Stochastic Game (Shapley, 1953) GT

$\langle Q, N, A, P, R \rangle$

Q = a <u>finite</u> set of strategic games

N = a finite set of Players

$A = A_1 \times A_2 \times \dots \times A_n$, A_i is a <u>finite</u> set of actions for agent i

$P = Q \times A \times Q \longmapsto [0,1]$ Transition Probability functions

$P(q, a, \hat{q})$ = Probability of moving from q to $\hat{q}$ on a

$R = r_1, \dots, r_n$ Payoff functions, r_i is Payoff function for Player i

Strategy Space for players is the same.

$h_t = q^0, a^0, q^1, a^1, \dots, a^{t-1}, q^t$) = t stages of game, a history of length t.

H_t = set of all possible histories of length t.

$\prod_t H_t A_i$ = set of deterministic strategies

A <u>behavioral strategy</u> $S_i(h_t, a_{ij})$ returns the Probability of playing action a_{ij} for history h_t.

A <u>markov Strategy</u> S_i is a Markovian Strategy in which $S_i(h_t, a_{ij}) = S_i(h'_t, a_{ij})$ if $q_t = q'_t$ where these are final states of h_t & h'_t.

A <u>Stationary Strategy</u> S_i is a Markov Strategy in which $S_i(h_{t_1}, a_{ij}) = S_i(h'_{t_2}, a_{ij})$ if $q_{t_1} = q_{t_2}$ with these as final states of h_{t_1} and h'_{t_2}.

<u>Theorem</u>: Every n-player, general-sum, discounted-reward Stochastic game has a Markov Perfect Equilibrium (MPE).

An MPE Consists of only Markov strategies and is a Nash Equil. regardless of the starting state.

Hexmoor
2010

e.g. Jean-Francois Mertens, 2002 Handbook of GT

Consider a 2-Player Zerosum game with 2 States ω_1 & ω_2

Payoff matrix for w_1 (Immediate Payoffs) $\begin{pmatrix} 10 & \overset{\dagger}{-1} \\ -1 & 10 \end{pmatrix}$ for w_2 $\begin{pmatrix} 0 & 6 \\ 3 & 0 \end{pmatrix}$

Transition Probabilities for w_1 $\begin{pmatrix} (1,0) & \overset{\dagger}{(\frac{1}{2},\frac{1}{2})} \\ (\frac{1}{2},\frac{1}{2}) & (1,0) \end{pmatrix}$ † prob of going to w_1, w_2 after Column gets 1 from row player

for w_2 $\begin{pmatrix} (0,1) & (0,1) \\ (0,1) & (0,1) \end{pmatrix}$

Let u = Value of starting at state $w=1$

v = " " " " " $w=2$ N.B. w_2 is an absorbing state, i.e., no exit from it.

$$v = \mathrm{Val}\begin{pmatrix} \underbrace{\frac{1}{2}\cdot 10}_{\text{Immediate}} + \underbrace{\frac{1}{2}(1\cdot v + 0\cdot u)}_{\text{future}} & \frac{1}{2}\cdot(-1) + \frac{1}{2}\left(\frac{1}{2}\cdot v + \frac{1}{2}\cdot u\right) \\ \frac{1}{2}\cdot(-1) + \frac{1}{2}\left(\frac{1}{2}\cdot v + \frac{1}{2}\cdot u\right) & \frac{1}{2}\cdot 10 + \frac{1}{2}(1\cdot v + 0\cdot u) \end{pmatrix}$$

$$u = \mathrm{val}\begin{pmatrix} \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot u & \frac{1}{2}\cdot 6 + \frac{1}{2}\cdot u \\ \frac{1}{2}\cdot 3 + \frac{1}{2}\cdot u & \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot u \end{pmatrix}$$

$$u = \frac{1}{2}\cdot u + \frac{1}{2}\cdot \mathrm{val}\begin{pmatrix} 0 & 6 \\ 3 & 0 \end{pmatrix} = \frac{1}{2}\cdot u + \frac{1}{2}\cdot 2 = 2$$

$$v = \mathrm{Val}\begin{pmatrix} 5 + \frac{1}{2}\cdot v & \frac{1}{4}\cdot v \\ \frac{1}{4}\cdot v & 5 + \frac{1}{2}\cdot v \end{pmatrix} = \frac{5}{2} + \frac{3}{8}\cdot v = 4$$

Row should play $(\frac{1}{2}, \frac{1}{2})$ at $w=1$ and $(\frac{1}{3}, \frac{2}{3})$ at $w=2$

Column " " " " " " $(\frac{2}{3}, \frac{1}{3})$ " ".