INCENTIVE-DUAL METHODS IN GAME THEORY: VIRTUAL UTILITY AND DUAL REDUCTION REVISITED by Roger B. Myerson <myerson@uchicago.edu>

Abstract: Incentive-compatible mechanisms and correlated equilibria in game theory are defined by systems of linear incentive constraints. The duals of these linear systems yield shadow prices for these incentive constraints which can be used to get insights into these games. This talk will review and integrate two strands of the research literature in which these duals have been applied. From the duals of the informational incentive constraints that characterize incentive-efficient mechanisms for a Bayesian game, we get virtual utility functions that intuitively characterize the costly signaling actions that may occur in these incentive-efficient mechanisms. This concept of virtual utility also enables us to extend cooperative solution concepts like the core, the Nash bargaining solution, and the Shapley value, to general games with incomplete information. As a application, these solutions for bargaining games with incomplete information can offer a more robust explanation of inefficient underemployment in labor contracts, even in situations where the employer is risk neutral, has the only private information, and has all the bargaining power. The duals of the strategic incentive constraints that characterize the correlated equilibria of a strategic-form game can be used to reduce any such game to an elementary game in which the knife-edge indifference problems of equilibrium-imperfection do not arise.

INCENTIVE-DUAL METHODS IN GAME THEORY: VIRTUAL UTILITY AND DUAL REDUCTION REVISITED by Roger B. Myerson <myerson@uchicago.edu>

For presentation at LAMES, Brazil, July 2002

0. Duals of resource constraints are economically important as a mathematical model of prices. Game theory has added incentive constraints to resource constraints as essential parts of the economic problem that can be recognized by mathematical economic theory.
Might duals of incentive constraints have some important economic meaning?

1. Incentive-efficiency in Bayesian choice problems with informational incentive constraints

$\Gamma = (N, C, (T_i)_{i\in N}, (u_i)_{i\in N}, P)$ where:

N={players}, C ={jointly feasible actions}, T_i={i's possible types}, (nonempty finite sets)

$T = \times_{i\in N} T_i$ = {type profiles}, $u_i: C\times T \to \mathbb{R}$ (i's utility function), $P \in \Delta(T)$ (probability distn).

Notation $T_{-i} = \times_{j\in N-i} T_j$, $p_i(t_i) = \sum_{t_{-i}\in T_{-i}} P(t_{-i},t_i)$, $P_i(t_{-i}|t_i) = P(t)/p_i(t_i)$, $t = (t_{-i},t_i)$.

Suppose each player also has a <u>nonparticipation</u> option that yields a payoff of 0.

Consider a <u>mechanism</u> $\mu: T \to \Delta(C)$ satisfying $\mu(d|t) \geq 0 \ \forall d\in C$, and $\sum_{c\in C} \mu(c|t) = 1, \ \forall t\in T$.

Let $U_i(\mu|t_i) = \sum_{t_{-i}\in T_{-i}} P_i(t_{-i}|t_i) \sum_{c\in C} \mu(c|t)\, u_i(c,t)$,

$\hat{U}_i(\mu,s_i|t_i) = \sum_{t_{-i}\in T_{-i}} P_i(t_{-i}|t_i) \sum_{c\in C} \mu(c|t_{-i},s_i)\, u_i(c,t)$.

μ is <u>incentive compatible</u> iff $U_i(\mu|t_i) \geq 0$ and $U_i(\mu|t_i) \geq \hat{U}_i(\mu,s_i|t_i), \ \forall s_i\in T_i, \ \forall t_i\in T_i, \ \forall i\in N$.

We consider only informational incentive constraints until near the end. That is, the players can surrender control of all actions to a mediator, who then chooses an action profile in C given reports from each player about his or her type, but players can misreport their types.
By the revelation principle, any equilibrium of any mechanism is equivalent to an incentive compatible mechanism, satisfying the informational incentive constraints above.
A mechanism is <u>(weakly) incentive-efficient</u> iff it is incentive compatible and no other incentive-compatible mechanism yields higher expected utilities for all types of all players.

2. Lagrangean conditions for incentive-efficiency and virtual utility

So an incentive-compatible μ is incentive-efficient iff there exists some $\lambda = (\lambda_i(t_i))_{i\in N, t_i\in T_i}$ such that $\lambda_i(t_i) \geq 0,\ \forall i\in N, \forall t_i\in T_i,$ with some $\lambda_i(t_i) > 0,$ and μ is a optimal solution to the problem

$$\text{maximize}_{\mu:T\to\Delta(C)} \sum_{i\in N}\sum_{t_i\in T_i} \lambda_i(t_i)U_i(\mu|t_i) \text{ subject to } U_i(\mu|t_i) \geq \hat{U}_i(\mu,s_i|t_i),\ \forall i\in N,\ \forall t_i\in T_i,\ \forall s_i\in T_i.$$

Given the utility weights λ, this problem is actually a linear programming problem in μ.

Let $\alpha_i(s_i|t_i)$ denote the Lagrange multiplier (or dual variable) for incentive constraint that type t_i of i should not want to pretend to be s_i. Them the Lagrangean for this optimization is:

$$\mathbf{L}(\mu,\alpha,\lambda) = \sum_{i\in N}\sum_{t_i\in T_i} [\lambda_i(t_i)U_i(\mu|t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i)(U_i(\mu|t_i) - \hat{U}_i(\mu,s_i|t_i))]$$

Given λ and α, this Lagrangean is a linear function of μ:

$$\begin{aligned}
\mathbf{L}(\mu,\alpha,\lambda) &= \sum_{i\in N}\sum_{t\in T} \Big(\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i)\Big) P_i(t_{-i}|t_i) \sum_{c\in C} \mu(c|t)\, u_i(c,t) \\
&\quad - \sum_{i\in N}\sum_{t\in T}\sum_{s_i\in T_i} \alpha_i(s_i|t_i)\, P_i(t_{-i}|t_i) \sum_{c\in C} \mu(c|t_{-i},s_i)\, u_i(c,t) \\
&= \sum_{i\in N}\sum_{t\in T} \Big(\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i)\Big) P_i(t_{-i}|t_i) \sum_{c\in C} \mu(c|t)\, u_i(c,t) \\
&\quad - \sum_{i\in N}\sum_{t\in T}\sum_{s_i\in T_i} \alpha_i(t_i|s_i)\, P_i(t_{-i}|s_i) \sum_{c\in C} \mu(c|t)\, u_i(c,(t_{-i},s_i)) \\
&= \sum_{i\in N}\sum_{t\in T} P(t) \Big(\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i)\Big) \sum_{c\in C} \mu(c|t)\, u_i(c,t)/p_i(t_i) \\
&\quad - \sum_{i\in N}\sum_{t\in T} P(t) \Big(\sum_{s_i\in T_i} \alpha_i(t_i|s_i) \sum_{c\in C} \mu(c|t)\, u_i(c,(t_{-i},s_i))\, P_i(t_{-i}|s_i)/P_i(t_{-i}|t_i)\Big)/p_i(t_i)
\end{aligned}$$

For any c and t, find all the terms for player i in this Lagrangean that are multiplied by $\mu(c|t)$.

