How to Cite:
Vimala, K., & Bharathi, S. (2022). Cooperative solutions for TU and NTU stochastic game-theoretical approach. *International Journal of Health Sciences*, *6*(S3), 431–455. https://doi.org/10.53730/ijhs.v6nS3.5532

Cooperative Solutions for TU and NTU Stochastic Game-Theoretical Approach

K. Vimala
Assistant Professor, Department of Mathematics Government Arts and Science College, Kadaladi-623703, Ramanathapuram-District, Tamil Nadu, India
Email: vimalamaths78@gmail.com

S. Bharathi
Assistant Professor & Research Supervisor, Department of Mathematics, Bharathiar University PG Extension and Research Center, Erode, Tamil Nadu, India
Email: bharathikamesh6@gmail.com

***Abstract*---**The player's payoff in stochastic games is a stochastic variable. In most studies, predicted payout is referred to as a payoff, implying that the players are risk averse. However, there may be risk-averse players that consider "risk" while calculating their stochastic payoffs. We offer a model of stochastic games using mean-variance reward functions, which are the sum of expectation and standard deviation multiplied by a coefficient indicating a player's risk aversion. By adopting a maxmin technique to define the characteristic function, we may create a cooperative version of a stochastic game with mean-variance preferences. This paper contains an analysis of a cooperative solution concept of stochastic games for both TU (transferable utility) and NTU (non-transferable utility) games. We proposed a cooperative game solution method of extension of two-person zero sum strategic game to n-agent stochastic stage game using both Shapley and Banzhaf index functions. Results are proved for both discounted and undiscounted payoff situations. Existing models of discrete-time stochastic games and ways to finding cooperative solutions in these games are extended in this study.

***Keywords*---**cooperative stochastic games, nash equilibrium, TU and NTU games, cooperative solutions.

***Corresponding author**: Vimala, K.; Email: vimalamaths78@gmail.com*
Manuscript submitted: 18 Nov 2021, Manuscript revised: 09 Feb 2022, Accepted for publication: 27 March 2022

Introduction

The concept of stochastic games (SG) was introduced by Shapley [23]. It is also considered as a repeated game or multistage game. At each stage, the game is in one of the finite number of states. Suppose there are *n* players involved in a game, each player chooses an action from a finite set of possible actions. The player selection actions and the state of the game at that time jointly gives a payoff to each player and transition probabilities are determined to reach the next state. The theory of stochastic games has been developed in many different directions, but one of the method is seeing it as an extension of MDP problem with single agent(player) to multi–agent game. Cooperative game solution is another complicated topic where multiple agents are divided into groups having same mind set in decision making. Only a limited work has been done on the interplay between stochastic game and cooperative game theory (Elan Kohlberg [12,13], Abraham Neyman [19]). This present work is leading towards a new direction along an untraveled path.

The building blocks for this work are taken from Nash [18], Harsanyi [10]. andThe original idea proposed by Nash [19] takes account of both the competitive and cooperative aspect of stochastic games. Nash proved the existenceand uniqueness of such solution. "Bargaining with variable threats " is the route taken to get a solution for cooperative games. The main idea of getting a cooperative game solution startswith an initial competitive stage in which each players declares a threat strategy to be used if negotiation break down. The deployment of these strategies lead to a "disagreement point". Consequently, cooperative strategy has been evolved to achieve *Pareto optimal* outcome and share the gains over the disagreement point. Principle of fairness is implemented to share the payoff.

Extension of the Nash solution to n-player strategic games is our objective. It needs several ideas and planning. Some ideas independently discussed in the works of Harsanyi[10], Shapley [24] Aumann [2],and Neyman [18]. However, there is no comprehensive treatment available for the cooperative game situation. In this paper, we designed a simple solution strategy, and tried to prove the possible generalization of the *Nash solution*. We refer it to as new co- cooperative solution of a strategic game. We do it in two different steps.

TU and NTU Games

TU Games As a first step, consider the alterative way of two players game under the assumption of Transferable Utility (TU), the outcome of "bargaining with variable threats can be described very simply due to Shapley [23] and Kalai [11]. Kalai used it to define co-co-value of Shapley value.Let *s* denote the maximal sum of the player's playoff in any entry of the payoff matrix, and let *d* denote difference in payoff in the two person zero- sum game.

Now the Nash solution splits the total amount s in such a way that the difference in payoff is d. Hence

$$\text{Payoff} \begin{cases} \rightarrow \dfrac{s}{2} + \dfrac{d}{2} \\[2ex] \rightarrow \dfrac{s}{2} - \dfrac{d}{2} \end{cases}$$

This two player game can be extended and generalized to n - player TU games as follows:

Let s denote the total sum of payoffs of any entry of payoff Matrix A=(a_{ij}).

Let d_S denote the minimax value of the zero sum game between the subsets of players S and the complement N - S where $N = \{1,2,3, \dots\}$,and $S \subseteq N$. Players in each of these subsets collaborate as a single player and the opponent player is N - S. The cooperative solution splits total payoff ***s*** into parts which corresponds to the relative strength of subsets S. Specifically the cooperative solution is the Shapley value of the coalition game value v., where $v(N) = s$ for $S \subset N, \quad v(S) - v(N - S) = d_S$.

