ON A THEOREM OF KALAI AND SAMET

When Do Pure Equilibria Imply a Potential Function?

Tim Roughgarden (Stanford)

THE WEIGHTED SHAPLEY VALUE

The Shapley Value

Coalitional game: set N, set function $c : 2^N \rightarrow R$

Shapley value: (equivalent definitions)

- *linear algebraic:* define for unaniminity games (basis for all games), extend by linearity
- *axiomatic:* efficiency, symmetry, dummy additivity
- *probabilistic:* expected marginal contribution w.r.t. uniformly random player ordering

The Weighted Shapley Value

[Shapley's PhD thesis, pp. 66-67, 1953]

It is easy to imagine games or game-like situations in which the symmetry assumption (d) is not appropriate, because of differences in the <u>external</u> characteristics of the players. (Internal differences are accounted for in the function v!) For example, individuals might be competing with corporations, or governments, or there might be differences in "bargaining ability", or some other skill factor. These cases might be handled by means of imputation operators based on measures other than the symmetric measure j. The effect would be to calibrate the players according to their performance in the "pure bargaining" game

The Weighted Shapley Value

Weighted Shapley value: (equivalent definitions)

- [Shapley's PhD thesis, 1953] *linear algebraic:* define for unaniminity games (proportional to weights), extend by linearity
- [Kalai/Samet, 1987] *axiomatic:* efficiency, dummy, additivity, positivity, partnership
- [Kalai/Samet, 1987] *probabilistic:* expected marginal contribution w.r.t. suitable (non-uniform) random player ordering

next

3 Probabilistic Definition of Weighted Shapley Values

Theorem: [Kalai/Samet 87] $(\phi_\omega)_i(v) = E_{P_\omega}(C_i(v,\cdot))$

I.e., weighted Shapley values (for weights $w_1,\ldots,w_n > 0$) = expected marginal contributions w.r.t. random order:

- pick a player w/probability proportional to weight, make this the *last* player
- pick next player of those left w/probability proportional to weight, make *second-to-last*
- etc.

Weight systems: handle zero-weight players recursively.

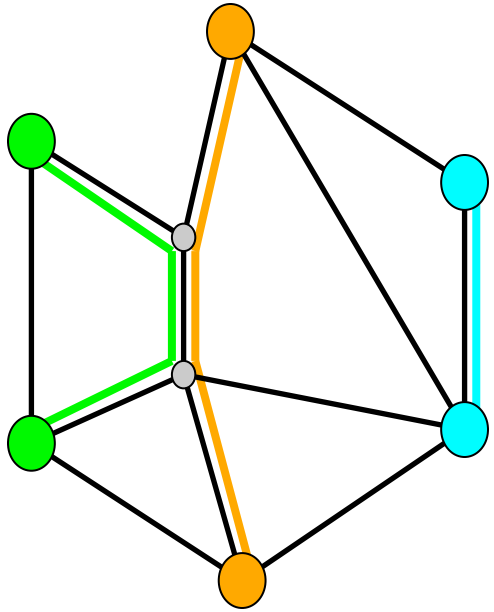
NETWORK COST-SHARING GAMES

Network Cost-Sharing

Given: graph $G = (V,E)$

- k players = vertex pairs (s_i,t_i)
- each picks an s_i-t_i path

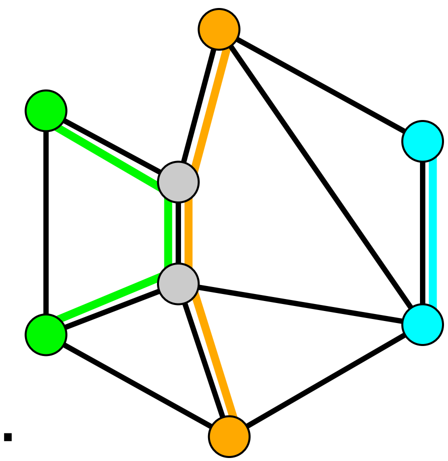
  [Anshelevich/Dasgupta/Kleinberg/Tardos 03]



Cost of outcome: number of edges used by least one player

Goal: budget-balanced method of sharing the cost (users of an edge should jointly pay 1 for it).

Symmetric Cost Sharing



Assumption: Cost of each edge shared independently.

Symmetric cost sharing:

Players using e share costs evenly:

$$c_i(P) = \sum_{e \in P} 1/k_e$$

[Anshelevich et al FOCS 04]

- players' objectives: minimize individual cost
- global objective: minimize total network cost

A Potential Function

Claim: symmetric cost sharing => special case of a congestion game [Rosenthal 73].

Potential function: let $f(S) = 1+1/2+1/3+\ldots+1/|S|$ and $P(A_1,\ldots,A_n) = \sum_r f(S_r)$.

Observe that $\Delta P = \Delta c_i$ for every player i, every possible deviation.

Corollary: best-response dynamics converges to a pure Nash equilibrium.

Stable Cost-Sharing

Definition: *cost-sharing method* χ specifies cost share $\chi(i,S) \geq 0$ for every non-empty S and i in S, subject to budget-balance: $\Sigma_{i \text{ in } S}\, \chi(i,S) = 1$.

Note: choice of χ and G = (V,E) induces a game.

Definition: χ is *stable* if, for every G=(V,E), the game induced by G and χ has a pure Nash eq.

