COOPERATIVE GAME THEORY: CORE AND SHAPLEY VALUE

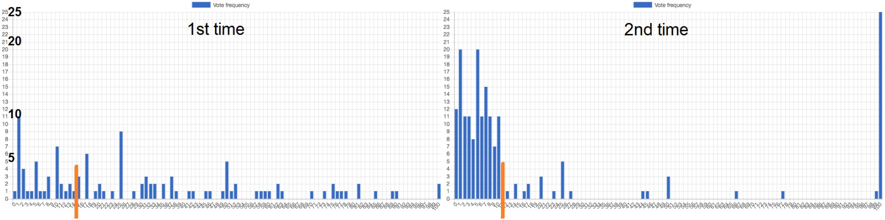
Heinrich H. Nax & Bary S. R. Pradelski

hnax@ethz.ch bpradelski@ethz.ch

February 27, 2017: Lecture 2

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

What you did last week...



There appear to be traces of both, more strategic behavior (more mass around 0 and 1) and more collusive/cooperative behavior (the 100s), when you played the game the second time.

The **two** branches of game theory

Non-cooperative game theory

- No binding contracts can be written
- Players are individuals
- Nash equilibrium

Cooperative game theory

- Binding contract can be written
- Players are individuals and coalitions of individuals
- Main solution concepts:
  - Core
  - Shapley value
- **The focus of today!**

Reminder: the ingredients of a **noncooperative game**

- **Players**: $N = \{1, 2, ..., n\}$
- **Actions / strategies**: each player chooses s_i from his own finite strategy set; S_i for each $i \in N$
  - **Outcome**: resulting strategy combination: $s = (s_1, ..., s_n) \in (S_i)_{i \in N}$
- **Payoff outcome**: payoffs $u_i = u_i(s)$

The Theory of Games and Economic Behavior (1944)

John von Neumann (1903-1957) and Oskar Morgenstern (1902-1977)

The ingredients of a **cooperative game**

- **Population of players**: $N = \{1, 2, ..., n\}$ (finite)
- **Coalitions**: $C \subseteq N$ form in the population and become players
  - resulting in a coalition structure $\rho = \{C_1, C_2, ..., C_k\}$
- **Payoffs**: *–we still need to specify how–* payoffs $\phi = \{\phi_1, ..., \phi_n\}$ come about:
  - something like this: $\phi_i = \phi(\rho,$“**sharing rule**”$)$

Cooperative games in **characteristic function form** (CFG)

- **The game:** A CFG defined by 2-tuple $G(v, N)$
- **Players:** $N = 1, 2, ..., n$ (finite, fixed population)
- **Coalitions**: *disjoint* $C \subseteq N$ form resulting in a coalition structure/ partition ρ
  - $\emptyset$ is an *empty coalition*
  - N is the *grand coalition*
  - The set of all coalitions is 2^N
  - $\boldsymbol{\rho}$ is the set of all partitions
- **Characteristic function**: v is the characteristic function form that assigns a *worth* $v(C)$ to each coalition
  - v: $2^N \rightarrow R$
  - (and $v(\emptyset) = 0$)

3-player example

- N=1,2,3
- v(i)=0
- v(1,2)=v(1,3)=0.5
- v(2,3)=0
- v(N)=1

"Transferable utility" and **feasibility**

- **The game:** CFG defined by 2-tuple $G(v, N)$
- **Outcome:** *Coalition structure*
  - *partition* $\rho = \{C_1, C_2, ..., C_k\}$ and
  - *payoff allocation*/imputation $\phi = \{\phi_1, ..., \phi_n\}$
- Importantly, v(C) can be "shared" amongst $i \in C$ (**transfer of utils**)!
- **Feasibility**: in each C, $\sum_{i \in C} \phi_i \leq v(C)$

3-player example: some feasible outcomes

- Outcome 1: {(1,2),3} and {(0.25,0.25),0}
- Outcome 2: {N} and {0.25,0.25,0.5}
- Outcome 3: {N} and {0.8,0.1,0.1}

Superadditivity assumption

Superadditivity

If two coalitions C and S are disjoint (i.e. $S \cap C = \emptyset$), then $v(C) + v(S) \leq v(C \cup S)$

i.e. "mergers of coalitions weakly improve their worths"

- Superadditivity implies *efficiency* of the grand coalition: for all $\rho \in \boldsymbol{\rho}$, $v(N) \geq \sum_{C \in \rho} v(C)$.
- In our example: $v(N) > v(1,2) = v(1,3) > v(2,3) = v(1) = v(2) = v(3)$.