Thus the solution of game results in a single vector of payoffs, it follows a n-person TU strategic game that has a cooperative solution. Specifically, a three player game has the payoff splitting as follows:

$$\text{Pay off} \begin{cases} \rightarrow \text{Player 1}: \dfrac{1}{3}s + \dfrac{1}{3}\left(d_1 + \dfrac{1}{2}[d_2 + d_3]\right) \\[2ex] \rightarrow \text{Player 2}: \dfrac{1}{3}s + \dfrac{1}{3}\left(d_2 + \dfrac{1}{2}[d_3 + d_1]\right) \\[2ex] \rightarrow \text{Player 3}: \dfrac{1}{3}s + \dfrac{1}{3}\left(d_3 + \dfrac{1}{2}[d_1 + d_2]\right) \end{cases}$$

Non-Transferable Utility Games

In this section we carried a more general case of non –transferable utility (NTU) games. Since the criterion for NTU game is not clear, it became a challenging one. Still it is possible to generalize the Nash solution to NTU games. The lambda transfer strategy proposed by Shapley [23] is a best method by imagining the transferability, after multiplying by some scaling factors: $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \geq 0$. The corresponing TU game is smoother and a cooperative solution can be obtained in that actual transferable utility. These solutions can be defined as a cooperative solution of the NTU game.

In fact, the 2 person NTU games, with the idea of "bargaining with variable threats" coincide with the lambda-transfer method. Thus Nash's original proof establishes existence and uniqueness of the cooperative solutions in any two-person strategic game. Next we prove that a cooperative solution exists for a n-person strategic game. As a major contribution, we extend these results to stochastic games. The main results are definitions and existence theorem for the cooperative game value for TU/NTU stochastic games (both discounted and undiscounted case).

The plan of proof in SG (strategic game) be canfield as follows

co-operative value	existence	uniqueness
TU Games	√	√
NTU Games	√	____

The reminder of the paper is presented as follows: In section 2, we describe the TU (Transferable Utility) and NTU (Non- Transferable Utility) games and the techniques to make th NTU into TU by ascaling vector. In section 3, a plain review of the coalitional game with notations and *Shapleyvalue* of such game is defined. Section 4, depicts the well-defined cooperative solutions concept for TU and NTU games, and the existence and uniqueness theorems. In section 5, we give a brief review the notion of minimax value in undiscounted situation of stochastic game and the definition of MDP with single agent decision making problem is reviewed. Section 6 gives a full account of definition for cooperative solutions for TU and NTU. In section 7, we present some numerical examples and asymptotic propertiesof this solution concept.

Shapley Values and Banzhaf index of Coalition Games
Shapley values in coalition games

In this section, we consider the game model with coalition concept with TU (transferable utility). A coalition game with TU is a pair (N, v), where $N = \{1,2,3, \ldots, n\}$, , v the value of game defined as a map $v : 2^N \rightarrow \mathbb{R}$ s
uch that
, $(\Phi) = 0$.

A coalition is defined by a non-empty set $S \subseteq N$ and $v(S)$ be interpreted as the total utility the players in S can achieve on their own, under the assumption thatutility is transferable among players.Shapley [18] proposed a priori value assessment of what the play of thegame is worth to each player. This kind of value assessment is done through amapping $\varphi: \mathbb{R}^{2^N} \rightarrow \mathbb{R}$, that assigns to each coalitional game ν, a vector of individual utilities φ_{ν} such that the unique value Shapley map is given by

$$\nu_i = \frac{1}{n!} \sum_R [(P_i^R \cup \{i\}) - (P_i^R)] \quad (1)$$

where the summation is taken over all possible permutations of members ofset N in n! ways and P^R denote the subset consist of all j in N that precede I in the ordering R He also attached four properties satisfied by this mapping.
For coalitional games (N, ν), (N, ω)

1) Efficiency: $\sum_{i\in N} \phi_i v = (N)$ (2)
2) Linearity: $(\alpha u + \beta\omega) = \alpha\phi_v + \beta\phi_\omega \quad \forall \alpha.\ \beta\ \epsilon\mathbb{R}$ (3)
3) min $(v(S\cup\{i\}) - v(S)) \le \phi_i v \le \max_{S\subset N, i\notin S}, (v(S \cup \{i\}) - v(S))$ (4)
And
4) $\phi_i v = \sum_{S\subset N} \frac{(s-1)!(n-s)!}{n!} ((S) - (N - S))$, where $s = |S|$ and (5)

$$\phi v_i = \frac{1}{n} \sum_{k=1}^{n} \frac{(k-1)!\,(n-k)!}{(n-1)!} \sum_{S\subset N, i\in S, S \ni i} d_S,$$

$$= \frac{1}{n} \sum_{k=1}^{n} d_{i,k} \quad (6)$$

where $d_S = v(S) - v(N\setminus S)$ and $d_{i,k}$ is the average of the d_S aver all k- player solutions that include i.

The property (4) is well known, this can be proved logically with the following notations. $S \ni$, means that in an ordering of elements in S, i is the last element. The marginal contribution of player i *is* the s is given by $(S) - (S \setminus \{i\})$) In the revenue during the marginal contribution of player i is given by

(N \ (SU{i}))– (N \ S) ,

since, the set of reverse- ordering is same as the set of ordering. We can replace each summand in the right hand side of (1), by the average marginal contribution in the ordering and its reverse,

Here

$$\phi_i v = \frac{1}{n} \sum_{S \ni i} \sum_{R: P_i^R \to S-i} \frac{1}{2} \{((v(S) - v(N \setminus S)) + v(N - S \cup \{i\}) - v(S \setminus \{i\}))\}$$

As the set S ranges over the subset of N, that include , i, $N - \{S \cup \{i\}\}$ also need same property.

We have

$$\phi_i v = \frac{1}{n!} \sum_{S \to i} (s-1)!\,(n-s)!\,(v(S) - v(N \setminus S))$$

Where $|S| = s$ *and* $|N| = n$. Thus the result (4) is proved.

Remark: By defining $d_S = ((S) - (N\setminus S)$, we can choose the value function v, such that

$(\phi) = 0, (N) = d_N$

and for

$\phi \neq S \subset N v(S) = \frac{1}{2} d_S$ *and*

$$v(N \setminus S) = -\frac{1}{2} d_S. \qquad (7)$$

Applying general formula (1) we can obtain Shapley value.

