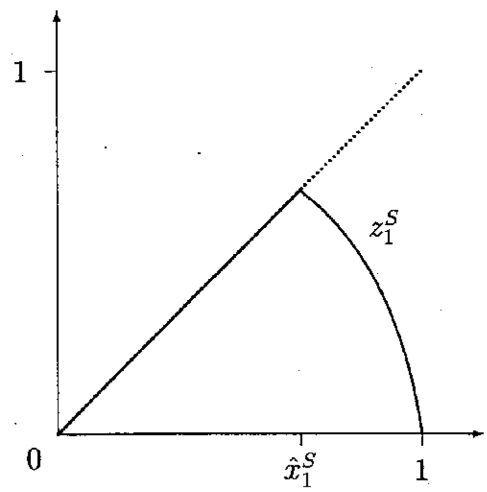

Figure 2: The concession function for player 1

Since we can identify preference profiles with bargaining problems (see definition of the bargaining context above), each $S \in \mathcal{B}$ together with the mechanism G^* induces a non-cooperative game in strategic form $\Gamma^S := ([0,1],[0,1],v_1^S,v_2^S)$ with payoff functions

$$v_i^S(x_1,x_2) := u_i^S(g^*(x_1,x_2)) = (g^*(x_1,x_2)(S))_i .$$

In the sequel we will analyze the structure of these induced games and show that each game has a *unique equilibrium in dominant strategies.* To prove this theorem we have to analyze the structure of the concession function.

Lemma 2.1 *For any bargaining problem $S \in \mathcal{B}$ there exists a unique maximizer of the minimal concession functions z_i^S $(i = 1,2)$.*

Proof: We show that for any $S \in \mathcal{B}$ there exists $\hat{x}^S \in [0,1]$ such that $z_1^S(\hat{x}^S) > z_1^S(x)$ holds for all $x \in [0,1]$, $x \neq \hat{x}^S$.
Let $S \in \mathcal{B}$ be given. Define

$$\hat{x}^S := \max\left\{x \in [0,1] \mid (x\,\bar{\beta}_1(S), x\,\bar{\beta}_2(S)) \in S\right\}. \tag{1}$$

Since S is closed the maximum in (1) is attained and by definition of z_1^S we have $z_1^S(\hat{x}^S) = \hat{x}^S$. For $x \in [0,1]$ consider two cases:

Case 1: $x < \hat{x}^S$
By comprehensiveness of S we have $(x\,\bar{\beta}_1(S), x\,\bar{\beta}_2(S)) \in S$, which implies $z_1^S(x) = x$, hence

$$z_1^S(x) = x < \hat{x}^S = z_1^S(x^S).$$

Case 2: $x > \hat{x}^S$
By definition of $\hat{x}^S$ we have $(x\,\bar{\beta}_1(S), x\,\bar{\beta}_2(S)) \notin S$, implying

$$z_1^S(x) = \max\left\{y \in [0,1] \,|\, (y\,\bar{\beta}_1(S), x\,\bar{\beta}_2(S)) \in S\right\} =: \bar{y}^S(x).$$

Note that $\bar{y}^S(x) < x$ (S is closed).
Assume $z_1^S(x) \geq z_1^S(\hat{x}^S)$, which is equivalent to $\bar{y}^S(x) \geq \hat{x}^S$. Since S is convex and of course $(\bar{\beta}_1(S), 0) \in S$, we have that

$$\lambda\,(\bar{y}^S(x)\,\bar{\beta}_1(S), x\,\bar{\beta}_2(S)) + (1-\lambda)\,(\bar{\beta}_1(S), 0) \in S$$

holds for all $0 \leq \lambda \leq 1$. For $\bar{\lambda} := \frac{1}{x - \bar{y}^S(x) + 1}$ we obtain

$$(\frac{x}{x - \bar{y}^S(x) + 1}\,\bar{\beta}_1(S), \frac{x}{x - \bar{y}^S(x) + 1}\,\bar{\beta}_2(S)) \in S.$$

which contradicts the maximality of $\hat{x}^S$, because $\frac{x}{x - \bar{y}^S(x) + 1}$ is obtained as a convex combination of $\bar{y}^S(x)$ and 1 and therefore is certainly larger than $\hat{x}^S$.
Hence we also get $z_1^S(x) < \hat{x}^S = z_1^S(\hat{x}^S)$, which completes the proof of the lemma. □

Note that by construction of the maximizer $\hat{x}^S$ of a concession function z_i^S it coincides with the referring KS fraction l(S).
The next theorem shows that these maximizers constitute an equilibrium in dominant strategies of the induced game.

Theorem 2.2 *For any bargaining problem $S \in \mathcal{B}$ the induced game Γ^S has a unique equilibrium in dominant strategies.*

Proof: Let $S \in \mathcal{B}$ be given. Define $\hat{x}_1^S = \hat{x}_2^S := \hat{x}^S$ as in the proof of lemma 2.1. We show that $(\hat{x}_1^S, \hat{x}_2^S) \in [0,1] \times [0,1]$ is a dominant strategy equilibrium of the induced game $\Gamma^S = ([0,1],[0,1], v_1^S, v_2^S)$. Let $x_1 \neq \hat{x}_1^S$ and $x_2 \in [0,1]$ be arbitrary. Then we have

$$v_1^S(x_1, x_2) \;\; = \;\; u_1^S(g^*(x_1, x_2))$$

$$
\begin{aligned}
&= \left(z_1^S(x_1), z_2^S(x_2)\right)_1 \\
&= z_1^S(x_1) \\
(2) \qquad \text{(by Lemma 2.1)} &< z_1^S(\hat{x}_1^S) \\
&= \left(z_1^S(\hat{x}_1^S), z_2^S(x_2)\right)_1 \\
&= u_1^S(g^*(\hat{x}_1^S, x_2)) \\
&= v_1^S(\hat{x}_1^S, x_2).
\end{aligned}
$$