We define $v_i(c,t,\lambda,\alpha)$, the <u>virtual utility</u> for player i from an action-profile c with types-profile t, to be the sum of these terms divided by the probability P(t). So the formula for virtual utility is

$$v_i(c,t,\lambda,\alpha) = \Big[\Big(\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i)\Big)u_i(c,t) - \sum_{s_i\in T_i} \alpha_i(t_i|s_i)u_i(c,(t_{-i},s_i))P_i(t_{-i}|s_i)/P_i(t_{-i}|t_i)\Big]/p_i(t_i)$$

With this definition, we can write the Lagrangean function more simply as

$$\mathbf{L}(\mu,\alpha,\lambda) = \sum_{t\in T} P(t) \sum_{c\in C} \mu(c|t) \sum_{i\in N} v_i(c,t,\lambda,\alpha)$$

That is, the Lagrangean is just the expected sum of the virtual-utility payoffs to all players.

3. Characterizing incentive efficiency by the virtual-utility hypothesis

In Lagrangean analysis, the optimality conditions for a mechanism μ then are:

[1] $\sum_{c\in C} \mu(c|t) \sum_{i\in N} v_i(c,t,\lambda,\alpha) = \max_{c\in C} \sum_{i\in N} v_i(c,t,\lambda,\alpha)$, $\forall t\in T$ (unconstrained Lagrange-optimality),

[2] $U_i(\mu|t_i) \geq \hat{U}_i(\mu,s_i|t_i)$, $\forall i$, $\forall t_i\in T_i$, $\forall s_i\in T_i$ (incentive compatibility), and

[3] $\alpha_i(s_i|t_i)\big(U_i(\mu|t_i) - \hat{U}_i(\mu,s_i|t_i)\big) = 0$, $\forall i$, $\forall t_i\in T_i$, $\forall s_i\in T_i$ (complementary slackness).

Let us say that a type t_i of player i jeopardizes another type s_i of player i, in the incentive-efficient μ, iff the constraint that t_i should not want to imitate s_i ($U_i(\mu|t_i) \geq \hat{U}_i(\mu,s_i|t_i)$) is binding and has a positive Lagrange multiplier in the optimality conditions for μ.

When the player's types are independent random variables, we have $P(t) = \prod_{j\in N} p_j(t_j)$ $\forall t$, and $P_i(t_{-i}|s_i) = \prod_{j\in N-i} p_j(t_j) = P_i(t_{-i}|t_i)$, and so virtual utility formula simplifies to

$$v_i(c,t,\lambda,\alpha) = \Big[\big(\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i)\big)u_i(c,t) - \sum_{s_i\in T_i} \alpha_i(t_i|s_i)u_i(c,(t_{-i},s_i))\Big] \big/ p_i(t_i).$$

Henceforth we assume such independence.

Multiplying a type's utility function by a positive constant is decision-theoretically inessential.

So essential difference between the virtual utility $v_i(c,t,\lambda,\alpha)$ for a type t_i of player i and the actual utility $u_i(c,t)$ is that the virtual utility of type t_i exaggerates the difference from the utilities of i's other types that jeopardize t_i.

So we have the following general proposition to help us to qualitatively understand the ex-post inefficiencies (signaling costs) that may be incurred in an incentive-efficient mechanism:

The incentive-efficient mechanism will be ex-post efficient in terms of the players' virtual utilities, where the virtual utility of any type t_i differs from the actual utility by exaggerating the difference from the other possible types that jeopardize t_i.

In this framework, a Coasian believer in ex-post efficiency could equivalently "explain" inefficient signaling costs by the following virtual-utility hypothesis:

In a situation of incomplete information where possible misrepresentation is problematic, individuals may act according to their virtual utilities, exaggerating the difference from the alternative types that jeopardize them.

4. Special cases where virtual utility is simple: independent types and utility transfers

The virtual-utility hypothesis is useful for predicting the nature of incentive-efficient signaling costs in situations where we have a good intuition about which incentive constraints will be binding and problematic.

For example, in an economic situation where a seller may be a good or bad type (selling high- or low-quality goods), we may intuitively expect that the bad type will jeopardize the good type, but not visa versa. If so, then the good type's virtual utility will differ from its actual utility by exaggerating the difference from the bad type. So the good type may do unpleasant activities that do not benefit anyone else, if his bad type would hate them more so that he gets positive virtual benefits from these activities. But the unjeopardized bad type will have virtual utility equivalent to actual utility, and so there will be no ex-post inefficiency for the bad type.

Now consider a situation where the players' types are independent $P(t) = \prod_{i\in N} p_i(t_i)$, the players can receive any profile of monetary transfers $x= (x_i)_{i\in N}$ that sum to zero $\sum_{i\in N} x_i = 0$, and each player has additively separable linear utility for money $u_i((d,x),t) = U(d,t) + x_i$.

For any $c = (d,x)$, the sum of virtual utilities in any state t can be written

$$\sum_{i\in N} v_i(c,t,\lambda,\alpha) = \sum_{i\in N} x_i \left[\left(\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i)\right) - \sum_{s_i\in T_i} \alpha_i(t_i|s_i)\right]/p_i(t_i) + \textit{(terms independent of x)}$$

Virtual ex-post efficiency of any finite vector x (constrained only $\sum_{i\in N} x_i = 0$) requires than the coefficients $\left[\left(\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i)\right) - \sum_{s_i\in T_i} \alpha_i(t_i|s_i)\right]/p_i(t_i)$ must be the same for all i, $\forall t\in T$. We may take this constant to be 1, without loss of generality. So in this case of independent types with linear utility for money, the vectors λ and α may be taken to satisfy the equations

$$\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i) = \sum_{s_i\in T_i} \alpha_i(t_i|s_i) + p_i(t_i) \quad \forall i\in N, \ \forall t_i\in T_i; \text{ and } \sum_{s_i\in T_i} \lambda_i(s_i) = 1.$$

These equations describe a "hydraulic" inflow to any type t_i, from $p_i(t_i)$ and the Lagrange multipliers $\alpha_i(s_i|t_i)$ of the types s_i that jeopardize t_i, which must be balanced by an outflow to the $\lambda_i(t_i)$ and the Lagrange multipliers $\alpha_i(s_i|t_i)$ of the types s_i that are jeopardized by t_i.