Banzhaf index in coalition games

Like Shapley values an alternative method can be used in cooperative solution of cooperative sense. Banzhaf index for each player i can be used to get optimal solution. In this model we proposed this new solution method. The Banzhaf index $B_i(\Gamma) = \frac{1}{2^{n-1}} \sum_{C \subseteq N} (v(C \cup i) - v(C))$. Here the number of

subsets of $N = \{1,2,3, \ldots, n\}$ not containing the player i is 2^{n-1} ' The marginal contribution of player i in this game is given by $v(C \cup i) - v(C)$ $for\ C \subseteq N$.

Cooperative solutions of TU strategic games

A finite strategic game is a triple $G = (N, A, g)$,where, N= {1, 2, ….,n} is afinite set of players, A, the finite set of players pure strategies and g=$(g^i : i \epsilon N)$, where $g^i = A^N \rightarrow \mathbb{R}$, is player i^s playoff function, Without loss of generality, we may assume that the set of pure strategies (finite) isthe same for all players.

Notation: Linear extension can be denoted by the map g, with g^i: $\Delta(A^N) \rightarrow \mathbb{R}^N$;

where $\Delta(K)$ denotes the probability distribution on an arbitrary set K, also we denote
$A^i = A$, $A^s = \prod_{i \in S} A^i$and $X^S = \Delta(A^s)$ (correlated strategies of the players in S)

In this kind of games, first we assume players can make extra payments, which means the utility transformation can be done from one player to another.
Now the maximum sum of payoff to all is given by,

$(N) = max \sum_{igN} g(x)$________________________________(8)

Next we have to answer the problem of allocation/ distribution of total payoff v (N) over players. First we try to assess the strength of any coalition S. Consider a two person zero-sum between set S and its complements. Since each of these coalition acts as single player, we have

$$(S) - (N \setminus S) = \max_{x \in X} \min_{y \in X^{N-S}} \left(\sum_{i \in s} g^i(x, y) - \sum_{i \notin s} g^i(x, y)\right) \qquad (9)$$

We can now apply the Shapley value to solve the game.

Definition: 4.1

The co-operation solution of the TU game G is nothing but the Shapley value v, where v is a coalitional game satisfying the relation (8).

Theorem 4.2

Every TU strategic game has a unique co-operation solution.
Applying relation (4), we can prove the following proposition.

Proposition 4.3

Let G be a transferable utility (TU) game which is strategic one. It's co-operative solutions $\psi \epsilon\ \mathbb{R}^N$, may be found as follows.

$$= \frac{1}{} \sum^{n} \quad d$$

where d
denotes the average of the

$$i \quad 2 \quad k=1$$

i,k

i,k

$$d_S = \max\ \min\ (\sum (x, y) - \sum g^i(x, y))$$

$$x \in xX^S \quad y \in X$$

$i \in S$

over all k- player coalitions S that includes i.

The NTU strategic Game

$i \notin S$

Von-Neumann Morgenstern utilities of a general strategic game gives payoff. Two different games where the payoff of each player differ a positive factor are the same. Thus we cannot assume transferable utility (TU). So NTU games are natural and in such games addition of payoffs is meaningless. Our focus of attention is no longer in the maximum sum of payoff buton Pareto optimization of the feasible set.

Let $F = \{(x) \colon x \ \epsilon\ x^N\} = conv\ \{(a) \colon x \epsilon A^N\}$

which consist of all the payoff vectors and can be attained, when players correlatetheir strategies.

To get NTU game solutions, we deploy the lambda-transfer method proposed by Shapley[17] with scaling factors $\lambda = (\lambda_1, \lambda_2, \dots \lambda_n) \geq 0$ *and* $\lambda \neq 0$.

In analogy with TU games, solutions procedure, we can take the total payoff available to all players is

$$(N) = \max \lambda_i \; g^i(x)$$
$$x g X^N$$

The relative strength of a coalition s is captured by the relation.
(10)
$(S) - v_\lambda(N - S) = \max$

Min

$\sum_{iss} g^i(x, y) - \sum_{i \notin s} g^i(x, y))$ (11)

$x g X^S \; y s X^{N-S}$

(Here $S = N$ gives relations (9),and $N - N = \phi$)

There are two difficulties arises when we define G, by taking Shapley value ϕV_λ.

- How the scaling factors λ is chosen.
- How the ruled out assumption NTU can be used to attain ϕv_λ (TU case).

The λ – transfer method is not done by a single λ value, one has to accept all values λ for which Shapley value can be implemented without actual TU. Thus we require that (v_λ) be a λ- rescaling of an allocation in the feasible set such that $\phi v_\lambda = \lambda * \psi$, when $\psi \in$,where the * operations is defined as follows:

$f * g = (f_i g_i) \; i \epsilon N \forall f, g \epsilon \mathbb{R}$, simple dot (.) multiplication of vectors/innerproduct.

Linearity of Shapley value yields a functional relation. For every λ scalingfactors, and $\alpha > 0$ if $(v_\lambda) = \lambda * \psi$, then $(u_{\alpha\lambda}) = \alpha\lambda * \psi$

Hence we can normalize the vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ so that it lies in thesimplex $\Delta = \{\lambda \in \mathbb{R}^n, \lambda \geq 0, \sum_{i \in \mathbb{N}} \lambda_i = 1\}$

We now summarize the results as follows:

Definition 4.5

Let v_λ be the coalitional game satisfying

$$(S) - v_\lambda(N - S) = \max$$

$$\min_{x\in X^S}\max_{y\in X^{N-S}}\Big(\sum_{i\in S}\lambda^i g^i(x,y)-\sum_{i\notin S}\lambda^i g^i(x,y)\Big)$$

and $\psi \in F = conv\,\{g(a): a \in A^N\}$ is a cooperative solution of the strategic game $\mathcal{G}$ if $\exists\lambda \in \Delta$ such that $\phi(v_\lambda) = \lambda * \psi$.

