نظریه بازیها
Game Theory

ارائه کننده: امیرحسین نیکوفرد
مهندسی برق و کامپیوتر دانشگاه خواجه نصیر

دانشگاه صنعتی خواجه نصیرالدین طوسی

Extensive form

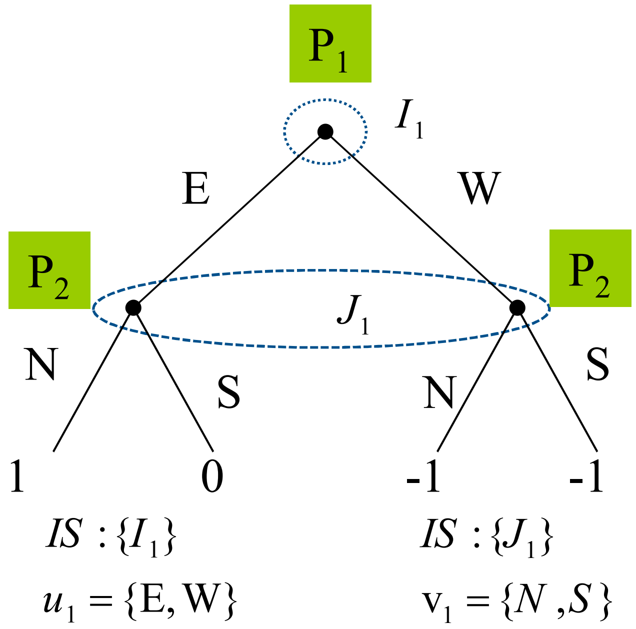
Material

- Dynamic Non-cooperative Game Theory: Second Edition
  - Chapter 2.4
- An Introductory Course in Non-cooperative Game Theory
  - Chapter 7,8

Zero sum games

- Zero sum games
  - Definitions
  - Pure strategies
  - Dominating strategies
  - Mixed strategies
  - Linear programs to find NE
  - Extensive form and information structure
  - Feedback games: pure, mixed, and behavioral strategies

Extensive form :example



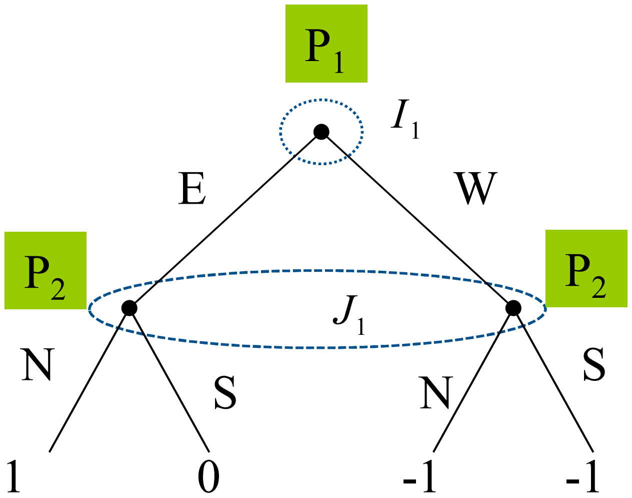
Example of a single-stage game

- Player 1 plays "first"
- Player 2 plays "second"
- Player 2 does not know the action of Player 1
- Pure strategy

  $\gamma(I_1) = E \qquad \sigma(J_1) = S$

Simultaneous play

Extensive form :example



$IS: \{I_1\}$

$u_1 = \{E, W\}$

$IS: \{J_1\}$

$v_1 = \{N, S\}$

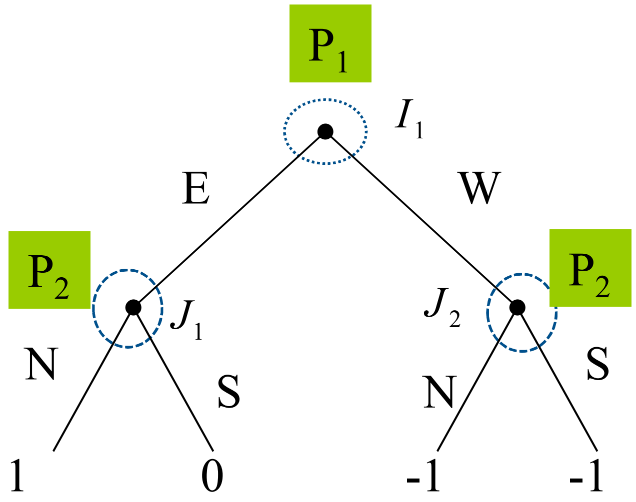
From extensive form to matrix form :