Therefore $\hat{x}_1^S$ is a dominant strategy in Γ^S for player 1. An analogue computation shows that $\hat{x}_2^S$ is a dominant strategy for player 2. Since we have strict inequality in (2) the strategy profile $(\hat{x}_1^S, \hat{x}_2^S)$ is the unique equilibrium of Γ^S in dominant strategies. This completes the proof of the theorem. □

For $S \in \mathcal{B}$ let $DSE(\Gamma^S) := \left\{(\hat{x}_1^S, \hat{x}_2^S)\right\}$ denote the set consisting of the unique dominant strategy equilibrium of Γ^S. We now show that the mechanism G^* DSE–implements the KS choice rule. Therefore we have to show that for any bargaining problem $S \in \mathcal{B}$ the outcome function g^* applied to the dominant strategy equilibrium of Γ^S yields the KS bargaining solution of S.

Theorem 2.3 (Implementation) *For each $S \in \mathcal{B}$ the "implementation conditions"*

$$
\begin{aligned}
g^*(DSE(\Gamma^S)) &\subseteq \mathcal{K}(S) \\
u^S(g^*(DSE(\Gamma^S))) &= u^S(\mathcal{K}(S)) = \kappa^f(S)
\end{aligned}
$$

are satisfied[3].

Proof: Let $S \in \mathcal{B}$ be given. To show that applying the outcome function g^* to the dominant strategy equilibrium $(\hat{x}_1^S, \hat{x}_2^S)$ of Γ^S yields a socially desired outcome (in the sense of the KS choice rule), we have to check the equation

$$
(3) \qquad \left(g^*(\hat{x}_1^S, \hat{x}_2^S)\right)(S) = \kappa^f(S).
$$

Then $g^*(\hat{x}_1^S, \hat{x}_2^S)$ is a bargaining solution (an element of A) that yields KS utilities for the bargaining problem S and therefore belongs to $\mathcal{K}(S)$. Equation (3) holds, since

$$
\left(g^*(\hat{x}_1^S, \hat{x}_2^S)\right)(S) = (z_1^S(\hat{x}_1^S), z_2^S(\hat{x}_2^S))
$$

[3]There is a slight abuse in formal correctness in the second condition. The expression on the left side is in fact a set that consists of just one element. This element, representing the utility from playing the dominant strategy equilibrium, is equal to the utility arising from the KS bargaining solution.

$$\begin{aligned} &= (\hat{x}_1^S, \hat{x}_2^S) \\ &= (l(S), l(S)) = \kappa^f(S). \end{aligned}$$

Hence $g^*(DSE(\Gamma^S)) \subseteq \mathcal{K}(S)$. The second part of the theorem follows easily, since $u^S(g^*(DSE(\Gamma^S))) = \left(g^*(\hat{x}_1^S, \hat{x}_2^S)\right)(S) = \kappa^f(S)$.[4] This completes the proof of the theorem. □

To close this section, note that for the pointwise definition[5] of the outcome function we make use of knowledge about the bargaining problem but we do **not** have to know how our problem S looks like to describe the general procedure of the mechanism.

3 Extension to the n–Player Case

This section is devoted to an extension of the mechanism for two players. Extending the framework to an n–player context will be straightforward. This is done in the first part. The mechanism for two players crucially bases on considerations of minimal concessions, which leads to the definition of two referring functions. The extended mechanism will again make use of such functions, but since there is more than one opponent involved, these functions have to be slightly modified, without changing the whole story behind it. This modification is discussed in part two. The last part presents the extended mechanism and proves that it really works and again yields the same outcomes as the Kalai–Smorodinsky bargaining solution.

3.1 The n–player framework

Here we do not have to do a lot of work, because all definitions from bargaining and implementation theory in general easily carry over to an n–player setup. Thereby we adjust the notations from the last section, e.g. $\mathcal{B}$ now denotes the set of all n–person bargaining problems with disagreement point 0. We just want to describe briefly the bargaining context for n players.

A context is described by a tuple $(N, A, \mathcal{P}, \mathcal{G})$ (cf. section 2). Here we take $N = \{1, \ldots, n\}$ as the set of player positions. The outcome space A consists of all mappings on $\mathcal{B}$ with

[4]Remember that the utility function u^S applied to some element of $\mathcal{K}(S)$ yields $\kappa^f(S)$, since u^S only considers what happens at S and there all solutions in $\mathcal{K}(S)$ coincide with the fractional KS solution.

[5]i.e. pointwise definition of the function $g^*(x_1, x_2) : \mathcal{B} \to A$ for fixed $(x_1, x_2) \in [0,1] \times [0,1]$.

values in $[0,1]^n$, i.e. $A := \{L : \mathcal{B} \to [0,1]^n\}$. The set of possible preference profiles is again described by means of preference relations $\succeq_i^S$ (cf. section 2.1). Therefore we get $\mathcal{P} := \left\{(\succeq_1^S, \ldots, \succeq_n^S) \mid S \in \mathcal{B}\right\}$. So again we can identify $\mathcal{P}$ with the set $\mathcal{B}$ of n–person bargaining problems itself. To complete the description of the context the set of strategic mechanisms $\mathcal{G}$ consists of triplets (N, Σ, g), where $\Sigma := \Sigma_1 \times \ldots \times \Sigma_n$ is the set of strategy profiles and $g : \Sigma \to A$ denotes the outcome function.