Theorem 4.6

Every finite strategic game has a co-operative solution. (Proof is obvious as a special case of Neyman [15]

Remark

- The implementations of a solution without actual side payments-arise from the data of the game- related to market clearing prices arise from the data of an exchange economy.
- The λ- transfer method has been applied in other context also.
- The theoretical aspects of lambda-transfer method have been discussed in Aumann [1] and [2] and Roth [20].

Stochastic Game and Markov Decision Processes

Game theory is a study of mathematical modeling of strategic behavior ofdecision making (players), in specific situations when one player decision may affect the other player in terms of payoff. In stochastic games, play proceeds in consecutive stages. At each stages game is in one of the finite number of states. Suppose there are n players in a SG, each of the players chooses an action inone of the finite set of possible actions. The player's action and the state jointly determined the payoff to each player and the transition probabilities to the succeeding states. Any stochastic game can be represented by a 5-tuple as described in the following definition:

Definition 5.1

- A finite stochastic game- form is a 5-tuple $\Gamma = (N, E, A, q,\ )$, $where\ N = \{1, 2, 3 \ldots n\}$ is a finite set of players.
- E is a finite set of states.
- A is the finite set of player's stage actions.
- q=(q^i), $i\epsilon N$, where q^i: $E \times A^N \to \mathbb{R}$ is the stage payoff to players i.and
- p: $E \times A^N \to \Delta(E)$ are the transition probabilities from present state tofuture stages.

We use new notation E and p different from the notations of strategic games. The different meanings they will take from the context. Further we assume that

- The set of actions A is same for all players.
- The set of actions is independent of the state of the system, when action is

taken that is if $A^i(z)$ denotes the player i^s set of actions in state z, then $A^i(z)$= A for all i and z.

Stochastic games can be defined by specifying the player's strategies and their payoffs. For the infinite game player's behavior (strategies) are given by

$r^i_t: (E \times A^N)^{-1} \rightarrow \Delta(A)$ and $r^i = (r^i_t)^\infty_{t=1}$, $r = (r^i)i\epsilon N$

The strategies r together with the initial state z determine a probability distribution P^z over the different players of infinite game.

This P^z_c denotes the probability distribution over the stream of payoffs. The expected payoff with respect to probability distribution is given by

$$E^z_c[Q] = \sum_{t=1}^{\infty} q^i P^z_c$$

where Q is random variable denoting the payoff.

Even though there are many possible valuation of the stream of payoffs, by fixing the numbers of stages k, we denote:

(i)

$$\beta^i_k(r)[z] = E^z_c \sum_{t=1}^{k} \frac{1}{k} q^i(z_t, a_t), \quad a_t \epsilon A, z_t \epsilon E$$

(ii)

$$\beta^i_k(r) = (\beta^i_k(r)[z]), z\epsilon E$$

(iii)

$$\beta_k(r) == (\beta^k_i(r)), \quad i \epsilon N.$$

Markov Decision Processes (MDP)

The MDP is the special case of Markov game or stochastic game. A single person stochastic game is known as a Markov Decision Process (MDP).One shot game played by a single person has a pure optimal strategy.
Previous section Theorem 7.3 has the -following

Corollary 5.2.1

In an r- discounted MDP there exists an optimal strategy that is stationary and pure. Above corollary may be considerably improved by Blackwell(1962) toprove a new theorem.

Theorem 5.2.2

For every MDP there exists a uniformly optimal pure stationary strategy (pure stationary strategy) , such that

(i) r^* is optimal in the r-discounted MDP for all $r < r_0$ for same $r_0 > 0$.
(ii) for every $\epsilon > 0$, $\exists\, k_g > 0$ such that r^* is the $\in$- optimal in the k-stage gamefor all $k > k_\in$ and
(iii) (iii)

$\bar{q} = \frac{1}{k}\sum_{t=1}^{k}$

$(_t^k, z_t^k)$, converges to probability distribution

P_C^* $a.e$ and $E_C^*[\lim \bar{q}_k \geq E_C^*[\lim \bar{q}_k$ for all r^*.

Remarks:

$k\to\infty$
$k\to\infty$

(a) $\lim_{k\to\infty} \bar{q}_k$ exists and equals P_C^* a.e.

(b) Stationary strategy induces fixed transition probabilities on the statesresulting in a Markov Chain.
(c) As r^* is $\in$- optimal in Γ_k, it follows that

$\lim v_k = \lim E_C(\bar{q}_n = E_C^*[\lim \bar{q}_k$ exits. Thus converges to the same limit.

$k \to \infty$
$k \to \infty$
$k \to \infty$

(d) Theorem statement (b) is equivalent to r^* is ε-optimal in the k- stage game for all but finitely many values of k.

Two zero- sum game-discounted

For a two person-zero sum game with un-disconnected cast (payoff) factor isnot a defined value and optimal strategies. As a first step we try to find analogy with theorem for MDPs. Define the pair of strategies to, for players 1 and 2 respectively, to beoptimal if there exist $k_0 > 0$ and $k_0 > 0$ such that for all r, 6.

(i) $(r_0, \sigma) \geq \beta_r(r, \sigma_0)\ \forall\ 0 < r < r_o$
(ii) $(r_0, \sigma) \geq \beta_k(r, \sigma_0)\ \forall\ k > k_o$
(iii) $E_{c_0,6}$ [lim inf $\overline{q}_n \geq E_{c_0,6}$ [lim sup $\overline{q}_n$

$k \to \infty$

$k \to \infty$

(Here (ii) holds good in MDPs within ε)Finally, we have
Let σ, denotes stratergies of player 1 and r, r_g denote strategies of player 2,then we define stochastic games.