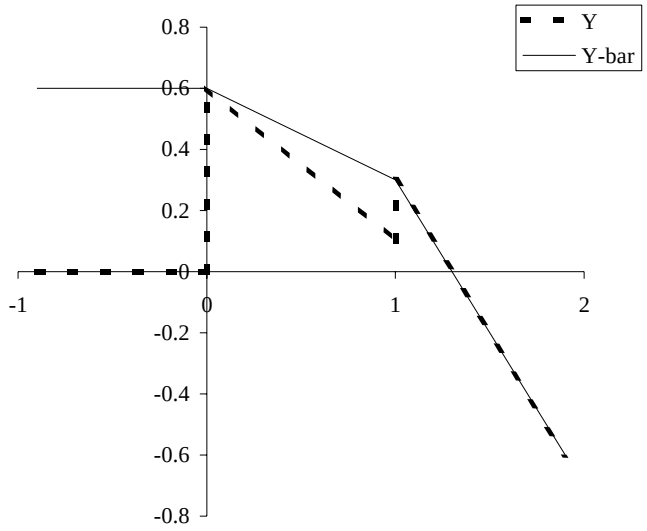
When the types in T_i are ordered on a line, it may be sufficient to consider only local incentive constraints between types that are adjacent to each other on the line. (Local IC => Global IC.) Suppose that t_i and $t_i+\varepsilon$ are adjacent types in T_i, with $\varepsilon>0$. Then the hydraulic equations imply $\alpha_i(t_i+\varepsilon|t_i) - \alpha_i(t_i|t_i+\varepsilon) = \sum_{s_i\le t_i} (p_i(s_i) - \lambda_i(s_i))$.

If higher types do not jeopardize lower, then $\alpha_i(t_i|t_i+\varepsilon) = 0$, $\alpha_i(t_i+\varepsilon|t_i) = \sum_{s_i\le t_i} (p_i(s_i) - \lambda_i(s_i))$.

5. More specialized cases where virtual utility is simple

Consider bargaining between seller(1) and buyer(2) where the seller has all private information.
The seller's type $t_1 = t$ is his value for an indivisible good that would be worth g(t) to the buyer.
Suppose $T = T_1 = \{A, A+\varepsilon, A+2\varepsilon,..., B\}$ consists of multiples of $\varepsilon>0$ from A to B (A<B),
each with positive probability, and utility is linear in money.
Then the incentive-efficient frontier is flat: all incentive-efficient mechanisms can be supported by the same λ, as shown in my paper in Alvin Roth's 1985 book. The construction is as follows:
Let $F(t) = \sum_{s\le t} P(s)$ = (cumulative probability of 1's type at t).
Let $Y(t) = \sum_{s\le t} P(s)(g(s)-t)$ = (2's EU from offer to buy for t).
At each t, Y(t) is continuous from above but may have discontinuities in limits from below.
Let $Y_-(t) = \sum_{s<t} P(s)(g(s)-t) = \lim_{\tau\uparrow t} Y(\tau) = Y(t) - P(t)(g(t)-t) = Y(t-\varepsilon) + F(t-\varepsilon)\varepsilon$.
Let $\bar{Y}:\mathbb{R}\to\mathbb{R}$ be the least concave function that is nonincreasing and is never less than Y.
Let $\lambda(t) = (\bar{Y}(t)-\bar{Y}(t-\varepsilon))/\varepsilon - (\bar{Y}(t+\varepsilon)-\bar{Y}(t))/\varepsilon = (2\bar{Y}(t) - \bar{Y}(t+\varepsilon) - \bar{Y}(t-\varepsilon))/\varepsilon$,
$\alpha(t|t+\varepsilon) = (\bar{Y}(t)-Y(t))/\varepsilon$,
$\alpha(t+\varepsilon|t) = (\bar{Y}(t+\varepsilon) + F(t)\varepsilon - Y(t))/\varepsilon = (\bar{Y}(t+\varepsilon) - Y_-(t+\varepsilon))/\varepsilon$.
Then 1's virtual values of keeping the object are
$v(t) = [(\lambda(t)+\alpha(t+\varepsilon|t)+\alpha(t-\varepsilon|t))t - \alpha(t|t+\varepsilon)*(t+\varepsilon) - \alpha(t|t-\varepsilon)*(t-\varepsilon)]/P(t) = g(t) \ \forall t\in T\backslash\{A,B\}$,
$v(A) = [(\lambda(A)+\alpha(A+\varepsilon|A))*A - \alpha(A|A+\varepsilon)*(A+\varepsilon)]/P(A) = g(A) - \bar{Y}(A)/P(A)$,
$v(B) = [(\lambda(B)+\alpha(B-\varepsilon|B))*B - \alpha(B|B-\varepsilon)*(B-\varepsilon)]/P(B) = g(B) + (\bar{Y}(B)-Y(B))/P(B)$.
(The extra terms in v(A) and v(B) are because $\varepsilon\alpha(A|A-\varepsilon)$ and $\varepsilon\alpha(B|B+\varepsilon)$ are missing.)
In any incentive-efficient mechanism, each incentive constraint must bind if it gets a positive dual value α from a $\bar{Y} > Y$ inequality, type B must not sell if $\bar{Y}(B) > Y(B)$,
and type A must sell for sure if $\bar{Y}(A) > 0$ but must not sell if $\bar{Y}(A) < 0$.
If $\alpha(t+\varepsilon|t)>0$ and $\alpha(t|t+\varepsilon)>0$ then incentive-efficient mechanisms must treat t and $t+\varepsilon$ the same.
Notice $\bar{Y}(A)=\max_{t\in T} Y(t)$.

For example, suppose $\varepsilon=1$, A=0, B=1, T = {0,1}, g(0) = 1.2, g(1) = 1.4, P(0) = 1/2 = P(1).
Then $Y(t) = 0 \ \forall t<0$, $Y(t) = (1.2-t)/2 \ \forall t\in[0,1)$, $Y(t) = 1.3-t \ \forall t\ge 1$.
So $\bar{Y}(t)=0.6 \ \forall t<0$, $\bar{Y}(t) = 0.6-0.3t \ \forall t\in[0,1)$, $\bar{Y}(t)=0.3-t \ \forall t\ge 1$.
All incentive-efficient mechanisms are supported by
$\lambda_1(0)=0.3$, $\lambda_1(1)=0.7$, $\alpha(0|1)=0$, $\alpha(1|0)=0.2$, which yield virtual values of keeping the object
$v(0) = ((0.3+0.2)*0-0.0)/0.5 = 0$, $v(1) = (0.7*1-0.2*0)/0.5 = 1.4$ for the two types of seller.