To close this subsection we define the KS choice rule for n players. What is socially desired depends on an underlying preference profile (i.e. "real players"), hence on the underlying bargaining problem. So we just claim that players should agree on the KS outcome in the underlying problem and put no further restrictions on what happens in other situations. This leads us to the definition of the KS choice rule for n players

$$\mathcal{K}(S) := \left\{L \in A \mid L(S) = \kappa^f(S) = (l(S), \ldots, l(S))\right\}.$$

$l(S)$ denotes the (common) fraction of blisslevels that each player gets at the KS solution. (cf. definition of th fractional KS solution at the end of section 1.)

3.2 Minimal concessions

The idea of the mechanism when two players are involved is to guarantee a certain amount of utility to the opponent. This amount is determined by one's own announcement and guaranteeing this amount eventually leads to concessions one has to make[6]. To let this idea work, we exploited the special structure of two person bargaining problems. For **any** announcement x_1 of player 1 it is possible to find a point in S that ensures a fraction of x_1 of 2's blisslevel to player 2. Consider now a three person bargaining problem. If player 1's announcement x_1 is too bold there might be no possibility to assure this to players 2 and 3 simultaneously. In this case we will set the value of his concession function to 0. Otherwise we can define concession functions z_i^S in the usual way. Figure 3 illustrates the construction of the minimal concession function for player 1 when 3 players are involved.

Let $S \in \mathcal{B}$ be a given bargaining problem. For $i \in N$ define the minimal concession

[6]Again we use the term "having to make a concession" although actually not players "have to concede" but the payoff function returns a smaller value.

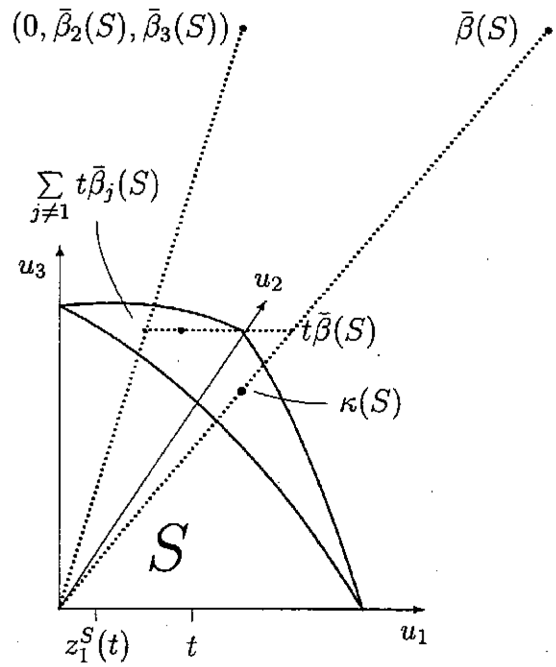

Figure 3: Concession function for player 1 with three players involved

function $z_i^S : [0,1] \to [0,1]$ by

$$z_i^S(t) := \begin{cases} \max\left\{0, \max\left\{y \in [0,1] \,|\, y\,\bar{\beta}_i(S)\,e^i + \sum_{j \neq i} t\,\bar{\beta}_j(S)\,e^j \in S\right\}\right\} & , \; t\,\bar{\beta}(S) \notin S \\ t & , \; t\,\bar{\beta}(S) \in S \end{cases}$$

with the convention $\max \emptyset = -\infty$. Remember that $\bar{\beta}(S) \in I\!R^n$ denotes the blisspoint.

Due to convexity of utility possibility sets, each concession function has a specific shape that is depicted in figure 4. Roughly speaking we can partition player i's strategy set $[0,1]$ into three intervals $I_1, .., I_3$ as shown.

On the interval I_1 i's concession function coincides with the identity. In this interval his announcements are humble enough in the sense that guaranteeing his announcement to all players simultaneously is no problem. An announcement in the intervall I_2 cannot be guaranteed to all players but to all opponents of player i by means of an appropriate concession of player i. To determine the shape of the concession function in this area we have to connect the diagonal (connecting blisspoint and disagreement point) with the

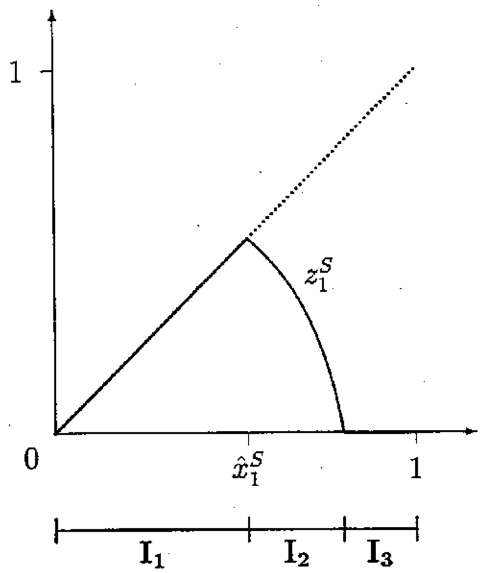

Figure 4: Shape of concession functions in the n–player case

diagonal in the $n-1$ dimensional bargaining problem where i is left out (cf. figure 3). And so convexity of S gives a concave shape of the concession function. In the interval I_3 there are those announcements that are so bold that they even cannot be guaranteed to all opponents simultaneously. But since we want each player to have an incentive to take part in the mechanism, we have to save individual rationality and so z_i^S should return 0 in this case.

Again we want to remark that a fixed fraction $x_i \in [0,1]$ yields a mapping $z_i^{(\cdot)} : \mathcal{B} \to [0,1]$.

We close this part with a lemma analogous to Lemma 2.1. It shows that even in the n–player world concession functions have a unique maximizer.

Lemma 3.1 *In the n–player setup there exists for any bargaining problem $S \in \mathcal{B}$ a unique maximizer of the concession functions $z_i^S \quad (i = 1, .., n)$.*

The proof can easily be carried over from the proof of lemma 2.1.