Definition 5.3.1

Let $v\ \epsilon \mathbb{R}^z$ is the (minimax) value of a two=persons zero-sumstochastic game if $\forall \in > 0$. There exists $r_\in, \sigma_\in, > 0$ and $k_\in > 0\ s.t\ \forall r, \sigma$

(i) $\in + (r_\in, \sigma) \geq v \geq \beta_r(r, \sigma_\in) - \in$ for $0 < r < r'$
(ii) $\in +(r_\in, \sigma) \geq v \geq \beta_k(r, \sigma_\in) - \in\ for\ k > k_\in^g$
(iii) $\in + E_{c_0,6}$ [lim inf $\overline{q}_k \geq v \geq E_{c,\sigma_0}$ [lim sup$\overline{]}q$- $\in$.

$k \to \infty$

$k \to \infty$

Remarks

(a) Here condition (i) is redundant because it is a consequence of condition (ii).
(b) The real value u (value of game) is the uniform and limiting average value of atwo- person zero stochastic game if $\forall \in > 0$, *there exists a* $r_\in, \sigma_\in$ and $k_\in > 0$ s.t $\forall\ r, \sigma$,

(ii) respectively
(iii) holds

(c) If the uniform value of the limiting average value exists then it's unique.
(d) If the minimax value v exists, then

$$\lim_{r\to 0} v_r = \lim_{k\to 0} v_k = v.$$

and every MDP has a value v also we prove that condition (ii) implies thefollowing conditions.

(e) for every $\in > 0$ there exist some $r_\in$, $\sigma_\in$ and $w_t > 0$ such that for every r, σ anda decreasing sequence of non- negative sequence $(w_t)_{t=1}^{\infty}$ that converges to 1.

(f)

If $w_1 < w_g$, then $\in +(r_\in, \sigma) \geq v \geq \beta_W(r, \sigma_\in) - \epsilon$, where

$$\beta_w(r)(z) = E^E_c [\sum_{t=1}^{\infty} w_t\ q(z_t, a_t)]$$

This result can be obtained from the following lemma.

Lemma 5.3.2

Any non-increasing sequence of non-negative numbers (w_t) that converges to 1 is an average of sequences of the form $((k))_{t=1}^{\infty}$ where $e(k)_{t=1}^{\infty} = \frac{1}{k}$ for $t \leq k$ and $(k)_t = 0\ for\ t > k.$

Proof:

We have $(w_t) = \sum_{t=1}^{\infty} (k)$ where $\alpha_k = k(w_k - w_{k+1})$

Since (w_k) is a decreasing sequence $(w_k - w_{k+1}) > 0$ and hence $(w_k - w_{k+1}) \geq 0$.

That is $\alpha_k \geq 0$.

So, $\sum_{k=1}^{\infty} \alpha_k = \sum_{k=1}^{\infty} (w_k - w_{k+1}) \geq 0$

$= 1.\ w_1 - w_2 + 2w_2 - 2w_3 + 3w_3 - 3w_4 +$

$= \sum_{k=1}^{\infty} w_k$

Hence the theorem ■. Now we consider the general theorem:

Theorem 5.3.3

Every finite, two-person zero-sum stochastic game has a minimax value. First step towards the proof of the above theorem was taken by Blackwell and Ferguson [6]. They proved the big match for any $t > 0$ there exist non-Markov strategies that satisfy (ii) within an ∈. ■

Corollary 5.3.4

Let v_r ($or\ v_k$) denotes the minimax value of the r- disconnected game(respectively the k- stage game). Then v= Val(Γ) iff $v = \lim_{r\to 0} v_r$ ($respectively\ v = \lim_{k\to\infty} v_k$).

Corollary 5.3.5

Consider the stochastic game Γ = (N, E, A, q, p)d Γ′ = (N, E, A, q′, p) then

$$||\mathrm{Val}(\Gamma) - \mathrm{Val}(\Gamma')||_\infty = \max_{z\in E} |Val(\Gamma)(z) - \mathrm{Val}(\Gamma)(z)| \le ||q - q'||_\infty$$

where $||g||_\infty = \max_{(z,a)\in E\times A} |q(z, a)|$.

Proof

We know that the stages payoffs in the games Γ and Γ′ differs by at most

$$||q - q'||_\infty \quad | \le \mathrm{Val}(\Gamma) - \mathrm{Val}(\Gamma')| \le ||q - q'||_\infty$$

Therefore $|\overline{\Gamma}_k \overline{q}\,\overline{q}|_\infty \le ||q - q'||_\infty$

Thus the optimal strategy is guarantee $v_k - ||q - q'||_\infty$ in Γ_k and vice versa.This implies $\|v_k - v_k'\| \le ||q - q'||_\infty$

Taking $k \to \infty$, in the previous corollary, we get

$$\|V(\Gamma) - \mathrm{Val}(\Gamma')\|_\infty \le ||q - q'||_\infty$$

Corollary 5.3.6

Every Markov Decision Process (MDP) has a uniform value u in a stochasticgame Γ Proof of this corollary also follows from Blackwell's theorem. ■.

Discounted stochastic game

Applying the discount rate r, where $0 < r < 1$, we denote

(i) $(r)(z) = E^z \sum^\infty$ $r(1-r)^{t-1} g^i(z \;, \; a \;)$
r c
$t=1$
t t

(ii) $(r) = (\beta^i(r)[z])$
r r zgE

(iii) $(r) == (\beta^k(r))$
r
igE

Here Γ_r is a family of games, $\Gamma_r{}^z$ is the game, parameterized by the initial state. We denote by u, and u_k the minimax value of Γ_r *respectively*.

