Introduction

Stag Hunt Problem

(With Two Players)

Stag Hunt Problem

	Stag	Hare
Stag	2,2	0,1
Hare	1,0	1,1

1. If both row-player and column-player hunt stag, since a stag is worth 4 "utils", they each get 2 "utils."

2. If both row-player and column-player hunt hares, since a hare is worth 1 "util", they each get 1 "util."

3. If row-player hunts hare, while column-player hunts stag (and hence fails to hunt any thing), then the row-player gets 1 "util" and the column-player gets 0 "util."

4. The other case is symmetric.

Note that if row-player is risk aversive, he will choose to hunt hare and thus guarantee that he gets 1 "util" independent of the choice column-player makes. Thus he will maximize the minimum utility under the two possible pure strategies ("hunt stag" with a minimum utility of 0 if the opponent hunts hare vs. "hunt hare" with a minimum utility of 1 regardless of what the opponent chooses to play) and choose to hunt hare. By symmetry, it is seen that in fact both players will choose to hunt hares.

Is this the truly optimal strategy?

Quoting Rousseau (*Discourse on the origin and Basis of Equality among Men*):

> "If a group of hunters set out to take a stag, they are fully aware that they would all have to remain faithfully at their posts in order to succeed; but if a hare happens to pass near one of them, there can be no doubt that if he pursued it without qualm, and that once he had caught his prey, he cared very little whether or not he had made his companions miss theirs."

Changing the discussion slightly, suppose that column-player will play a mixed strategy by playing "hunt stag" with some probability (say, y) and by playing the other strategy ("hunt hare") with probability $(1 - y)$. His best choice of these probabilities must be such that row-player is now "indifferent" to the choice of his own strategies. Thus, we have

$$2y + 0(1 - y) = 1y + 1(1 - y)$$

and $y = 1/2$. Thus one expects both row-player and column-player to play the strategies "hunt stag" and "hunt hare" with equal probabilities.

Why are these kinds of analysis important to us?

1. Economy

2. Evolutionary Biology
3. Large Scale Distributed Systems
4. Resource Allocation
5. Intelligent Agents

Prisoners' Dilemma

Prisoners' Dilemma

	C	D
C	0,0	-2,1
D	1,-2	-1,-1

There are two prisoners (row-player and column-player) arrested for a particular crime, but the prosecutor does not have enough evidence to convict them both. He relies one one of them testifying against the other in order to get a conviction and punish the second prisoner by sending him to jail. If both of them testify against the other (defections: "D, D") then they both go to jail for 1 year each, thus getting a "util" of -1. If, on the other hand, both maintain silence (cooperations: "C, C") then they go free with "util" of 0 each. If, on the other hand, row-player testifies (D) and column-player maintains silence (C), then row-player is rewarded with 1 util and column-player is punished with -2 util. The other case is symmetric.

The pay-offs can be made all non-negative by adding 2 utils to each and thus getting a pay-off matrix:

Prisoners' Dilemma (Modified Pay-offs)

	C	D
C	2,2	0,3
D	3,0	1,1

1. For column-player the strategy C is dominated by the strategy D independent of how row-player plays the game. Thus column player must defect.

2. Similarly, for row-player the strategy C is dominated by the strategy D independent of how column-player plays the game. Thus row player must defect.

Hence the equilibrium strategy for the players is to defect even when they could have each gotten better pay-offs by cooperating.

Second-Price Auction

1. Seller has one indivisible unit of object for sale.

2. There are I potential buyers (bidders) with valuations

$$0 \leq v_1 \leq v_2 \leq v_I.$$

   (Consider the case when $I = 2$.)

3. The bidders simultaneously submit bids

$$s_i \in [0, \infty].$$

4. The highest bidder wins the object.

5. But he only pays the second bid ($\max_{j \neq i} s_j$).

6. His utility is

$$v_i - \max_{j \neq i} s_j.$$

Consider the special case of just two players

$$v_1, v_2 = \text{valuations} \qquad s_1, s_2 = \text{bids}.$$

Pay-offs

$$\begin{aligned} u_1 &\equiv \text{ if } s_1 > s_2 \text{ then } v_1 - s_2 \text{ else } 0. \\ u_2 &\equiv \text{ if } s_2 > s_1 \text{ then } v_2 - s_1 \text{ else } 0. \end{aligned}$$

Let us look at the player 1's choices.

1. <u>Overbidding</u>
   (a) $s_1 \leq s_2$: The payoff is zero and the strategy is weakly dominated.
   (b) $s_2 \leq v_1$: The payoff is $v_1 - s_2$ and the strategy is weakly dominated with respect to bidding $s_1 = v_1$.
   (c) $v_1 < s_2 < s_1$: The payoff is $v_1 - s_2 < 0$ negative and the strategy is strongly dominated.
2. <u>Underbidding</u>
   (a) $s_2 \geq v_1$: The payoff is zero and the strategy is weakly dominated.
   (b) $s_1 \geq s_2$: The payoff is is $v_1 - s_2$ and the strategy is weakly dominated with respect to bidding $s_1 = v_1$.
   (c) $s_1 < s_2 < v_1$: The payoff is zero and the strategy is weakly dominated.

So the best strategy for player 1 is to bid exactly his own valuation ($s_1 = v_1$). And by a symmetric argument, the best strategy for player 2 is also to bid exactly his own valuation ($s_2 = v_2$).

Two Person Zero-sum Games

We define a *loss matrix* M as follows:

$M(s_i, s_j) = M(i, j) =$ Loss suffered by the row-player for the strategy profile (s_i, s_j).

Rock, Paper & Scissors

	R	P	S
R	1/2	1	0
P	0	1/2	1
S	1	0	1/2

Row-player's goal is to minimize the loss. Assume (without loss of generality) that all the losses are in the range $[0, 1]$.