3.3 The mechanism for n–players

The strategic mechanism that implements the Kalai–Smorodinsky choice rule for n players does not deviate in general from that for two players. Again we are going to construct a mechanism such that the derived (non-cooperative) games have a unique equilibrium in dominant strategies. Furthermore the outcome function applied to this equilibrium yields an outcome suggested by the KS choice rule, i.e. for any bargaining problem the referring DSE strategies lead to the (fractional) Kalai–Smorodinsky bargaining solution.

The exact definition of the mechanism is as follows:
As we have seen in the brief discussion of the n–player context a mechanism G^* is a tuple $(N, \Sigma_1 \times \ldots \times \Sigma_n, g^*)$ where Σ_i is player i's strategy set and $g^* : \Sigma \to A$ denotes the outcome function of the mechanism. As in section 2 we will again use the unit interval as the strategy set for each player, i.e. $\Sigma_i := [0,1] \quad (i \in N)$. Define the outcome function $g^* : [0,1]^n \to A$ by

$$\begin{array}{rcl} (x_1, \ldots, x_n) & \mapsto & g^*(x_1, \ldots, x_n) \quad \in \quad A \\ g^*(x_1, \ldots, x_n)(\cdot) & := & (z_1^{(\cdot)}(x_1), \ldots, z_n^{(\cdot)}(x_n)). \end{array}$$

The mechanism G^* together with a bargaining problem $S \in \mathcal{B}$ (i.e. a preference profile) induces an n–person game Γ^S in strategic form. We have

$$\Gamma^S := ([0,1], \ldots, [0,1], v_1^S, \ldots, v_n^S),$$

where the payoffs are compositions of the outcome function g^* and utility functions u_i^S, representing the "preference profile" S, i.e. $v_i^S := u_i^S \circ g^*$.

Theorem 3.2 *For each bargaining problem $S \in \mathcal{B}$ the induced game Γ^S has a unique equilibrium $(\hat{x}_1^S, \ldots, \hat{x}_n^S) \in [0,1]^n$ in dominant strategies. This is given by $\hat{x}_i^S := \kappa^J(S)_i = l(S)$ for all $i \in N$.*

Considering Lemma 3.1 and the proof of the related theorem 2.2 the proof of this theorem is straightforward.

Finally we want to state that the mechanism really works. That means we want to show the implementation conditions.

Theorem 3.3
In the n–player context the "implementation conditions"

$$\begin{aligned} g^*(DSE(\Gamma^S)) &\subseteq \mathcal{K}(S) \\ u^S(g^*(DSE(\Gamma^S))) &= \kappa^f(S) \end{aligned}$$

are satisfied for each $S \in \mathcal{B}$.

The proof is immediate from the proof of the analogous theorem 3.3.

4 Subgame–Perfect Implementation

This section constitutes the second part of this paper. Here we want to take the mechanisms constructed in the previous sections and derive a **game in extensive form** rather than a non–cooperative **one–shot game**. Here players do not act simultaneously but one after another. The result we will get from the analysis will be that the Kalai–Smorodinsky solution is implemented by a **subgame perfect equilibrium**. And this equilibrium will be (quasi–)unique[7].

4.1 The induced two stage game

Implementation in dominant strategies means that we have to construct a strategic mechanism such that the induced games are one–shot games. Now we want to derive games in extensive form, so the "rules of the game" say that players choose their strategies one after another. Again we first concentrate on the two player case. This will lead to some form of a *Divide and Choose* mechanism[8], where the first player is the divider, who suggests two possible utility allocations, and the second player chooses among these alternatives. Which points does the divider propose? Let us remind of the procedure in the last two sections. Player 1 announces a number $x_1 \in [0, 1]$. Depending on whether the

[7]The term "quasi–unique" will be discussed and formalized later.
[8]For an extensive discussion of such mechanisms see Brams & Taylor (1996).

point $x_1\,\bar{\beta}(S)$ is feasible, he eventually has to make concessions to ensure a feasible utility allocation. Suppose we have a fixed bargaining problem $S \in \mathcal{B}$ and an announcement $x_1 \in [0,1]$. Starting from the point $x_1\,\bar{\beta}(S)$ there are two possible cases. First, this point is not feasible. Then there are **two** canonical ways to achieve a Pareto–efficient point in S, namely either player 1 or player 2 has to make a concession. So one player gets a fraction of x_1 of his blisslevel and the other gets less than that. In the other case, where $x_1\,\bar{\beta}(S)$ belongs to S this can be done the other way round. There are also two "natural" efficient points, because we can either increase player 1's or player 2's utility until the Pareto frontier is reached. So again, one player gets a fraction of x_1 of his blisslevel and the other gets more than that. Figure 5 illustrates this idea.