The Core (Gillies 1959)

The Core

The Core of a superadditive $G(v, n)$ consists of all outcomes where the *grand coalition* forms and payoff allocations ϕ^* are

Pareto-efficient: $\sum_{i \in N} \phi_i^* = v(N)$

Unblockable: for all $C \subset N$, $\sum_{i \in C} \phi_i^* \geq v(C)$

- *individual rational*: $\phi_i^* \geq v(i)$ for all i
- *coalitional rational*: $\sum_{i \in C} \phi_i^* \geq v(C)$ for all C

3-player example

- Outcome 1: {(1,2),3} and {(0.25,0.25),0}
- Outcome 2: {N} and {0.25,0.25,0.5}
- Outcome 3: {N} and {0.8,0.1,0.1}

Properties of the Core

- A system of weak linear inequalities defines the Core, which is therefore closed and convex.
- The core can be
  - empty
  - non-empty
  - large

Core empty

- $v(i) = 0$
- $v(i,j) = 0.9$
- $v(N) = 1$

Core unique

- $v(i) = 0$
- $v(i,j) = 2/3$
- $v(N) = 1$

Core large

- $v(i) = v(i,j) = 0$
- $v(N) = 1$

Bondareva-Shapley Theorem

Bondareva 1963 and Shapley 1967

The Core of a cooperative game is *nonempty* if and only if the game is *balanced*.

Balancedness:

- *Balancing weight:* Let $\alpha(C) \in [0, 1]$ be the balancing weight attached to any $C \in 2^N$
- *Balanced family:* A set of balancing weights α is a balanced family if, for every i, $\sum_{C \in 2^N : i \in C} \alpha(C) = 1$
- Balancedness in a superadditive game then requires that, for all balanced families,
  - $v(N) \geq \sum_{C \in 2^N} \alpha(C) v(C)$

Limitations of the Core

1. Core empty

- $v(i) = 0$
- $v(i,j) = 5/6$
- $v(N) = 1$

2. Core non-empty but very inequitable (1, 0, 0)

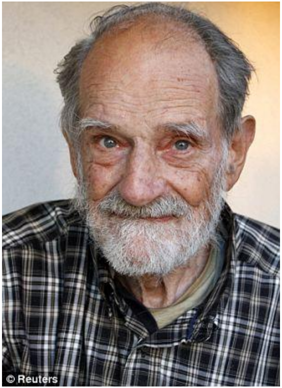
- $v(i) = v(2, 3) = 0$
- $v(N) = v(1, 2) = v(1, 3) = 1$

3. Core large (any split of 1)

- $v(i) = v(i,j) = 0$
- $v(N) = 1$

So is the Core a *descriptive* or a *prescriptive/normative* solution concept?

What about an explicitly *normative* solution concept?



Lloyd Shapley (1923-2016)

Shapley value (Shapley 1953)

Axioms. Given some $G(v, N)$, an acceptable allocation/value $x^*(v)$ should satisfy

- **Efficiency.** $\sum_{i \in N} x_i^*(v) = v(N)$
- **Symmetry**. if, for any two players i and j, $v(S \cup i) = v(S \cup j)$ for all S not including i and j, then $x_i^*(v) = x_j^*(v)$
- **Dummy player**. if, for any i, $v(S \cup i) = v(S)$ for all S not including i, then $x_i^*(v) = 0$
- **Additivity**. If u and v are two characteristic functions, then $x^*(v + u) = x^*(v) + x^*(u)$

Shapley's characterization

The unique function satisfying all four axioms for the set of all games is

$$\phi_i(v) = \sum_{S \in N, i \in S} \frac{(|S|-1)!(n-|S|)!}{n!}[v(S) - v(S \setminus \{i\})]$$

So what does this function mean?

Shapley value

The Shapley value pays each player his *average marginal contributions*:

- For any S: $i \in S$, think of the *marginal contribution* $MC_i(S) = v(S) - v(S \setminus i)$.
- And of $\sum_{S \in N, i \in S} \frac{(|S|-1)!(n-|S|)!}{n!}$ as some kind of "average" operator (more detail later).