The r- discounted game

In this section we consider a two persons zero-sum stochastic games N = $\{1, 2\}$ and $q^2 = -q^1$

In simplified notation, we denote $r^1 = r$, $r^2 = \sigma$ and $(r, \sigma) = \beta_r{}^1(r^1, r^2\;)$ and similarly $\beta_r(r, \sigma) = \beta_r{}^1(r^1, r^2\;)$

Definition: 5.4.2

Let Γ_r be a r- discounted game. If $\exists\, r_0, \sigma_0$ such that $\forall r, \sigma$.

$$\beta_r(r, \;\sigma) \geq v \geq \beta_r(r, \sigma_0 \qquad\qquad) \quad (respectively\ \beta_k(r, \sigma) \geq v \geq \beta_k(r, \sigma_0\;))$$

where u[z] is the minimax value of the game with initial state z, and $Val(G)$ denotes the minimax value of a two-persons zero-sum strategic game G.

Theorem 5. 4.3 (Shapley 1953)

Let Γ_r be a two persons zero sum r-discounted stochastic games, then

(i) $\Gamma_r$has the minimax value and stationary optimal strategies.
(ii) If $[z]nd\ y_r[z]$ are optimal strategies for player 1 and 2 respectively in the(one shot) game $G_r[z, \;v]$, then the stationary stratergies $r_1 = x_r$, $\sigma_t = y_r\ \forall t$ are optimal strategies in Γ_r.
(iii) $(v[z])_{zgE}$ is the unique solution of the equations

$[z] = Va(G_r[z, v])\forall z\epsilon E$____(12) if it v_r is the minimax value of Γ_r with initial state z.

Here we $v_r$denote the minimax value of Γ_r and v_k. The minimax value of Γ_k a finite game.

Cooperative solution of stochastic game
Introduction

In this section we define cooperative solution of a stochastic process game in analogy with the definition. Notations: Let Γ be a Stochastic game and $S \subset N$ = { 1, 2, …n } ,where n number ofplayers involved in the game.

- $X^S = \Delta(A^S)$ the set of all correlated stage actions of the players in the set S.
- r^s: $(E \times A^N)^{-1} \times E \rightarrow X^s$, a correlated stage strategy of players in S at $t$time t.
- $r^S = (r^s)^{\infty}_{=1}$ a correlated behavior strategy of the players in S.
- $\Sigma^s = \{r^S\}$ set of all correlated behavior strategies of the players in S
- $\sum^{N}_{s.p}$

the finite set of stationary pure strategies of the players in $\sum^N$.

Solutions

The existence and uniqueness theorem of Nash [14] requires only the existence of minimax value in two person strategic games. Following theoremgives the proof in a lucid manner.

Theorem 6.3

Every two- person Stochastic game (discounted) or un-discounted) TU orNTU has a unique cooperative solution.

TU Stochastic Game

By applying the following modification in the stochastic game we get thefollowing result:

Replace $(N) = \max_{x\in X^N} (\sum_{i\in N} q(x))$ by $u(N) = \max_{c\in\sum^N} (\sum_{i\in N} q^i(x))$______________(13)

and replace u(S) – u (N- S) of section 3, by u (S) – u (N-S) = Val (Γ^s)____________(14)

where Γ^s denotes the two person-zero sum stochastic game, played between S and N-S, where the pure stage action is A^S and A^{N-S} respectively.

The stage payoff to S is given by

$$\sum q^i(z, a^s, a^{N-S}) - \sum_{i \in S} q^i(z, a^s, a^{N-S})$$

This proves the following theorem.

Theorem 6.5:

$i \notin S$

Every TU Stochastic game, discounted or un-discounted has some unique cooperative solutions.

NTU discounted stochastic game

A non-transferable utility Stochastic game Γ has a unique solution with feasible set M_r as follows:

$$M_r = \{(r) : r \in \textstyle\sum^N \} \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (15)$$

$$= conv\{\beta_r(r) : r \in \textstyle\sum^n_{s.p} \}.$$

Since every two-persons zero sum r- discounted stochastic game has a unique solution.

Definition 6.6.1

$\psi_r \in F_r$ is an NTU -value of the r- discounted stochastic game Γ_r if $\exists\ \lambda \in \Delta$, such that $(v_{r,}) = \lambda * \psi$, where $v_{r,\lambda}$ is a coalition game satisfying

$$v_{r,}(S) - v_{r,\lambda}(S) = vVl\ (\Gamma_{r,\lambda}{}^S) \forall S \in N \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (16)$$

6.7 NTU un-discounted stochastic game

In this section, we try to find the cooperative solutions of NTU-un- discounted games. We define the feasible set $M_0 \subseteq \mathbb{R}^N$,where

$$M_0 = \{x : \exists\ r \in \sum_{s.p}^{N}$$

$$s.\ t.\ x = \lim_{r \to 0} (r)\}$$

This feasible set contains the co-operative solution of NTU game.

Lemma 6. 7.1

For a NTU game Γ the following feasible set is convex.

$$M_0 = conv\,\{x : \exists\, r \in \sum_{s.p}^{N}$$

$s.\,t.\quad x = \lim_{r\to 0} (r)\}$

$$= \{x : \exists\, r \in \sum_{s.p}^{N}$$

$s.\,t.\quad x = \lim_{r\to 0} (r)\}$

$$= conv\,\{x :\ \exists\, r \in \sum^{N}$$

$s.\,t.\,x = \lim \beta$

$(r)\}$________(17)

Proof:

$s.p$

$r\to 0 \quad r$

First we prove that M_0 is convex polytope. Let x', x'' ∈ M_0. Them r', r'' $\in \Sigma^N$ such that $x' = \lim_{r\to 0} (r')$ and $x' = \lim_{r\to 0} \beta_r(r')$

By Kuhn's theorem $\exists\, r \in \Sigma^N$ that induces the same distribution on the players of the game as the mixed strategy 1