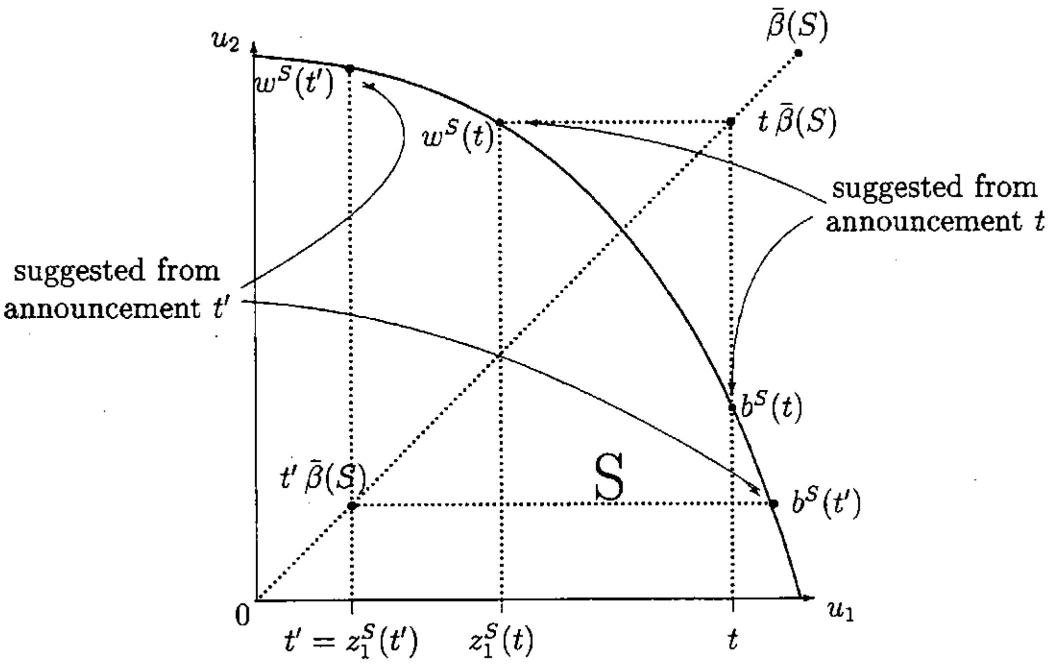

Figure 5: "Divide and Choose" for two players

So for a bargaining problem $S \in \mathcal{B}$ we derive the following two–stage game:

First stage: Player 1 annonunces a number t in the unit interval. This announcement defines two points $b^S(t)$ and $w^S(t)$ on the Pareto frontier (cf. figure 5).

Second stage: Player 2 chooses one of these points.

Final outcome: The point that player 2 has chosen.

Player 1 has a continuum of strategies whereas player 2 has to decide among two alternatives. This is depicted in Figure 6.

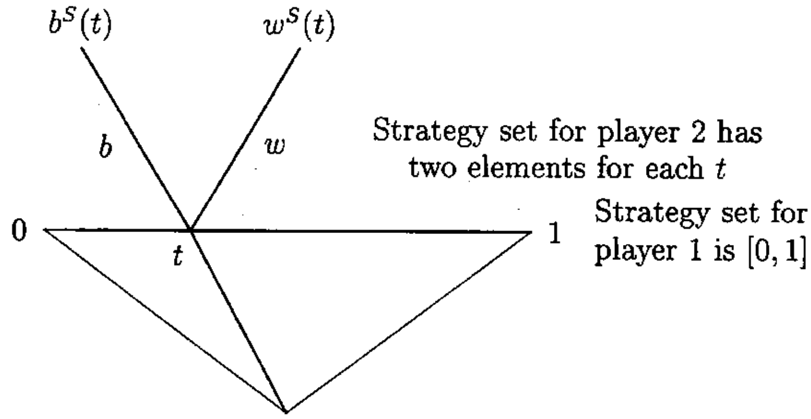

Figure 6: The two-stage-game

4.2 Implementation

Since the derived game is a game with finite horizon we know from Kuhn's theorem that it posesses a subgame–perfect equilibrium. This equilibrium can be easily described. Any announcement t from player 1 suggests one point where player 1 gets a fraction that is at least as large as his opponent's one. The other point gives him a fraction that is at most as large as his opponent's one. Let us denote the former by $b^S(t)$ and the latter by $w^S(t)$ ($b^S(t)$ is the best point for player 1 and $w^S(t)$ is the worst point for him). From player 2's point of view the notion of "good" and "bad" is just the other way round. This is due to convexity of utility possibility sets. Let us determine the subgame–perfect equilibrium by backwards induction:
Let $S \in \mathcal{B}$ be given and let t be a fixed announcement from player 1. Facing t player 2 has to choose in stage 2 either $b^S(t)$ or $w^S(t)$. Since $w^S(t)$ is **always** at least as good for him as $b^S(t)$ he always chooses $w^S(t)$.
Having this in mind it's up to player 1 to choose a strategy t from the unit interval. If he chooses a fraction that is higher than the "KS–fraction" $l(S)$, then by nature of the

construction of b^S and w^S and convexity of S the point $w^S(t)$ (that player 2 will choose) gives him a fraction less than $l(S)$. If player 1 announces $t < l(S)$, then he can do better by increasing his announcement, because in this case he will end up with exactly his announcement. So player 1 will choose the number $l(S)$. In that case $b^S(t)$ and $w^S(t)$ coincide (with the KS bargaining solution).

Therefore we get a (quasi–)unique subgame–perfect equilibrium, claiming that player 1 chooses the KS fraction and player 2 always chooses the point $w^S(t)$. The term "quasi" refers to a slight abuse in uniqueness, because player 2 can also choose $b^S(l(S))$ when player 1 announces $l(S)$ since in this case, and only in this case, player 2 has to decide between identical points. So to be precise, we have in fact two subgame–perfect equilibria but both yield the same outcome for both players, namely the KS bargaining solution. Moreover there is no coordination problem arising from the existence of two equilibria, because player 1 has a unique equilibrium strategy and equality of outcomes in equilibrium makes it unimportant, which of the equilibria player 2 chooses.

4.3 Extension to the n–player case

In the last part of this section we want to look at the n–player case and also derive an n–stage game, similar to that for two players.

In the two–player case player 1 makes an announcement in the unit intervall that gives player 2 a problem of choosing between 2 points. How did we get these two points? In the case where player 1's announcement is too bold, we can either let player 1 or player 2 make a concession to obtain two points on the Pareto boundary. If his announcement yields a feasible point, we can increase either 1's or 2's utility until the Pareto frontier is reached. In the n–player case we can apply a similar method. Here we have to distinguish between three types of announcement of player 1. First consider an announcement $t \in [0, 1]$ such that $t\,\bar{\beta}(S)$ belongs to S. Here we can increase one of the n player's utility and therefore get n points on the Pareto frontier. In the second case the point $t\,\bar{\beta}(S)$ is not feasible, but all "projections", i.e. all points $\sum_{j \neq i} t\,\bar{\beta}_j(S)\,e^j$ $(i \in N)$ belong to S. Here we again get n efficient points by forcing one player to make a concession. The remaining case is somewhat difficult. In this case at least one "projection" is not feasible and we want the disagreement point, namely 0. to be the only point selected as a form of punishment.
The n–stage game that we want to derive from this is as follows:

First stage: Player 1 makes an announcement in the unit interval. As described above this anoncement yields n (not necessarily different) points in the utility allocation set S.