Then,

$$\phi_i(v) = \text{average}\ (MC_i(S))$$

An alternative characterization

Young (1985): a set of equivalent axioms is

- **Efficiency**. $\sum_{i \in N} x_i^*(v) = v(N)$
- **Symmetry**. if, for any two players i and j, $v(S \cup i) = v(S \cup j)$ for all S not including i and j, then $x_i^*(v) = x_j^*(v)$
- **Monotonicity**. If u and v are two characteristic functions and, for all S including i, $u(S) \geq v(S)$, then $x_i^*(u) \geq x^*(v)$

A more attractive set of axioms...

1. Core empty

- $v(i) = 0$
- $v(i,j) = 5/6$
- $v(N) = 1$

Shapley value

$(1/3, 1/3, 1/3)$

2. Core non-empty but very inequitable (1, 0, 0)

- $v(i) = v(2,3) = 0$
- $v(N) = v(1,2) = v(1,3) = 1$

Shapley value

$(4/6, 1/6, 1/6)$

3. Core large (any split of 1)

- $v(i) = v(i,j) = 0$
- $v(N) = 1$

Shapley value

$(1/3, 1/3, 1/3)$

Room-entering story (Roth 1983)

Average MC in this sense...

- $N = \{1, 2, ..., n\}$ players enter a room in some order.
- Whenever a player enters a room, and players $S \setminus i$ are already there, he is paid his marginal contribution $MC_i(S) = v(S) - v(S \setminus i)$.
- Suppose all $n!$ orders are equally likely.
- Then there are $(s-1)!$ different orders in which these players in $S \setminus i$ can precede i
- and $(n-s)!$ order in which the others may follow
- hence, a total of $(s-1)!(n-s)!$ orders for that case of the $n!$ total orders.
- Ex ante, the payoff of a players is $\sum_{S \in N, i \in S} \frac{(s-1)!(n-s)!}{n!} MC_i(S)$ – the Shapley value.

Relationship between the Core and the Shapley value

Put simply, none...

- the Shapley value is normative
- the Core is something else (hybrid)
- when the Core is non-empty, the SV may lie inside or not
- when the Core is empty, the SV is still uniquely determined

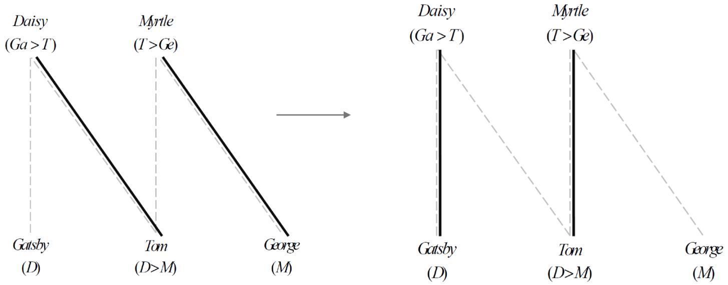
Other cooperative models

Non-transferable-utility cooperative game

- **As before**: CFG defined by 2-tuple $G(v, N)$
- **Outcome**: *partition* $\rho = \{C_1, C_2, ..., C_k\}$ *directly* (w/o negotiating how to share) implies a payoff allocation/imputation – $\phi_i = f_i(C_i)$
- There are no side-payments and the worth of a coalition cannot be (re-)distributed.

Agents have preferences over coalitions.

Stable Marriage/Matching problem



Stable Marriage/Matching problem

2-sided market

- *Men* $M = \{m_1, ..., m_n\}$ on one side, *women* $W = \{w_1, ..., w_n\}$ on the other.
- Each m_i: preferences (e.g. $w_1 \succ w_2 \succ ... \succ w_n$) over women
- Each w_i: preferences (e.g. $m_n \succ m_1 \succ ... \succ m_{n-1}$) over men

We want to establish a stable matching: forming couples (man-woman) such that there exists no alternative couple where both partners prefer to be matched with each other rather than with their current partners.

Deferred acceptance

Gale-Shapley 1962

For any marriage problem, one can make all matchings stable using the deferred acceptance algorithm.

Widely used in practice (e.g. Roth & Sotomayor 1990, Roth et al. ...):

- Resource allocations/doctor recruitment for hospitals
- Organ transplantations
- School admissions/room allocation
- Assigning users to servers in distributed Internet services
- ...

DA "pseudo-code"

Initialize : all $m_i \in M$ and all $w_i \in W$ are *single*.

Engage : Each single man $m \in M$ *proposes* to his *preferred woman* w to whom he has *not yet proposed*.