P_2

P_1

	$\sigma(J_1) = N$	$\sigma(J_1) = S$
$\gamma(I_1) = E$	1	0
$\gamma(I_2) = W$	-1	-1

Extensive form :example



$IS : \{I_1\}$ $\quad$ $IS : \{J_1, J_2\}$

$u_1 = \{E, W\}$ $\quad$ $v_1 = v_2 = \{N, S\}$

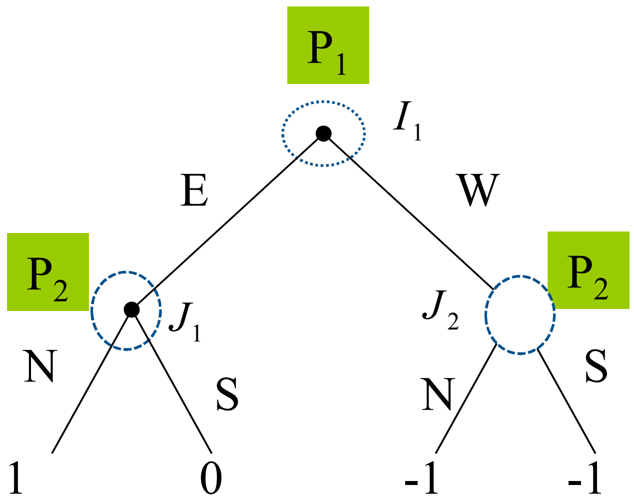
Example of a single-stage game

- Player 1 plays “first”
- Player 2 plays “second”
- Player 2 knows the action of Player 1
- v_1 *and* v_2 could be different
- Pure strategy

$\gamma(I_1) = E \qquad \sigma(J_1) = S \quad \sigma(J_2) = S, N$

Sequential play

Extensive form :example



From extensive form to matrix form :

$IS:\{I_1\}$ $\quad$ $IS:\{J_1, J_2\}$

$u_1 = \{E, W\}$ $\quad$ $v_1 = v_2 = \{N, S\}$

P_2

P_1

	$\sigma(J_1)=N$ $\sigma(J_2)=N$	$\sigma(J_1)=N$ $\sigma(J_2)=S$	$\sigma(J_1)=S$ $\sigma(J_2)=N$	$\sigma(J_1)=S$ $\sigma(J_2)=S$
$\gamma(I_1)=E$	1	1	0	0
$\gamma(I_2)=W$	-1	-1	-1	-1

Actions and strategies

Actions $\neq$ strategies

- Let $I_1, \ldots, I_r$ be the **information sets** of Player 1.
- Let U_i be the set of **actions** available to Player 1 in the information set I_i
- A pure strategy γ for Player 1 is a map that assigns an action to each information set.

$$\gamma : \{I_1, \ldots, I_r\} \rightarrow \bigcup_i u_i$$

$$I_i \mapsto \gamma(I_i) \in u_i$$

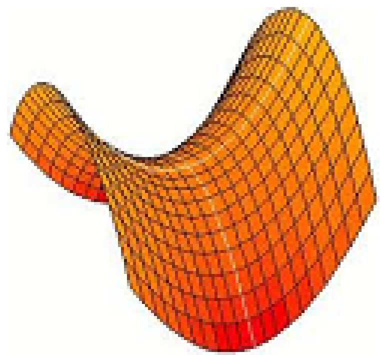
- For **player 2** $\sigma : \{J_1, \ldots, J_r\} \rightarrow \bigcup_i v_i$

Saddle-point

Definition: Saddle point equilibria (pure equilibrium strategy)