Example: A=1, B=3, $\varepsilon = 1$, g(t) = t+1 $\forall$t, p(1)=0.4, p(2)=0.1, p(1)=0.5.

Y(0) = 0 = Y_(0), Y(1) = 0.4, Y_(2) = 0, Y(2) = 0.1, Y_(3) = −0.4, Y(3) = 0.1, Y(4) = −0.9.

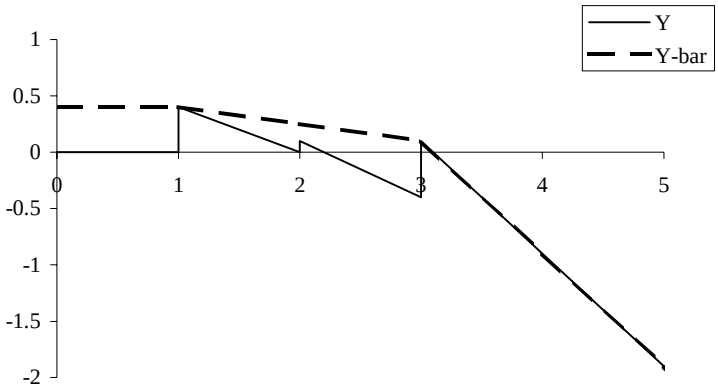
$\bar{Y}(0) = \bar{Y}(1) = 0.4$, $\bar{Y}(2) = 0.25$, $\bar{Y}(3) = 0.1$, $\bar{Y}(4) = -0.9$.

$\lambda(1) = (2*0.4-0.4-0.25)/1 = 0.15$,

$\lambda(2) = (2*0.25-0.4-0.1)/1 = 0$,

$\lambda(3) = (2*0.1-0.25--0.9)/1 = 0.85$,

$\alpha(1|2) = 0$, $\alpha(2|1) = 0.25$, $\alpha(2|3) = 0.15$, $\alpha(3|2) = 0.5$.



5(b) This characterization of the incentive-efficient mechanisms for seller-buyer problems with one-sided information uses two crucial primal-metheod lemmas:
(1) All informationally incentive-compatible mechanisms can be expressed as lotteries among of mechanisms of the form "sell for price r if the seller's value is below r, otherwise do not sell" plus some fixed monetary sidepayment. (Y(r) is the buyer's expected payoff from this mechanism, but where Y is discontinuous we must distinguish two mechanisms that differ by whether the indifferent type r sells or not.)
(2) Such a mechanism of selling only at price r is not incentive-efficient if $Y<\bar{Y}$ at r, because then we can construct a random price with expected value r such that the buyer would get strictly a higher expected utility from the random price mechanism, and no type of seller would be worse off from the change. See my article in Roth's 1985 book.

We may also consider a discrete Myerson-Satterthwaite (1983) bilateral bargaining problem where the types(values) of seller(1) and buyer(2) are multiples of $\varepsilon>0$ from A to B (A<B).
For the seller (1), let $\lambda_1(B) = 1$ and consider α where $\alpha_1(s_1|t_1)>0$ only when $s_1=t_1+\varepsilon$ (each type jeopardizes only the next higher type). Then $\alpha_1(t_1+\varepsilon|t_1) = \sum_{s_1\le t_1} p_1(s_1) = F_1(t_1) \ \forall t_1<B$,
and so type t_1's virtual value for keeping the object is
$v(t_1) = [F_1(t_1)\, t_1 - F_1(t_1-\varepsilon)\,(t_1-\varepsilon)]/p_1(t_1) = t_1 - F_1(t_1-\varepsilon)\,\varepsilon/p_1(t_1)$.
For the buyer (2), let $\lambda_2(A) = 1$, and consider α where $\alpha_2(s_2|t_2)>0$ only when $s_2=t_2-\varepsilon$.
Then $\alpha_2(t_2-\varepsilon|t_2) = \sum_{s_2\ge t_2} p_2(s_2) = 1 - F_2(t_2-\varepsilon), \ \forall t_2>0$,
and so type t_2's virtual value for the object is
$[(1-F_2(t_2-\varepsilon))\, t_2 - (1-F_2(t_2))\,(t_2+\varepsilon)]/p_2(t_2) = t_2 - (1-F_2(t_2))\,\varepsilon/p_2(t_2)$.
With $f_i=p_i/\varepsilon$, these are the virtual utilities in M-S's 1983 characterization of $U_1(\mu|B)+U_2(\mu|A)$.

6(a) From 1973 to 1983, I searched for natural ways to generalize cooperative solution concepts to games with incomplete information.
A critical problem was how to account for incentive constraints for subcoalitions.
(If we apply incentive constraints only to the grand coalition N that actually forms, then we lose superadditivity. If we apply incentive constraints to a subcoalition, we find that its incentive-feasible set may vary discontinuously in the actions of the complementary coalition.)
The key, I found, was to use the virtual utility hypothesis.
I first stumbled on this idea while searching for a generalized Nash bargaining solution, but now let me try introducing it here in a concept of <u>core</u> for incomplete-information games.

6. Virtual utility and the core with incomplete information

To formulate a generalized core with incomplete information, let us consider a game satisfying the usual assumption of orthogonal coalitions. That is, when a player joins coalition S, his payoff depends only on the actions that are chosen by the players in S from a given feasible set C(S). Then the cooperative Bayesian game is $\Gamma = (N, (C(S))_{S\subseteq N}, (T_i)_{i\in N}, (u_i)_{i\in N}, P)$, where $u_i(c,t) \in \mathbb{R}$ for any $S\subseteq N$, $i\in S$, $c\in C(S)$, $t=(t_i)_{i\in N} \in T = \times_{i\in N} T_i$. We write: $t_S = (t_i)_{i\in S} \in T_S = \times_{i\in S} T_i$.
Let's also assume independent types, so $p_S(t_S) = \prod_{j\in S} p_j(t_j)$, $P_i(t_{-i}|t_i) = p_{N-i}(t_{-i}) = \prod_{j\neq i} p_j(t_j)$.
The core is about allocations which could not be blocked by a deviating coalition that offers higher expected payoffs to all players when they join it. With incomplete information, we should allow that the identity of this blocking coalition and its actions might depend on the players' information in some probabilistic way, but this dependence must be incentive-compatible, and a blocking coalition should only use information available to its members.
In a random blocking rule δ, let $\delta(S,c|t_S)$ denote the probability that the members of S will form the blocking coalition and choose c in C(S), when t_S in T_S is their type-profile.
Consider any allocation $w = (w_i(t))_{i\in N, t\in T}$, where $w_i(t)$ denotes an ex-post payoff for player i when the type-profile is t. We say that w is strongly inhibitive iff no δ exists such that