- If w is single, she will become *engaged* with her *preferred proposer*.
- Else w is already engaged with m'.
  - If w prefers her preferred proposer m over her current engagement m', then (m, w) become engaged and m' becomes single.
  - Else (m', w) remain engaged.
- All proposers who do not become engaged remain single.

Repeat : If there exists a single man after **Engage**, repeat **Engage**; Else move to **Terminate**.

Terminate : *Marry* all engagements.

Proof sketch

Trade up : Women can *trade up* until every woman (hence also every man) is engaged, which is when they all get married.

Termination : No singles can remain, because every man would eventually propose to every woman as long as he remains single, and every single woman, once proposed to, becomes engaged.

Termination with stability?

Proof sketch

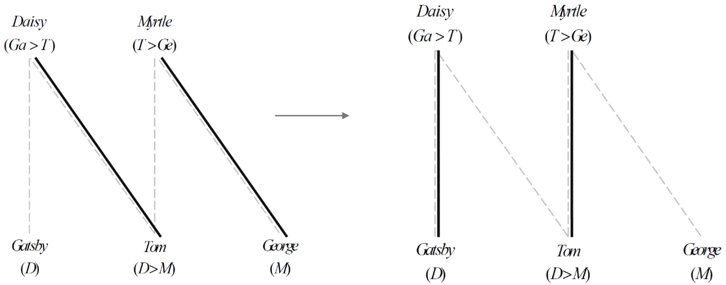
Stability : *The resulting matching is stable.*

Proof : Suppose the algorithm terminates so that there exists a pair (m, w) whose partners are engaged to $w' \neq w$ and $m' \neq m$ respectively.

Claim : It is not possible for both m and w to prefer each other over their engaged partner. because

- If m prefers w over w', then he proposed to w before he proposed to w'. At that time,
    - *Case 1:* If w got engaged with m, but did not marry him, then w must have traded up and left m for someone she prefers over m, and therefore cannot prefer m over m'.
    - *Case 2:* Else, if w did not get engaged with m, then she was already with someone she prefers to m at that time, and can therefore not prefer m over m'.
- Hence, either m prefers w' over w, or w prefers m' over m.

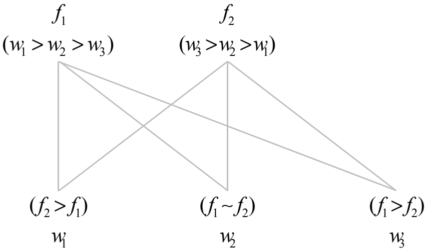
Back to the Great Gatsby



From NTU to TU matching

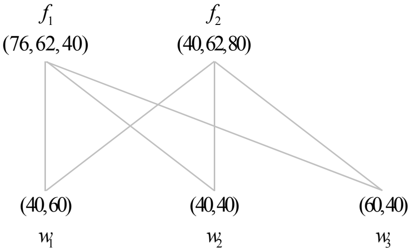
Non-transferable utility

Pairwise stable outcomes always exist (Gale & Shapley 1962)



Transferable utility

Pairwise stable and *optimal* outcomes (core) always exist (Shapley & Shubik 1972)

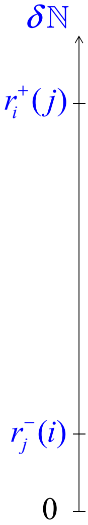


Firms/workers and their willingness to pay/accept

Firms $i \in F$ and workers $j \in W$ look for partners ($|F| = |W| = N$)

- Firm i is willing to pay at most $r_i^+(j) \in \delta\mathbb{N}$ to match worker j
- Worker j is willing to accept at least $r_j^-(i) \in \delta\mathbb{N}$ to match firm i

$\delta > 0$ is the minimum unit ('dollars')

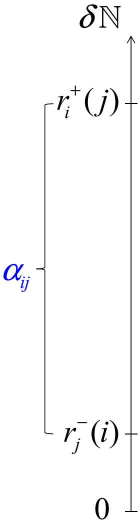


The resulting match values $\boldsymbol{\alpha}$

The match value for the pair (i, j) is

$$\alpha_{ij} = (r_i^+(j) - r_j^-(i))_+$$

Define matrix $\boldsymbol{\alpha} = (\alpha_{ij})_{i \in F, j \in W}$



Match values and the assignment matrix define the game

Let $\mathbf{A} = (a_{ij})_{i \in F, j \in W}$ be the assignment such that *each agent has at most one partner* and

$$\text{if } (i,j) \text{ is } \begin{cases} \text{matched} & \text{then} \quad a_{ij} = 1 \\ \text{unmatched} & \text{then} \quad a_{ij} = 0 \end{cases}$$