Second stage: Player 2 "sees" these n points and identifies those that give him the same utility. Then he chooses one of the resulting *utility levels* and therefore he chooses the referring set of points.

i–th stage: Player i now faces a set of points that player $i-1$ has chosen. Again he sorts them by his utility and chooses one of these utilities resulting in a subset of points.

Final outcome: Player n has chosen among those points that player $n-1$ has selected. So we will end up with a subset of those points that player 1 originally suggested by his announcement t. But this subset does not contain two different points (see argument below). Then this unique point player n has chosen yields the outcome of the game.

We have to clarify why there is exactly one point that player n chooses by choosing a utility level. Suppose there are still two points in S chosen by player n, say v and w. From the choice procedure above we know that all players 2,..,n are indifferent between v and w. So they can at most differ in their first coordinate. But since all points suggested by player 1's announcement are efficient (or all are equal to the disagreement point) v and w have to coincide.

Let us see how the mechanism works for three players. Figure 7 depicts the construction of the three points suggested by an announcement of player 1 and figure 8 illustrates the derived 3–stage game.

Consider the announcement t'. This gives us three points $I, II, III \in S$. Points I and II assign the same utility to player 2. He does not have to distinguish between them. Point III gives him a different utility level. So in stage 2 of the game he has to choose between the utility level given by points I and II and the utility level given by point III. If he decides to take the first one, he actually chooses two out of the original three points. In the third stage player 3 faces these two points. The two (different) points I and II give him different utilities (cf. figure 7). Now he decides between these two utility levels and we end up with one point.
If player 2 chooses the utility level arising from point III then things are quite easy for

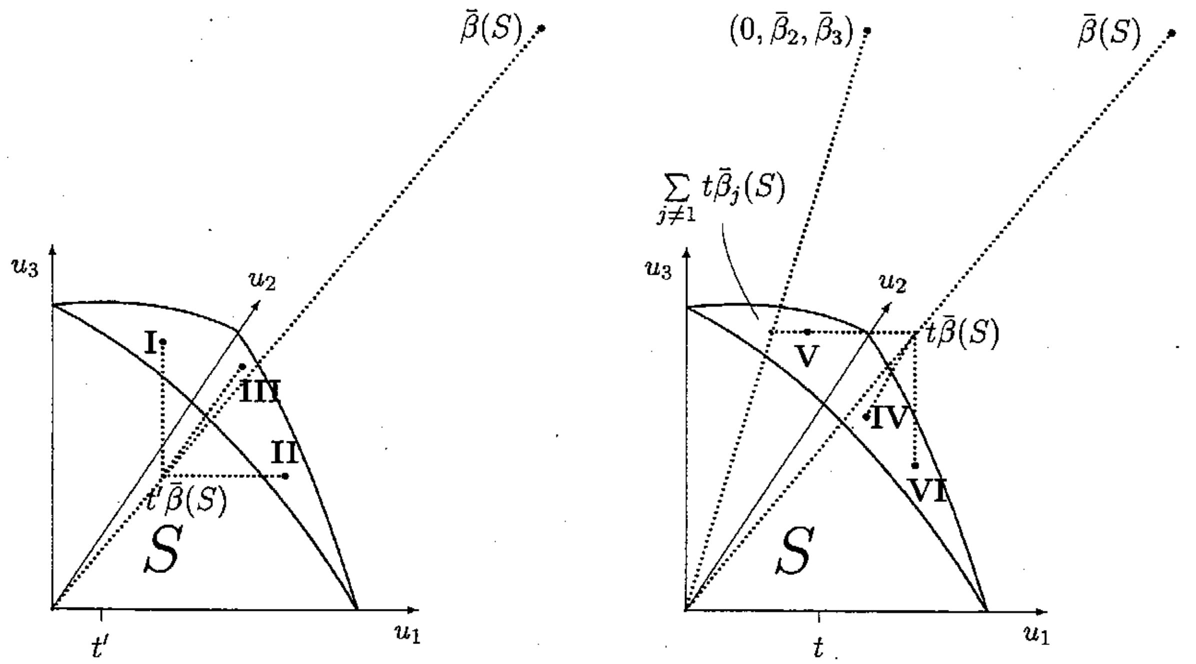

Figure 7: Divide and choose for n players

player 3, because he has no choice problem at all and we end up with point III.
What about optimal strategies in this game? Again we refer to the example depicted in figures 7 and 8. The first thing to remark is that we can partition player 1's strategy set $[0, 1]$ into three intervals I_1, I_2, I_3. For announcements t in the first interval the point $t\,\bar{\beta}(S)$ lies in S. Here we get the three points by increasing one player's utility. So in the second stage player 2 faces two different utility levels. One gives him a fraction of t and the other gives him at least this fraction[9]. So it is optimal for him to choose the point where his utility is increased. But then all players entering the scene in later stages have no choice problem at all and we will finally end up with the point that player 2 determined.

The situation is more complex for an announcement t in the interval I_2. Here the three points for stage two are determined by individual concessions. Again player 2 faces (at most) two different utilities, depending on whether his utility is reduced or not. Optimal

[9]In fact player 2 only faces one level, if $t = l(S)$, since in this case, and only in this case, the three points coincide.