$$\delta(S,c|t_S) \geq 0,\ \forall S\subseteq N,\ \forall c\in C(S),\ \forall t_S\in T_S,\ (\text{and } \textstyle\sum_{(S,c)} \delta(S,c|t_S) \leq 1,\ \forall t\in T)$$

$$\sum_{t_{-i}\in T_{-i}} \sum_{S\supseteq\{i\}} \sum_{c\in C(S)} P_i(t_{-i}|t_i)\,(u_i(c,t) - w_i(t))\,\delta(S,c|t_S) \geq 0,\ \forall i\in N,\ \forall t_i\in T_i,\ \text{with some strict } >, \qquad [\bullet\lambda_i(t_i)]$$

$$\sum_{t_{-i}\in T_{-i}} \sum_{S\supseteq\{i\}} \sum_{c\in C(S)} P_i(t_{-i}|t_i)\,(u_i(c,t) - w_i(t))\,(\delta(S,c|t_S) - \delta(S,c|t_{S-i},r_i)) \geq 0,\ \forall i\in N,\ \forall t_i\in T_i,\ \forall r_i\in T_i. \qquad [\bullet\alpha_i(r_i|t_i)]$$

Theorem The payoff allocation w is strongly inhibitive iff there exist vectors λ and α such that $\lambda_i(t_i) > 0$ and $\alpha_i(s_i|t_i) \geq 0$, $\forall i\in N$, $\forall t_i\in T_i$, $\forall r_i\in T_i$, and

$$\sum_{t_{N\setminus S}\in T_{N\setminus S}} p_{N\setminus S}(t_{N\setminus S}) \sum_{i\in S} \Big[\Big(\lambda_i(t_i) + \sum_{r_i\in T_i} \alpha_i(r_i|t_i)\Big) w_i(t) - \sum_{r_i\in T_i} \alpha_i(t_i|r_i)\, w_i(t_{-i},r_i)\Big] / p_i(t_i)$$
$$\geq \sum_{t_{N\setminus S}\in T_{N\setminus S}} p_{N\setminus S}(t_{N\setminus S}) \sum_{i\in S} \Big[\Big(\lambda_i(t_i) + \sum_{r_i\in T_i} \alpha_i(r_i|t_i)\Big) u_i(c,t) - \sum_{r_i\in T_i} \alpha_i(t_i|r_i)\, u_i(c,(t_{-i},r_i))\Big] / p_i(t_i),$$
$$\forall S\subseteq N,\ \forall c\in C(S),\ \forall t_S\in T_S \qquad [\bullet p_S(t_S)\delta(S,c|t_S)]$$

So an allocation is inhibitive iff, when we convert payoffs to some virtual utility scale and assume that this virtual utility is transferable, no coalition could ever expect a higher sum of virtual payoffs. Then define: core = {feasible allocations} $\cap$ {limits of strongly inhibitives}.

7. Mechanism selection by an informed principal

Now suppose that player 1 can select the mechanism. He can convey information within his mechanism itself, so we can assume without loss of generality that his selection is inscrutable. If the other players would infer nothing about his type from his selection, then honest play in any incentive-compatible mechanism would be an equilibrium.

But then each type might want to select a different mechanism, thus revealing the type!

We need a theory that restricts what is plausibly inscrutable, so that any other mechanism that 1 might prefer is ruled out, by a perception that it gives some type of 1 less than is warranted.

My 1983 theory has been dismissed as "using techniques drawn from cooperative game theory." But the essence of 1's problem is to co-ordinate with other possible types of himself, to induce the other players to accept the selected mechanism as a reasonable compromise among the different interests of all 1's possible types, such that no subset of 1's types could guarantee themselves more by a revealing defection. Such coordination on a reasonable compromise is "cooperation"!

Let $\Gamma = (N, (C_i)_{i\in N}, (T_i)_{i\in N}, (u_i)_{i\in N}, P)$ have independent types $P_i(t_{-i}|t_i)=\prod_{j\in N-i} p_j(t_j)$, and suppose that each player also has a non-participation option that gives him a payoff of 0.

We will consider when an allocation w_1 of interim expected utilities for the principal's types could be blocked by an alternative plan (δ,θ) where $\theta(t_1)$ denotes the probability that the principal would select alternative if his type were t_1, and $\delta(c|t)$ is the probability of the outcome c under this alternative when the type-profile is t.

(δ,θ) is a <u>viable alternative</u> to a utility allocation $w_1 = (w_1(t_1))_{t_1\in T_1}$ for the principal's types iff

$\delta(d|t) \geq 0 \ \ \forall d\in C$, and $\sum_{c\in C} \delta(c|t) = \theta(t_1) \geq 0, \ \ \forall t\in T$,

$$\sum_{t_{-1}\in T_{-1}} \sum_{c\in C} P_1(t_{-1}|t_1)\, \delta(c|t)\, (u_1(c,t)-w_1(t_1)) \geq 0, \ \ \forall t_1\in T_1, \text{ with strict inequality for some } t_1,$$

$$\sum_{t_{-1}\in T_{-1}} \sum_{c\in C} P_1(t_{-1}|t_1)\, \delta(c|t)\, (u_1(c,t)-w_1(t_1)) \geq \sum_{t_{-1}\in T_{-1}} \sum_{c\in C} P_1(t_{-1}|t_1)\, \delta(c|t_{-1},s_1)\, (u_1(c,t)-w_1(t_1)), \quad \forall t_1\in T_1, \ \forall s_1\in T_1,$$

$$\sum_{t_{-i}\in T_{-i}} \sum_{c\in C} P_i(t_{-i}|t_i)\, \delta(c|t)\, u_i(c,t) \geq 0, \ \ \forall i\in N-1, \ \forall t_i\in T_i,$$

$$\sum_{t_{-i}\in T_{-i}} \sum_{c\in C} P_i(t_{-i}|t_i)\, \delta(c|t)\, u_i(c,t) \geq \sum_{t_{-i}\in T_{-i}} \sum_{c\in C} P_i(t_{-i}|t_i)\, \delta(c|t_{-i},s_i)\, u_i(c,t), \ \ \forall i\in N-1, \ \forall t_i\in T_i, \ \forall s_i\in T_i.$$

<u>Theorem</u> A utility allocation w_1 is has no viable alternatives iff there exist λ and α such that