2

r'

$+ \frac{1}{2} r''$

Thusβ

$(\bar{r}) =$

$\frac{1}{2}\ \ 1$

$(r + \ \ r'')$

$\bar{r} \quad r \quad 2 \quad 2$

$= \frac{1}{\ }\ \ ' \ \ 1 \ \ ''$

$\beta_r(r\ ) +\ \ \beta_r(r\ )$
$2\ \ \ 2$

Therefore

$M\ \ \ni \lim \beta$

$(\tilde{r})) = \ \ \frac{1}{}\lim$

$\frac{1}{r} + \lim\beta$ $(\ ')$

(r'')

$0\ \ r\to 0\ \ r$

$2\ r\to 0\ \ r$

$2\ r\to 0\ \ r$

$= 1\ x'2$

$+ \frac{1}{2} x''$

Since M_0 is convex, $M_0 \supseteq conv\{x: \exists\ r \in \sum^N_{s.p}$

$/\ x = \lim_{r\to 0}(r)\}$

Then there is a separating linear functional $y \in \mathbb{R}^N$ such that

$(y, x_0) > (y, x)\ \forall x = \lim_{r\to 0}(r)$ such that $r \in \sum^N_{s.p}$

But this contradicts Blackwell's theorem of with respect to the MDP having stage payoff $\langle(y, q)\rangle$. Similarly, we may show that the second set of limits is also a convex polytope, spanned by the limiting expected payoffs of the pure stationary strategies. Finally we note that if r is a stationary strategy then $\lim_{r\to 0}(r) = \lim_{k\to\infty} \beta_k(r)$

Thus the first and third sets in (17) are equal and there sets are identified as M_0. Next we get the following lemma.

Lemma 6.7.2

Let $M_0(\lambda) = \{\lambda * x: x \in M_0\}$ then

(i) If $y \in M_0(\lambda)\}$ then $y_i \le \lambda_i \|q^i\|\ \forall i \in \mathbb{N}$

(ii) The mapping $\lambda \to M_0(\lambda)$ is continuous. The co-operative solutions of the un-discounted stochastic game is analogous to definition 12.1 for stochastic games.

Definition 6.8

The point $\psi \in M_0$ is a co-operative solution of the stochastic game Γ if $\exists \lambda \in \Delta$ such that $(v_\lambda) = \lambda * \psi$ where v_λ is a coalitional game with $v_\lambda(S) -$

$(N - S) = Val(\Gamma_\lambda{}^S)\ \forall S \subseteq N.$

Remark

When S= N, $(N) = Val(\Gamma_\lambda{}^S)$ is the maximal expected payoff in the MDP with the single player N pure stage actions are A^N and the stage payoff is

$\sum_{i \in \mathbb{N}} \lambda_i q^i(z, a).$

Theorem 6.9

Every NTU un-disconnected stochastic game has a co-operative solution.
Proof: It is obvious as theorem 2 in section 3,

Theorem 6.10

Every finite stochastic game discounted or un-discounted TU or NTU has co-operative solutions.

Asymptotic Behavior

In this section first we take into account some general logical statements and general proof techniques. First we define the concept of atomic formula. Atomic formula is an expression of the form $p > 0$ *or* $p = 0$, where p is a polynomial with integer coefficients in one or more variables. An *elementary formula (sentence)* is an expression constructed in a finite number of steps from atomic formula by mean of conjunctions $(\wedge)$ disjunctions $(\vee)$ negations $(\sim)$ and quantifiers of the form "there exists" $\exists$ or "for all" $\forall$. The variables is free in a formula if somewhere in the formula it is not modified by a quantifier like $\exists$ *or* $\forall$.

Lemma 7.1

Prove that for fixed (N, E, A), the statement of *Theorem 6* is an *elementary sentence.*

Definition 7.2

An ordered field is said to be real–closed if no proper algebraic extension is ordered.

Tarski's Principle:An *elementary sentence* that is true over one real-closed field is true over every real-closed field. Example given in [3]. Consider a field of power series in a fractional power of r that converges for $r > 0$ sufficiently small, ordered ascending to the assumptions that r is infinitesimal. That is $r < a$ for any real number > 0, is real-closed ([3]).
This field is of the form,

$$\bar{F} = \left\{ \sum_{k=0}^{\infty} \alpha_k r^{k/M} : M \in Z^+, \text{ and } \alpha_k \in \mathbb{R} \text{ series converges for } r > 0. \right\} \qquad (18)$$

The generic element of this field is of the form, $\sum_{k=K}^{\infty} \alpha_k r^{k/}$, $k \in \mathbb{Z}$, set of integers.

If the series remains bounded as $r \to 0$ then K = 0.
Theorem 6.10, Lemma 7.1 and Tarshi principle 7.3 yield the followingtheorem.

Theorem 7.4:

Fix the triplet (N, E, A) for all $0 < r < 1$ there exist a vector $\psi_r \in \mathbb{R}^N$, $\lambda_r \in \mathbb{R}^N$ and $u_{r,\lambda} \in \mathbb{R}^{2^N}$ satisfying the cooperative solution, satisfying

(i) $\psi_r \in \mathbb{R}$ (ii) $\lambda_r \in \Delta$

(iii) $u_{r,\lambda_r}(S) - u_{r,\lambda_r}(N \setminus S) = val\,(\Gamma^s_{r,\lambda_r})$, $\forall\, S \subseteq N$ and

(iv) $(u_r, \lambda_r) = \lambda_r * \psi_r$ and each co-ordinate of these variables has an asymptotic expression of the form $\sum_{k=0}^{\infty} \alpha_k r^{k/M}$.