Prices and payoffs

Price. Any price π_{ij} between the willingnesses to “buy” and “sell” of firm and worker is individual rational:

$$r_i^+(j) \geq \pi_{ij} \geq r_j^-(i)$$

Payoff. Given prices, payoffs are

The payoff to firm i is $\phi_i = r_i^+(j) - \pi_{ij}$

The payoff to worker j is $\phi_j = \pi_{ij} - r_j^-(i)$

If an agent i (no matter whether firm or worker) is single $\phi_i = 0$.

Assignment **A** and ***payoffs*** ϕ define an ***outcome*** of the game.

Solution concepts

Optimality. $\mathbf{A}$ is *optimal* if it maximizes total payoff:

$$\sum_{(i,j)\in F\times W} a_{ij} \cdot \alpha_{ij}$$

Pairwise stability. ϕ is *pairwise stable* if for all (i,j) matched

$$\phi_i + \phi_j = \alpha_{ij}$$

and for all k, l not matched

$$\phi_k + \phi_l \geq \alpha_{kl}$$

Core. The *Core* of an assignment game consists of all outcomes, $[\mathbf{A}, \phi]$, such that $\mathbf{A}$ is an optimal assignment and ϕ is pairwise stable.

Shapley-Shubik 1972

The Core of the assignment game is nonempty.

Example

Shapley-Shubik's **house-trading game**:

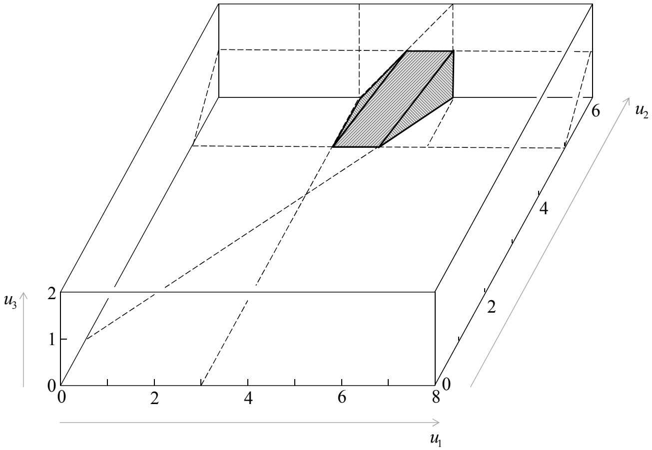
House	Sellers willingness to accept $q^+_{1j} = q^+_{2j} = q^+_{3j}$	Buyers' willingness to pay p^+_{i1}	p^+_{i2}	p^+_{i3}
1	$18,000$	$23,000$	$26,000$	$20,000$
2	$15,000$	$22,000$	$24,000$	$21,000$
3	$19,000$	$21,000$	$22,000$	$17,000$

These prices lead to the following match values, α_{ij} (units of 1000), where sellers are occupying rows and buyers columns:

$$\alpha = \begin{pmatrix} 5 & \mathbf{8} & 2 \\ 7 & 9 & \mathbf{6} \\ \mathbf{2} & 3 & 0 \end{pmatrix}$$

Unique optimal matching is shown in **bold** numbers.

Figure: Core imputation space for the sellers.



How to “implement” a core outcome?

Centralized market: Algorithm by central planner yields stable and optimal outcomes, e.g.,

- *deferred acceptance* (Gale & Shapley 1962) for the NTU game
- solving a LP-dual (Shapley & Shubik 1972) for the TU game
- used in practice: hospital matching, transplantations, school admissions, internet servers, communication networks, etc.

Decentralized market: Players trade and (re-)match repeatedly over time.

- random *blocking paths* (Roth & Vande Vate 1990) converge to the NTU core
- random *(re-)trade and price adjustment* (Nax & Pradelski 2015) converge to the TU core
- used in practice: “free” markets, internet auctions, labor market, etc.

THANKS EVERYBODY
Bary will take over for a few lectures next week!
and keep checking the website for new materials as we progress:
http://www.coss.ethz.ch/education/GT.html