I. Game Theory Basics

$N = \{1, \dots, n\}$ ← set of players

S_i ← set of actions for player i

$A = S_1 \times \dots \times S_n$ ← action profile
(set of every possible combinations of actions)

In case of RPS:
$A = \{(R,R), (R,P) \dots (S,P), (S,S)\}$

$\underbrace{u}_{\text{utility function}} : \underbrace{(a_1, \dots, a_n)}_{\in A} \longmapsto \underbrace{(\underbrace{u_1}_{\text{pay-off for player } i \text{ (reward)}}, \dots, u_n)}_{\text{vector of utilities for each player}}$

Pay-off table for RPS:

	R	P	S
R	0,0	-1,1	1,-1
P	1,-1	0,0	-1,1
S	-1,1	1,-1	0,0

zero-sum game - if each vector's sum is 0

cst-sum game - each vector has cst sum ← we can reformulate to 0-sum if we add a dummy player with always negative pay-off

pure strategy - if player chooses a single action with probability 1

mixed strategy - at least 2 actions with positive probability → we denote it like σ

$\sigma_i(s)$ ← probability that player i chooses action s
$(s \in S_i)$

$$\underbrace{u_i(\sigma_i, \sigma_{-i})}_{\text{Expected utility for both pl.}} = \sum_{s \in \underbrace{S_i}_{\text{player 1}}} \sum_{s' \in \underbrace{S_{-i}}_{\text{player 2}}} \sigma_i(s) \cdot \sigma_{-i}(s') \cdot u_i(s, s') \longrightarrow \text{returns a vector}$$

Expected utility for both pl., if player 1 play with strategy σ_i and player 2 plays with strategy σ_{-i} (In the 2-player case). ← The best strategy for every player is to maximize their expected utility

Nash Equilibrium ← Both players play best "response" strategies
No player can improve by changing strategy alone.

		Garry C (cinema)	Garry M (match)
Monica	C (cinema)	(2,1)	(0,0)
	M (match)	(0,0)	(1,2)

$\sigma_{gary}(C) = x$ ← probability that Gary chooses a movie

$\Rightarrow \sigma_{gary}(M) = 1 - x$

$\sigma_{monica}(C) = y$

$\Rightarrow \sigma_{monica}(M) = (1-y)$

$$\overset{\text{expected utility (Belief)}}{u_{monica}(C)} = \underset{\text{probability}}{\sigma_{garry}(C)} \cdot \overset{\text{action utility}}{u_{monica}(C,C)} + \underset{\text{probability}}{\sigma_{garry}(M)} \cdot \overset{\text{action utility}}{u_{monica}(\underset{\text{monica}}{C},\underset{\text{garry}}{M})} = x \cdot 2 + (1-x)\cdot 0 = 2x$$

$$u_{monica}(M) = \sigma_{garry}(C) \cdot u_{monica}(M,C) + \sigma_{garry}(M) \cdot u_{monica}(M,C) = x \cdot 0 + (1-x)\cdot 1 = 1-x$$

For monica to be indifferent: $u_{monica}(C) = u_{monica}(M)$

$$2x = 1-x$$

$$x = \frac{1}{3} \Rightarrow \sigma_{garry}(C) = \frac{1}{3}$$

$$\sigma_{garry}(M) = \frac{2}{3}$$

II. Regret Matching

1st round

pl	opp	utility for pl	regret for pl	strategy
R	P	-1	0	0
P	P	0	+1	1/3
S	P	+1	+2	2/3

Regret for player i of not choosing an action s_2 when choosing s_1: $r_i(s_1, s_2) = u_i(s_2) - u_i(s_1)$

Regret matching - select actions with probability proportional to positive regrets

Our goal is to minimize expected regrets over time

2nd round

pl	opp	utility for pl	regret for pl	previous regrets	cu ive regrets	strategy
S	R	-1	0	2	2	1/3
R	R	0	1	0	1	1/6
P	R	+1	2	1	3	1/2

3rd round

pl	opp	utility for pl	regret for pl	previous regrets	cumulative regrets	strategy
R	S	+1	0	2	2	1/2
P	S	-1	-2	1	0	0
S	S	0	-1	3	2	1/2

can these be negative? → should be negative and positive