More precisely there exist real Pursuer series $\bar{\psi}, \bar{\lambda}$ and $\bar{u}(S)$ such that for any $r > 0$ (sufficiently small) and $\forall i \in N$ and $S \subset N$, these series converges to ψ_r^i, λ_r^i and $u_{r,\lambda_r}(S)$ respectively.

We have $\psi_0 \in M_0$ and corollary 5.12 yields that v_{λ_0} is the uniform minimax value of (Γ^s), $\forall S \subseteq N$. Thus ψ_0, λ_0 *and* v_λ satisfy the requirement of *definition*

r, 0

4.10 ,hence ψ_0 is an NTU value of stochastic game Γ. At the outset, we get the following result:

Theorem 7.5

Every finite stochastic game Γ has a co-operative solution is the limiting case (as $r \rightarrow 0$) of co-operative solutions of r- discounted games. Furthermore these cooperative solutions as well as their scaling factors and the associated minimax value and optional strategies in the zero-sum scaled games are real Puisuex-series converging to their counter-parts of the game Γ.

Conclusion

This paper has initiated a deep study on extension principles of Harsaryi- Shapley cooperative solutions. It is based on the central idea of extension of Nash equilibrium for one-stage strategic games to finite stochastic game. We proved many results for:

- finitely many players, states and actions
- complete information bounded games
- perfect monitoring and current state and past actions of player's oriented games.

The *r*- discounted game has a cooperative solutions and it's proofs arediscussed. The existence of cooperative solutions in the un-discounted case depends on the uniform value for two person zero-sum models. We also tried a new solution concept based on Banzhaf index for the i^{th} player for cooperative games. This is an innovative idea for getting optimal solution for cooperative stochastic games.

References

1. Aumann, R. J. (1985), On the Non-transferable Utility Value: A Comment on the Roth-Shafer Examples, Econometrica, 53, 667-677.
2. Aumann, R. J. (1986), On the Non-transferable Utility Value: Rejoinder, Econometrica, 54, 985-989.
3. Aumann, R.J. (1985a) The axiomatization of a NTU value, Econometrica, 83, 599-612.
4. Benoit, J.P., and V. Krishna (1987) , Nash Equilibrium of finitely repeated Games, International Journal of Game Theory, 16, 197-204.
5. Bewley, T. and E. Kohlberg (1976), The Asymptotic Theory of Stochastic Games, Mathematics of Operations Research, 1, 197-208.

6. Bewley, T. and E. Kohlberg (1978), On Stochastic Games with Stationary Optical Strategies, Mathematics of Operations Research, 3, 104-125.
7. Blackwell, D. (1962), Discrete Dynamic programming, The Annals of Mathematical Statistics, 2, 724-738.
8. Blackwell, D. and T.S. Ferguson (1968), The Big Match, Annals of Mathematical Statistics, 39, 159-168.
9. Cherian, Jacob; Jacob, Jolly; Qureshi, Rubina; Gaikar, Vilas. 2020. 'Relationship between Entry Grades and Attrition Trends in the Context of Higher Education: Implication for Open Innovation of Education Policy' MDPI, Switzerland, Journal of Open Innovation Technology, Market and Complexity, Vol- 6, Issue- 4: 199
10. Gilette, D. (1957), Stochastic Games with Zero Stop Probabilities, Contributions to the Theory of Game 3, 179-187.
11. Harsanyi, J. (1963), A simplified Bargaining model for the n-Person Co-operative Games, International Economic Review 4, 194-220.
12. Kalai, A. and Kalai E. (2013), Cooperation in Strategic Games Revisited, Quarterly Journal of Economics, 128, 917-966.
13. Kohlberg, E. (1974), Repeated Games with Absorbing States, Annals of Statistics, 4, 194-220.
14. Kohlberg, E., and Abraham Neyman (2015) The Cooperative Solution of Stochastic Games, Working Paper, Harvard Business School, USA, pp- 1-39.
15. Martens, J.F. and Neyman, A. (1981), Stochastic Games, International Journal of Game Theory, 10, 53-66.
16. Martin, J.O., and Ariel , R. (1998) A course in Game Theory , The MIT Press , Cambridge, Massachusetts , USA
17. Milnor, J. (1952), Reasonable Outcomes for n-Persons Games, The Rand Corporation, 18, 196.
18. Nash, J. (1950), The Bargaining Problem, Econometrica, 18, 155-162.
19. Nash, J. (1953), Two-Person Cooperatives Games, Econometrica, 21, 128-140.
20. Neyman, A. (2003), From Markov Chains to Stochastic Games, Neyman A. and Sorin, S.(eds.), NATO ASI series 2003, Kluwer Academic Publishes, pp. 58-75.
21. Roth, A. (1985), On the Non-transferable Utility Value: A Reply to Aumann, Econometrica, 54, 981-984.
22. Sameer A., Gaikar V. (2019), 'Study of Demographic Variables on Financial Goal of Urban Individuals' in International Journal of Research, Vol. 9 (1), July-December 2019, Pp. 24 – 34.
23. Satheeskumar, R., and Elango, C. (2010) Semi-Markov Decision Processes in Service Facilities Holding Perishable Inventory, International Journal of Computer Applications, V-9 348- 365.
24. Satheeskumar, R., and Elango,C. (2012) Markov Decision Processes in Service Facilities holding perishable inventory, OPSEARCH, (4), 348- 365.
25. Shapley, L. (1953), Stochastic Games, Proceedings of the National Academy of Sciences, U.S.A., 39, 1095-1100.
26. Shapley, L. (1984), Mathematics 147 Game Theory, UCLA Department of Mathematics, 1984, 1988 and 1990.
27. Vilas Gaikar, Sameer Aziz Lakhani. (2020). 'Demographic Variables Influencing Financial Investment of Urban Individuals: A Case Study of

Selected Districts Of Maharashtra State', International Journal of Advanced Science and Technology, 29(05), Pp.962 – 974.