- example: χ = symmetric cost-sharing (has a potential function and hence a PNE)

What We Want and Why

Goal: characterize the set of stable cost-sharing methods.

Applications: identify "optimal" cost-sharing method, subject to stability requirement.

- want to minimize worst-case inefficiency of equilibria ("price of anarchy/stability")
  - positive externalities [Chen/Roughgarden/Valiant 10]
    - directed graphs: symmetric cost-sharing is optimal
    - undirected graphs: "priority" cost-sharing is optimal
      - order players arbitrarily; first player present pays full cost

What We Want and Why

Goal: characterize the set of stable cost-sharing methods.

Applications: identify "optimal" cost-sharing method, subject to stability requirement.

- want to minimize worst-case inefficiency of equilibria ("price of anarchy/stability")
  - positive externalities [Chen/Roughgarden/Valiant 10]
  - negative externalities [Gkatzelis/Kollias/Roughgarden 16]
    - generalizes routing games [Rosenthal 73]
    - unweighted Shapley value is optimal

STABLE COST-SHARING VIA THE WEIGHTED SHAPLEY VALUE

Public Excludable Good

Public excludable good:

- C(S)=1 for S non-empty, C(Φ)=0
  - a.k.a. *representation game* in [Kalai/Samet 87]
  - dual of unanimity game

Weighted Shapley cost-sharing method:

- χ (i,S) := probability i is the final survivor of S.

Equivalent: associate exponential(w_i) random variable X_i with each player i. Then:

$$\chi(i,S) := \Pr[i = \mathrm{argmax}_{j \text{ in } S} X_j]$$

Potential for Weighted Shapley

Question: other stable cost-sharing methods?

Claim: [Hart/Mas-Colell 89, Monderer/Shapley 96] every method χ_w derived from a weighted Shapley value (for some w) is stable.

Potential for Weighted Shapley

Claim: [Hart/Mas-Colell 89, Monderer/Shapley 96] every method χ_w derived from a weighted Shapley value (for some $w > 0$) is stable.

Proof: Order players arbitrarily.

- define $f(S) = \Sigma_i \varphi_w(i,S_i)/w_i$ [S_i = 1st i players of S]
- lemma: f(S) independent of ordering!
- can again define $P(A_1,\ldots,A_n) = \Sigma_r f(S_r)$ so that $\Delta P = \Delta c_i/w_i$ for every deviation by every i

Proof of Lemma

Lemma: [Hart/Mas-Colell 89] the function $f(S) = \Sigma_i \; \chi_w(i,S_i)/w_i$ [where S_i = 1st i players of S] is well defined (independent of player ordering).

Original proof:: check for unanimity games (easy), extend by linearity.

Alternative:: (via [Kalai/Samet 87]) let $X_i \sim \exp(w_i)$, so $\chi_w(j,S_i) = \Pr[j = \arg\max_{l \in S_i} X_l]$. For any ordering i=1,2,...,k of the players:

$$f(S) = E[\max_{j\in S} X_j] = \sum_{i=1}^{k} E[\max_{j\in S_i} X_j - \max_{j\in S_{i-1}} X_j]$$

$$= \sum_{i=1}^{k} \Pr[i = \arg\max_{j\in S_i} X_j] \bullet E[\max_{j\in S_i} X_j - \max_{j\in S_{i-1}} X_j \mid i = \arg\max_{j\in S_i} X_j]$$

($\Pr[\ldots]$ = $\chi_w(i,S_i)$; $E[\ldots]$ = $1/w_i$)

On Potentials vs. PNE

Previous work: characterizes coalitional values that lead to potential games.

- [Monderer/Shapley 96], [Qin 96], [Ui00], [Slikker 01]

Critique: requiring a potential overly strong.

- potential function = means to an end
  - existence of PNE, convergence of better-responses
  - many non-potential games have these properties

Question: what if we only want existence of PNE?

MAIN CHARACTERIZATION

Characterization

Theorem: [Chen/Roughgarden/Valiant 10] a cost-sharing method χ is stable *if and only if* it corresponds to a weighted Shapley value.

- recall: χ is *stable* if, for every G=(V,E), the game induced by G and χ has a pure Nash equilibrium
- general weight systems allowed

Thus: guaranteed existence of potential ⇔ guaranteed convergence of best-response dynamics ⇔ guaranteed existence of PNE!

Taste of Proof

1st Milestone: if χ is a stable cost-sharing method, then χ is *monotone*: χ (i,S) only decreases with S.

Step 1: failures of monotonicity are symmetric (i makes j worse off **=>** converse also holds).

- basic reason: else can encode matching pennies

Step 2: no (symmetric) failures of monotonicity.

- basic reason: otherwise contradict budget-balance (sum of all cost shares fixed)

Generalization

Theorem: [Gopalakrishnan/Marden/Wierman 14] for *every* cost function (not just public excludable goods), a cost-sharing method χ is stable *if and only if* it corresponds to a weighted Shapley value.

- χ is *stable* if, for every G=(V,E), with cost function C(.) on each edge, the game induced by G and χ has a pure Nash equilibrium
- χ defines cost share χ(i,S) for every non-empty S and i in S, subject to budget-balance: $\Sigma_{i \text{ in } S}\, \chi(i,S) = C(S)$.

HAPPY BIRTHDAY EHUD!