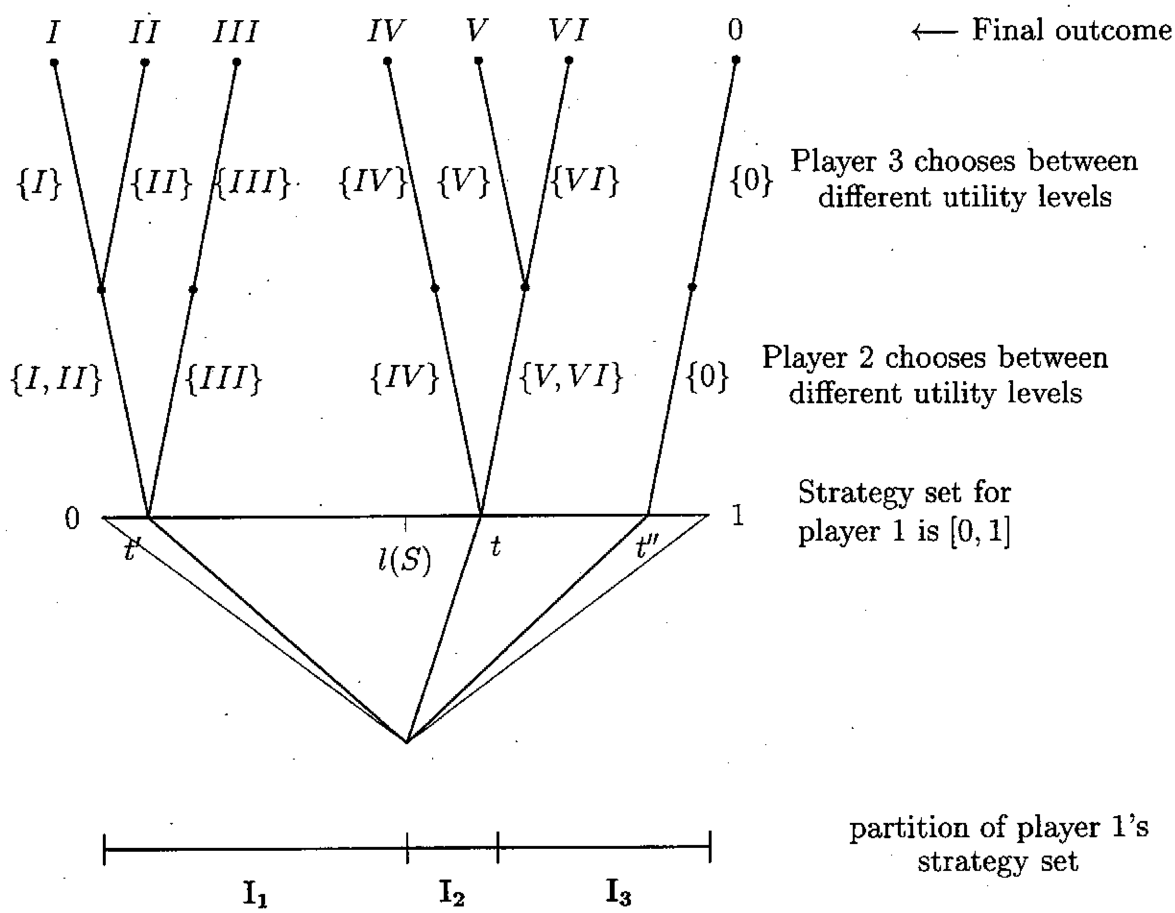

Figure 8: The n–stage game

behaviour claims that he chooses all points where someone else has to make a concession and he gets a fraction of t. In our example he chooses points V and VI. These points are delivered to the third stage, where player 3, playing optimal, chooses the one, where he does not have to make the concession. So we always end up with the point where player 1 has to concede. (In the example we end up with point V.)

The situation for "high announcements" $t'' \in I_3$ is trivial, since only the disagreement point is delivered to the second stage and so no real choice takes place and we end up with the disagreement point. Note that the frontier between I_1 and I_2 is given by $l(S)$.

But which behaviour is optimal for player 1, knowing what his opponents will choose?

The answer is "playing Kalai–Smorodinsky", i.e. announcing l(S). First of all he will not announce someting in I_3, because he can assure himself more than 0 by announcing $t \in I_1$ and getting this fraction t[10]. For any announcement $t \leq l(S)$ he gets exactly this fraction. So he can assure himself a fraction of $l(S)$ by announcing this. If he announces $t > l(S)$ he will be the one who concedes in the end and by the analysis in the whole paper[11] we know that this means getting less than $l(S)$.
We conclude that the following strategies constitute a subgame–perfect equilibrium of the n–stage game:

Player 1: He announces the KS fraction $l(S)$.

Players 2 to n: They choose the highest possible utility level and therefore deliver the referring point(s) to the next stage.

Final outcome: The Kalai–Smorodinsky bargaining solution, because all points delivered to stage two coincide with the KS solution and therefore we have no choice problems at all.

Again we see that there can be only one subgame–perfect equilibrium. In contrast to the two–player case we get "real" uniqueness because we slightly modified the game by letting the agents choose between utilities rather than between the referring points.

We close this section with an interesting observation. The SPE–implementation shows that like in zero–sum games there is a "max–min" problem behind the solution[12]. In fact, if we consider the induced one–shot games for two players (cf. section two), we can describe player 1's problem to choose the "optimal strategy" by the "maximization problem"

$$\max_{x_1 \in [0,1]} \min \left\{ b_1^S(t), w_1^S(t) \right\} \tag{4}$$

(with $b_1^S(t), w_1^S(t)$ as in this section). So, knowing the min in (4) for all t, player 1 faces an ordinary maximization problem. And this problem has a unique solution, namely the Kalai–Smorodinsky fraction $l(S)$.