$\lambda_i(t_i)\geq 0, \ \alpha_i(s_i|t_i) \geq 0, \ \forall i\in N, \ \forall t_i\in T_i, \ \forall s_i\in T_i$, and $\lambda_1(t_1) > 0, \ \forall t_1\in T_1$,

$$[(\lambda_1(t_1) + \sum_{s_1\in T_1} \alpha_1(s_1|t_1))w_1(t_1) - \sum_{s_1\in T_1} \alpha_1(t_1|s_1)\, w_1(t_{-1},s_1)]/p_1(t_1)$$

$$\geq \sum_{t_{-1}\in T_{-1}} P_1(t_{-1}|t_1) \max_{c\in C} \sum_{i\in N} [(\lambda_i(t_i) + \sum_{s_i\in T_i} \alpha_i(s_i|t_i))u_i(c,t) - \sum_{s_i\in T_i} \alpha_i(t_i|s_i)u_i(c,(t_{-i},s_i))]/p_i(t_i), \quad \forall t_1\in T_1.$$

7(b)

Lemma: If

$$[(\lambda_1(t_1) + \sum_{s_1 \in T_1} \alpha_1(s_1|t_1))w_1(t_1) - \sum_{s_1 \in T_1} \alpha_1(t_1|s_1)\, w_1(t_{-1},s_1)]/p_1(t_1)$$

$$\geq [(\lambda_1(t_1) + \sum_{s_1 \in T_1} \alpha_1(s_1|t_1))\hat{w}_1(t_1) - \sum_{s_1 \in T_1} \alpha_1(t_1|s_1)\, \hat{w}_1(t_{-1},s_1)]/p_1(t_1), \ \forall t_1 \in T_1$$

then $w_1(t_1) \geq \hat{w}_1(t_1) \ \forall t_1 \in T_1$.

(Method of proof: If not, then consider the sum, over all t_1 such that $w_1(t_1) < \hat{w}_1(t_1)$, of the given inequalities multiplied by $p_1(t_1)$...)

8. Neutral optimal solutions for the informed principal

Let us say that a utility allocation $w_1 = (w_1(t_1))_{t_1 \in T_1}$ for the principal's types is strongly inhibitive iff it has no viable alternatives. In terms of the virtual utility hypothesis, the above theorem asserts that a strongly inhibitive allocation w_1 gives, in some virtual-utility scales, each type of the principal a virtual payoff that not less than the maximal expected sum of all players' virtual payoffs that this type can generate.

Let us say that an allocation is inhibitive iff it is a limit of strongly inhibitive allocations.

A mechanism μ is a neutral optimum for the informed principal, in the sense of my paper in Econometrica 1983, iff μ is incentive-compatible and there exists some inhibitive allocation w_1 such that $U_1(\mu|t_1) \geq w_1(t_1)$, $\forall t_1 \in T_1$.

Such μ is plausibly inscrutable by the criterion that "each type t_1 should get at least $w_1(t_1)$", and an infinitesimal perturbation of this criterion rules out any other mechanism that some t_1 prefer.

Theorem A neutral optimum exists.

How compelling is this solution concept? I characterized it in 1983 as the smallest solution concept satisfying certain properties (strong solutions, extensions, domination, continuity). This minimality means to me that, if you can find a more compelling theory of informed mechanism selection, then I expect that your theory will include my neutral optima as a subset.

Example: Suppose that 1=firm, 2=union. 14 workers are available in the union.

The firm has all private information, has two types $T_1 = \{\text{high, low}\}$.

Firm's type	Probability	Firm's profit per worker	Worker's opportunity cost
high	1/2	170	110
low	1/2	120	100

(Isomorphic to a linear transformation of earlier example.)

Incentive-efficient plane supported by $\lambda_1(\text{high})=0.3$, $\lambda_1(\text{low})=0.7$, $\lambda_2=1$, $\alpha_1(\text{low}|\text{high}) = 0.2$, and $\alpha_1(\text{high}|\text{low}) = 0$. The firm's virtual profit per worker is then

$v(\text{high}) = [(0.3+0.2)*170 - 0*120]/0.5 = 170$, $v(\text{low}) = [(0.7+0)*120 - 0.2*170]/0.5 = 100$.

Maximum sum of virtual utilities is $14*(170-110)=840$ for high, $14*(100-100)=0$ for low.

$[(0.3+0.2)*w_1(\text{high}) - 0*w_1(\text{low})]/0.5 = 840$, $[(0.7+0)*w_1(\text{low}) - 0.2*w_1(\text{high})]/0.5 = 0$

yields warranted claims: $w_1(\text{high})=840$, $w_1(\text{low})=240$. These expected payoffs are achieved by the mechanism: hire all 14 workers for \$110 each if high, hire 12 workers for \$100 each if low.

9. Interim bargaining and virtual equity

Consider a Bayesian bargaining problem with N={1,2} and independent types, where each player can get 0 alone.

Let $\beta(1)$ and $\beta(2) = 1-\beta(1)$ denote r elative bargaining power of players 1 and 2.

The ex ante (nonsymmetric) Nash bargaining solution would select an incentive compatible μ to maximize $U_1(\mu)^{\beta(1)}\, U_2(\mu)^{\beta(2)}$, where $U_i(\mu) = \sum_{t_i \in T_i} p_i(t_i) U_i(\mu \mid t_i)$.

To study interim bargaining, let us say that an interim allocation $w=(w_i(t_i))_{i\in N, t_i \in T_i}$ is a vector of warranted claims iff there exist vectors λ and α satisfying the following virtual-equity conditions:

$$\lambda_i(t_i) > 0, \quad \alpha_i(s_i \mid t_i) \geq 0, \quad \forall i \in N=\{1,2\}, \quad \forall t_i \in T_i, \quad \forall s_i \in T_i,$$

$$\left[\left(\lambda_i(t_i) + \sum_{s_i \in T_i} \alpha_i(s_i \mid t_i)\right) w_i(t_i) - \sum_{s_i \in T_i} \alpha_i(t_i \mid s_i)\, w_i(t_{-i}, s_i)\right] \big/ p_i(t_i)$$

$$\geq \beta(i) \sum_{t_{-i} \in T_{-i}} p_{-i}(t_{-i}) \max_{c \in C} \sum_{j \in \{1,2\}} \left[\left(\lambda_j(t_j) + \sum_{s_j \in T_j} \alpha_j(s_j \mid t_j)\right) u_j(c,t) - \sum_{s_j \in T_j} \alpha_j(t_j \mid s_j) u_j(c,(t_{-j}, s_j))\right] \big/ p_j(t_j),$$

$$\forall i \in \{1,2\}, \quad \forall t_i \in T_i.$$

In my 1984 Econometrica paper, μ is a neutral bargaining solution iff μ is incentive compatible there is a limit w of such warranted-claim vectors such that $U_i(\mu \mid t_i) \geq w_i(t_i)$ $\forall i \in N$ $\forall t_i \in T_i$.