Row-player's expected loss

$$\begin{aligned} &\sum_{i,j} \sigma_r(s_i)\sigma_c(s_j)M(s_i, s_j) \\ &= \sum_{i,j} \sigma_r(i)M(i,j)\sigma_c(j) \\ &= \sigma_r^{\mathrm{T}} M \sigma_c = M(\sigma_r, \sigma_c). \end{aligned}$$

$$\begin{aligned} \sigma_r(s_i) &= \text{Probability that the row player plays } s_i \\ \sigma_c(s_j) &= \text{Probability that the column player plays } s_j \end{aligned}$$

Similarly,

$$M(\sigma_r, j) = \sum_{i,j} \sigma_r(i)M(i,j) \quad \text{and} \quad M(i, \sigma_c) = \sum_{i,j} \sigma_c(j)M(i,j).$$

Row-player's strategy

$$\min_{\sigma_r} \max_{\sigma_c} M(\sigma_r,\ \sigma_c).$$

A mixed strategy σ_r^* realizing this minimum is called a *minmax* strategy.

Theorem 1.5.1 The MINMAX theorem: von Neumann

$$\min_{\sigma_r} \max_{\sigma_c} M(\sigma_r,\ \sigma_c) = \max_{\sigma_c} \min_{\sigma_r} M(\sigma_r,\ \sigma_c).$$

Obstacles

1. **Imperfect Information**
   M (pay off) may be unknown.
2. **Computational complexity**
   M is so large that computing a minmax strategy using a linear program is infeasible.
3. **Irrationality**
   Opponent (column-player) may not be truly adversarial.

Repeated Play (with learning)

M unknown

1. The game is played repeatedly in a sequence of rounds.
2. On round $t = 1, \ldots, T$:
   (a) The learner (row-player) chooses mixed strategy $\sigma_{r,t}$.
   (b) The opponent (column-player) chooses mixed strategy $\sigma_{c,t}$.
   (c) Row-player observes all possible losses
   $$M(i, \sigma_{c,t}) = \sum_{i,j} \sigma_{c,t}(j) M(i, j),$$
   for each row i.
   (d) Row-player suffers loss $M(\sigma_{r,t}\ \sigma_{c,t})$.

Row-player's cumulative expected loss:

$$\sum_{t=1}^{T} M(\sigma_{r,t},\ \sigma_{c,t}).$$

The expected cumulative loss of the best strategy

$$\sum_{t=1}^{T} M(\sigma_r^*,\ \sigma_{c,t}) = \min_{\sigma_r} \sum_{t=1}^{T} M(\sigma_r,\ \sigma_{c,t}).$$

Learning Algorithm

Parameter β to be chosen. Initially,

$$\begin{aligned} W_1(i) &= 1, \qquad \forall i \\ W_{t+1}(i) &= W_t(i)\beta^{M(i,\sigma_{c,t})} \\ \sigma_{r,t}(i) &= \frac{W_t(i)}{\sum_i W_t(i)}. \end{aligned}$$

Analysis of Learning Algorithm

Inequality 1

$$\begin{aligned} \sum_i W_{t+1}(i) &= \sum_i W_t(i)\beta^{M(i,\ \sigma_{c,t})} \\ &= \left(\sum_i W_t(i)\right) \cdot \sum_i \sigma_{r,t}\beta^{M(i,\ \sigma_{c,t})} \\ \Rightarrow \quad \frac{\sum_i W_{t+1}(i)}{\sum_i W_t(i)} &= \sum_i \sigma_{r,t}\beta^{M(i,\ \sigma_{c,t})} \\ &\leq \sum_i \sigma_{r,t}(1 - (1 - \beta)M(i,\ \sigma_{c,t})) \\ &= 1 - (1 - \beta)M(\sigma_{r,t},\ \sigma_{c,t})). \end{aligned}$$

After telescoping, we get

$$\frac{\sum_i W_{T+1}(i)}{\sum_i W_1(i)} \leq \prod_t (1 - (1 - \beta)M(\sigma_{r,t},\ \sigma_{c,t}))$$

Hence,

$$\begin{aligned}\ln\left(\frac{\sum_i W_{T+1}(i)}{n}\right) &\leq \sum_t \ln(1-(1-\beta)M(\sigma_{r,t},\ \sigma_{c,t})) \\ &\leq -(1-\beta)\sum_t M(\sigma_{r,t},\ \sigma_{c,t}).\end{aligned}$$

Inequality 2

$$\begin{aligned}\sum_i W_{T+1}(i) \geq W_{T+1}(j) &= \beta^{\sum_t M(j,\ \sigma_{c,t})} \\ &\geq \beta^{\sum_t M(\sigma_r^*,\ \sigma_{c,t})}.\end{aligned}$$

Hence

$$\ln\left(\frac{\sum_i W_{T+1}(i)}{n}\right) \geq (\ln\beta)\sum_t M(\sigma_r^*,\ \sigma_{c,t})) - \ln n.$$

Final Result

Combining the two inequalities:

$$(1-\beta)\sum_t M(\sigma_{r,t},\ \sigma_{c,t}) \leq \ln n + (\ln 1/\beta)\sum_t M(\sigma_r^*,\ \sigma_{c,t})).$$

and,

$$\begin{aligned}\sum_t M(\sigma_r^*,\ \sigma_{c,t}) &\leq \sum_t M(\sigma_{r,t},\ \sigma_{c,t}) \\ &\leq \frac{(\ln 1/\beta)}{1-\beta}\sum_t M(\sigma_r^*,\ \sigma_{c,t})) + \frac{\ln n}{1-\beta}.\end{aligned}$$