[10]He can get more than 0, because we have assumed that the set of feasible utility allocations includes one strictly positive point. Then he can choose $t > 0$ small enough and end up with t.

[11]especially in the proof of lemma 2.1

[12]The author owes this viewpoint to Prof. Trockel.

5 Concluding Remarks

In the implementation literature our notion of implementation is often referred to as *weak implementation*, because we do not claim that every socially desired outcome can be achieved by application of the outcome function to an appropriate equilibrium of the induced game. What we have is therefore a subset condition rather than equality in the implementation condition (see theorems 2.3 and 3.3).
But having the whole story in mind, this actually means no weakening of our results. To see this, remember that the KS choice rule assigns a set of "bargaining solutions" (which need not be feasible in every bargaining problem) to a preference profile (represented by a bargaining problem $S \in \mathcal{B}$). But all mappings in $\mathcal{K}(S)$ coincide at S with the KS bargaining solution $\kappa^f(S)$. So, once "real" players face the bargaining problem S that is determined by their preference profile, this situation is common knowledge and only this situation really matters. And here the dominant strategy equilibrium yields exactly the only utility allocation for this situation that is given by the KS choice rule, namely the KS bargaining solution of S.[13]

Considering the one shot games induced by the mechanisms of sections two and three, one may complain about two things. First the Kalai–Smorodinsky outcome is the only efficient point that can be achieved by playing the game and second payoffs do not depend on strategies of other players, so that there is no real interaction in the game. Both facts are true so far. But let us briefly discuss a slight modification of the mechanism. We take the payoff of the game (a point in the utility allocation set) and make a "proportional adjustment", i.e. we increase both players utility according to the ratio of their blisslevels until the pareto frontier is reached[14]. So we equally distribute additional percentages of blisslevels. This intersection point is defined to be the payoff of the new game. In this modified game each outcome is an efficient point and payoffs do depend on the other players' strategies. But even in this case "playing Kalai–Smorodinsky" is still the unique equilibrium in dominant strategies. Roughly speaking, this is true, because the "max–

[13]Mathematically every bargaining problem S induces an equivalence relation on the outcome space A, meaning that two elements $L, L' \in A$ are *S–equivalent*, if they yield the same outcome in S, i.e. $L(S) = L'(S)$. Taking this into account we can get *full implementation* by assigning the equivalence class which is represented by the KS bargaining solution to $\mathcal{K}(S)$. Then the S–equivalence class generated by $g^*(DSE(\Gamma^S))$ **equals** $\mathcal{K}(S)$.

[14]Formally we consider the ray starting at the original payoff and lying parallel to the diagonal $[0, \bar{\beta}(S)]$ and determine its intersection with the set of efficient points in S.

min" structure[15] behind the solution in the simple setup is preserved in its essence. But since the analysis of these games or the referring mechanism is more complicated and does not give new insights, we decided to concentrate on the "simpler" mechanism presented in the second and third section.

It is obvious that the analysis of the games in extensive form (cf. section 4) does **not** depend on a special ordering of the players. We can even construct a game where random chooses one divider and an ordering of the choosers. Even in this game there is only one subgame–perfect equilibrium, which is mainly given by the equilibrium described above. And in this SPE we end up with the KS solution.

To summarize, we have shown that the strategic mechanisms of sections two and three with derived one–shot games also give an idea to construct an extensive mechanism with derived n–stage games. Whereas the strategic mechanism implements the KS bargaining solution in dominant strategies, the extensive one implements it by a subgame–perfect equilibrium which is (quasi–)unique.

Thinking of bargaining problems with a specific underlying structure (e.g. cake–cutting problems), one could imagine that the mechanisms described above lead to some division rule. Since any choice rule that is implementable in dominant strategies is also truthfully DSE–implentable (cf. Mas-Colell, Whinston & Green (1995, Proposition 23.C.1)), we may hope to get a mechanism in a setup, where preferences are common knowledge among the players but are not observable by a court that registers contracts, and no player can take an advantage from cheating in his own preferences. This will be an interesting line for future research.

[15]see last remark in section four.

References

BRAMS, S. J. & A. D. TAYLOR (1996): *Fair Division – from cake cutting to dispute resolution.* Cambridge University Press, Cambridge.

CRAWFORD, V. P. (1978): "A Procedure for generating Pareto–Efficient Egalitarian–Equivalent Allocations", *Econometrica*, 47, 49–60.

DEMANGE, G. (1984): "Implementing Efficient Egalitarian Equivalent Allocations", *Econometrica*, 52, 1167–1177.

KALAI, E. & M. SMORODINSKY (1975): "Other Solutions to Nash's Bargaining Problem", *Econometrica*, 43, 513–518.

MAS-COLELL, A., M. D. WHINSTON & J. R. GREEN (1995): *Microeconomic Theory.* Oxford University Press, New York – Oxford.

MOULIN, H. (1984): "Implementing the Kalai–Smorodinsky Bargaining Solution", *Journal of Economic Theory*, 33, 32–45.

NASH, J. F. (1950): "The Bargaining Problem", *Econometrica*, 18, 155–162.

——— (1953): "Two–person Cooperative Games", *Econometrica*, 21, 128–140.

OSBORNE, M. J. & A. RUBINSTEIN (1996): *A Course in Game Theory.* MIT Press, Massachusetts.

TROCKEL, W. (1998): "An Exact Implementation of the Nash Bargaining Solution in Dominant strategies", in *Functional Analysis and Economic Theory*, ed. by Y. Abramovich, E. Avgerinos & N. C. Yannelis. Springer, Heidelberg.

VAN DAMME, E. (1991): *Stability and Perfection of Nash Equilibria.* Springer Verlag, Heidelberg, second edn.