That is, in some virtual-utility scale, each type of each player i has a warranted claim to at least a $\beta(i)$ share of the maximal sum of virtual payoffs that could be expected given his type.

This solution is derived from axioms, including: "If a strong solution exists for each player-as-principal and randomizing between these two solutions is incentive efficient, then doing each i's solution with prob'y $\beta(i)$ is a bargaining solution." (Also: extensions, domination, continuity.)

10. Example to compare ex-ante and interim bargaining solutions

Suppose that 1=firm, 2=union. 14 workers are available in the union.

The firm has all private information, has two types T_1 = {high, low}.

Firm's type	Probability	Firm's profit per worker	Worker's opportunity cost
high	1/2	170	110
low	1/2	120	100

Incentive-efficient plane supported by $\lambda_1(\text{high})=0.3$, $\lambda_1(\text{low})=0.7$, $\lambda_2=1$, $\alpha_1(\text{low}|\text{high}) = 0.2$, and $\alpha_1(\text{high}|\text{low}) = 0$. The firm's virtual profit per worker is then

$v(\text{high}) = [(0.3+0.2)*170 - 0*120]/0.5 = 170$, $v(\text{low}) = [(0.7+0)*120 - 0.2*170]/0.5 = 100$.

Maximum sum of virtual utilities is $14*(170-110)=840$ for high, $14*(100-100)=0$ for low.

Let $\beta(1)$ denote the firm's relative bargaining power, $\beta(2) = 1-\beta(1)$. Warranted claims satisfy:

$[(0.3+0.2)*w_1(\text{high}) - 0*w_1(\text{low})]/0.5 = \beta(1)*840$,

$[(0.7+0)*w_1(\text{low}) - 0.2*w_1(\text{high})]/0.5 = \beta(1)*0$,

$w_2 = \beta(2)*(840*1/2 + 0*1/2)$.

Interim warranted claims are $w_1(\text{high}) = \beta(1)*840$, $w_1(\text{low}) = \beta(1)*240$, $w_2 = (1-\beta(1))*420$.

For any $\beta(1)$, these interim expected payoffs are achieved by the neutral bargaining solution $\mu_{I,\beta}$:

hire all 14 workers for $\$170-\beta(1)*60 = \$110+\beta(2)*60$ each if high,

hire $12*\beta(1)$ workers for \$100 each if low.

The ex ante (nonsymmetric) Nash bargaining solution $\mu_{A,\beta}$ depends on $\beta(1)$ as follows:

For $\beta(1)\geq 0.625$: always hire all 14 workers for $\$145-40*\beta(1)$ each.

For $\beta(1) \in [0.5, 0.625]$: always hire all 14 workers for \$120 each.

For $\beta(1)< 0.5$: hire all 14 workers for $\$170-\beta(1)*100$ if high,

hire $\beta(1)*28$ workers for \$120 each if low.

The ex-ante bargaining solution cannot generate ex-post inefficient underemployment unless the workers have substantial bargaining power (=>"blame the unions for unions for layoffs"), but then it yields \$0 gains for the low-type firm (which is less vulnerable to strikes).

The interim bargaining solution is better for the low-type firm, but it must signal its type by costly ex-post inefficient underemployment, even when the firm has all the bargaining power.

$U_1(\mu_{A,\beta}|\text{high})>U_1(\mu_{I,\beta}|\text{high})$, $U_1(\mu_{A,\beta}|\text{low})<U_1(\mu_{I,\beta}|\text{low})$. $U_2(\mu_{A,\beta})\geq U_1(\mu_{I,\beta})$ (= if $\beta(1)\leq 0.5$).

10(b)

Firm's type	Probability	Firm's profit per worker	Worker's opportunity cost
high	1/2	170	110
low	1/2	120	100

Interim neutral optimum for firm is: the high type hires 14 workers for $110 each,
the low type hires 12 workers for $100 each, and so the workers break even with either type.
Ex ante optimum for the firm is: both types hire 14 workers for $105 each,
and so the workers lose $5 each with high type, gain $5 each with low type.
But only high type prefers ex-ante optimum over interim neutral optimum!

The core requires that the firm satisfy an ex-post virtual participation constraint, which implies that the low type cannot pay more than $100 per worker.
To see why, consider the incentive-efficient mechanism in which money profits are shared equally in each state:
the high type hires 14 workers for $140 each, the low type hires 7 workers for $110 each.
Firm's expected allocation is then w(high) = 14*30, w(low)= 7*10.
This would be vulnerable to the following random blocking plan, for any ε between 0 and 1/2:
with proby ε ask both, if high then 14 work for $138 each, if low then 14 work for $109 each;
with prob'y ε only ask firm, if high then no block (get w), if low then firm blocks alone (gets $0).
So if asked, union expects 0.5*14*28+0.5*14*9 = $259 (> $231= 0.5*14*30+0.5*7*10).
High firm, if asked, expects 0.5*14*32+0.5*w(high) from honesty, which is incentive compatible (> 0.5*14*61+0.5*0) when w(high) = 14*30.
Low firm, if asked, expects 0.5*14*11+0.5*0, which is better than w(low) = 7*10,
and also better than hiring for $138 (>$120) with a lie.

11. Virtual utility in strategic-form games with strategic incentive constraints

Above we have considered games with informational incentive constraints (adverse selection) but no strategic incentive constraints (moral hazard). Consider now the design of communication mechanisms with strategic incentive constraints but no informational incentive constraints.

Let $\Gamma = (N, (C_i)_{i\in N}, (u_i)_{i\in N})$ be a finite strategic-form game, where $C_i = \{\text{i's pure strategies}\}$ and $u_i: C \to \mathbb{R}$ is i's utility function. Notation: $C = \times_{i\in N} C_i$, $C_{-i} = \times_{j\in N-i} C_j$, $c = (c_{-i}, c_i)$.

Communication allows any correlated equilibrium $\mu \in \Delta(C)$ satisfying the strategic incentive constraints $\sum_{c_{-i}\in C_{-i}} \mu(c)u_i(c) \geq \sum_{c_{-i}\in C_{-i}} \mu(c)u_i(c_{-i},d_i)$, $\forall i\in N$, $\forall c_i\in C_i$, $\forall d_i\in C_i$.

Such μ is incentive-efficient iff it is not Pareto-dominated by any other correlated equilibrium.

Any incentive-efficient μ must maximize some weighted sum of the players' expected utilities $\sum_{i\in N} \lambda_i U_i(\mu)$, where $U_i(\mu) = \sum_{c\in C} \mu(c)u_i(c)$, subject to the strategic incentive constraints above.

The Lagrangean for this optimization problem can be written

$$\mathbf{L}(\mu,\lambda,\alpha) = \sum_{i\in N} \lambda_i U_i(\mu) + \sum_{i\in N} \sum_{c_i\in C_i} \sum_{d_i\in C_i} \alpha_i(d_i|c_i)\Big(\sum_{c_{-i}\in C_{-i}} \mu(c)u_i(c) - \sum_{c_{-i}\in C_{-i}} \mu(c)u_i(c_{-i},d_i)\Big)$$

$$= \sum_{c\in C} \mu(c) \sum_{i\in N} v_i(c,\lambda,\alpha),$$

where $v_i(c,\lambda,\alpha) = u_i(c) + \sum_{d_i\in C_i} \alpha_i(d_i|c_i)\big(u_i(c) - u_i(c_{-i},d_i)\big)$ may be called i's virtual utility of c.

The Lagrangean conditions for an optimum then are

$$\sum_{c\in C} \mu(c) \sum_{i\in N} v_i(c,\lambda,\alpha) = \max_{c\in C} \sum_{i\in N} v_i(c,\lambda,\alpha),$$

$$\alpha_i(d_i|c_i)\Big(\sum_{c_{-i}\in C_{-i}} \mu(c)u_i(c) - \sum_{c_{-i}\in C_{-i}} \mu(c)u_i(c_{-i},d_i)\Big) = 0, \ \forall i\in N, \ \forall c_i\in C_i, \ \forall d_i\in C_i.$$

We may say that a strategy d_i jeopardizes another strategy c_i for player i, in the incentive-efficient correlated equilibrium μ, iff the constraint that i should not be tempted to do d_i when c_i is recommended is binding in μ and has a positive Lagrange multiplier $\alpha_i(d_i|c_i)$.

So the incentive-efficient μ always maximizes the sum of the players' virtual utilities, where i's virtual utility of c differs from i's actual utility by exaggerating the difference from what i could get by unilaterally deviating to other strategies that jeopardize c_i.

Existence of correlated equilibrium means that the following inequalities have a solution:

$$\sum_{c_{-i}\in C_{-i}} \mu(c)(u_i(c) - u_i((c_{-i},d_i)) \geq 0 \ \ \forall i \ \forall c_i\in C_i \ \forall d_i\in C_i, \ \mu(c)\geq 0 \ \forall c\in C, \text{ and } \sum_{c\in C} \mu(c)\geq 1.$$

This holds iff the following dual inequalities have no solutions in $(\alpha,\beta)\geq 0$:

$$\beta_0 > 0, \ \beta_0 + \beta_c + \sum_{i\in N} \sum_{d_i\in C_i} \alpha_i(d_i|c_i)(u_i(c) - u_i(c_{-i},d_i)) = 0 \ \forall c\in C.$$

This impossibility can be proving by looking at stationary distributions to the α_i Markov chains.

12. Dual reduction and elementary games

Let us say that a strategic-form game Γ is elementary iff there exists some strictly incentive-compatible correlated equilibrium μ in $\Delta(C)$ such that

$$\sum_{c_{-i}\in C_{-i}} \mu(c)u_i(c) > \sum_{c_{-i}\in C_{-i}} \mu(c)u_i(c_{-i},d_i),\ \forall i\in N,\ \forall c_i\in C_i,\ \forall d_i\in C_i.$$

When Γ is elementary, problems of indifference in equilibrium (imperfection) vanish for almost all correlated equilibria (in the relative interior of the correlated-equilibrium set).

By theorems of the alternative for linear systems, Γ is not elementary if and only if there exists some dual vector α such that

$$\alpha_i(d_i|c_i) \geq 0\ ,\ \forall i\in N,\ \forall c_i\in C_i,\ \forall d_i\in C_i,\ \text{with at least one } \alpha_i(d_i|c_i) > 0,$$

$$\sum_{i\in N}\sum_{d_i\in C_i} \alpha_i(d_i|c_i)(u_i(c)-u_i(c_{-i},d_i)) \leq 0,\ \forall c\in C.$$

Given such a dual vector α, each subvector α_i can be interpreted as a matrix of transition rates in a Markov process on C_i.

This Markov process partitions C_i into a transient set and disjoint recurrent communicating sets.

On each recurrent communicating set, there is a unique α_i-stationary strategy σ_i such that

$$\sum_{c_i\in C_i} \alpha_i(c_i|d_i)\sigma_i(d_i) = \sum_{c_i\in C_i} \alpha_i(d_i|c_i)\sigma_i(c_i),\ \forall d_i.$$

Now consider the α-reduced game Γ/α in which the pure strategies for each player i are his α_i-stationary strategies in Γ. This is also a finite strategic form game, with fewer strategies.

Theorem For any dual vector α, every equilibrium of Γ/α is an equilibrium of Γ.

(This result holds both for Nash's concept of equilibrium and for correlated equilibrium).

Theorem Any finite strategic-form can be reduced by iterative dual reduction to an elementary game.

Dual reduction generalizes both elimination of dominated strategies and consolidation of unique subgame equilibrium strategies.

13. References

"Mechanism Design by an Informed Principal," Econometrica 51:1767-1797 (1983).

With M. Satterthwaite, "Efficient Mechanisms for Bilateral Trading," Journal of Economic Theory 29:265-281 (1983).

"Two-Person Bargaining Problems with Incomplete Information" Econometrica 52:461-487 (1984).

"Cooperative Games with Incomplete Information," International Journal of Game Theory 13:69-96 (1984).

"Bayesian Equilibrium and Incentive Compatibility" in Social Goals and Social Organization, edited by L. Hurwicz, D. Schmeidler, and H. Sonnenschein (Cambridge U. Press, 1985).

"Analysis of Two Bargaining Problems with Incomplete Information," in Game-Theoretic Models of Bargaining, edited by A. Roth (Cambridge U. Press, 1985).

"Credible Negotiation Statements and Coherent Plans," Journal of Economic Theory 48:264-303 (1989).

"Chapter 10: Cooperation under Uncertainty," in Game Theory (Harvard U. Press, 1991).

"Fictitious-Transfer Solutions in Cooperative Game Theory" in Rational Interaction, edited by R Selten (Springer-Verlag, 1992).

"Sustainable Matching Plans with Adverse Selection," Games and Economic Behavior 9:35-65 (1995).

"Dual Reduction and Elementary Games," Games and Economic Behavior 21:182-202 (1997).

By other authors:

E. Maskin and J. Tirole, "The Principal Agent Relationship with an Informed Principal: The Case of Private Values," Econometrica 58:379-409 (1990).

F. Forges, E. Minelli, and R. Vohra, "Incentives and the Core of an Exchange Economy: a Survey," to appear in Journal of Mathematical Economics.

http://home.uchicago.edu/~rmyerson/research/virtual.pdf