A pair of strategies γ^* and σ^* is a saddle point equilibrium if for any other policies γ and σ , the outcome $J(\gamma,\sigma)$ of the game satisfies :

$$J(\gamma,\sigma^*) \leq J(\gamma^*,\sigma^*) \leq J(\gamma^*,\sigma)$$



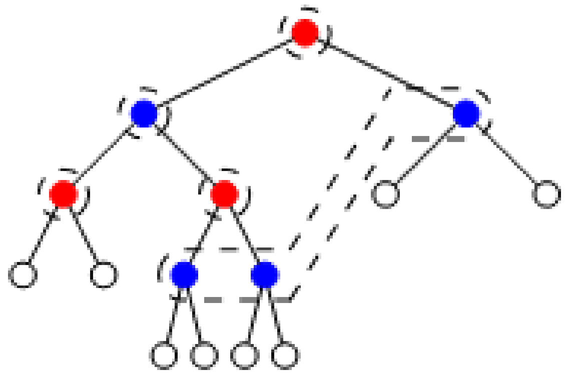
value $J(\gamma^*,\sigma^*)$ *is the saddle-point value;*

- Definitions extend to this larger class of games, as long as we use the new definition of strategy.

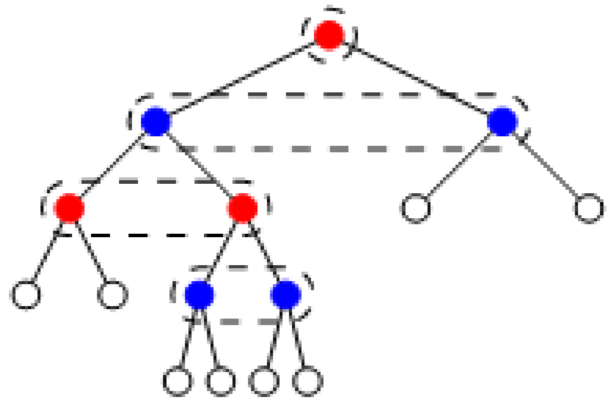
Feedback games

- A multi-stage game in extensive form is a feedback game if

1. no information set spans over multiple stages
2. each "Player 1" node is the root of a separated sub-game.



not a feedback game because the dash- dot information set spans over multiple stages, violets conditions C_1

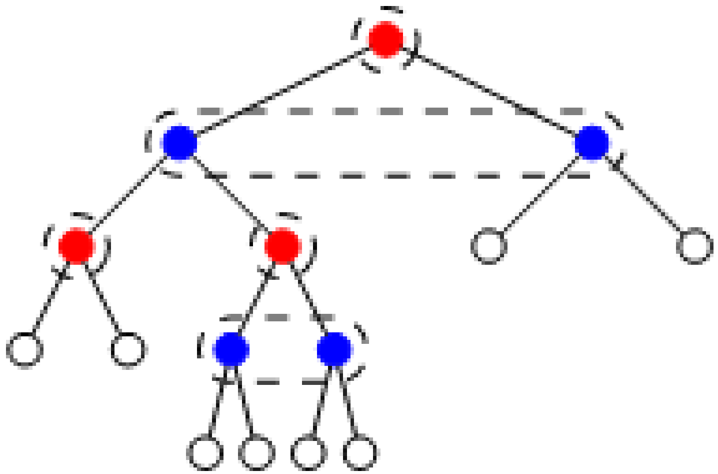


not a feedback game because at the start of the second stage P_1 does not necessarily know the action selected by P_2 in the first stage. the dash-dot information set violets conditions C_2

Feedback games

- A multi-stage game in extensive form is a feedback game if

1. no information set spans over multiple stages
   - Each player knows the current stage of the game
2. each "Player 1" node is the root of a separated sub-game.
   - Both player know what happened in the previous stages



Feedback games

- A multi-stage game in extensive form is a feedback game if

1. no information set spans over multiple stages
2. each "Player 1" node is the root of a separated sub-game.

Finding NE in feedback games

Backward induction can be used to find NE in feedback games.

Note: we already used backward induction for games with perfect information, which is a special case of feedback games.

- A very reasonable class for control applications.
- Think of stages as time steps in a discrete time system

Mixed strategies

- Take a game in extensive form for which

$$\Gamma_1 = \{\gamma_1, \gamma_2, \ldots, \gamma_m\} \qquad \Gamma_2 = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$$

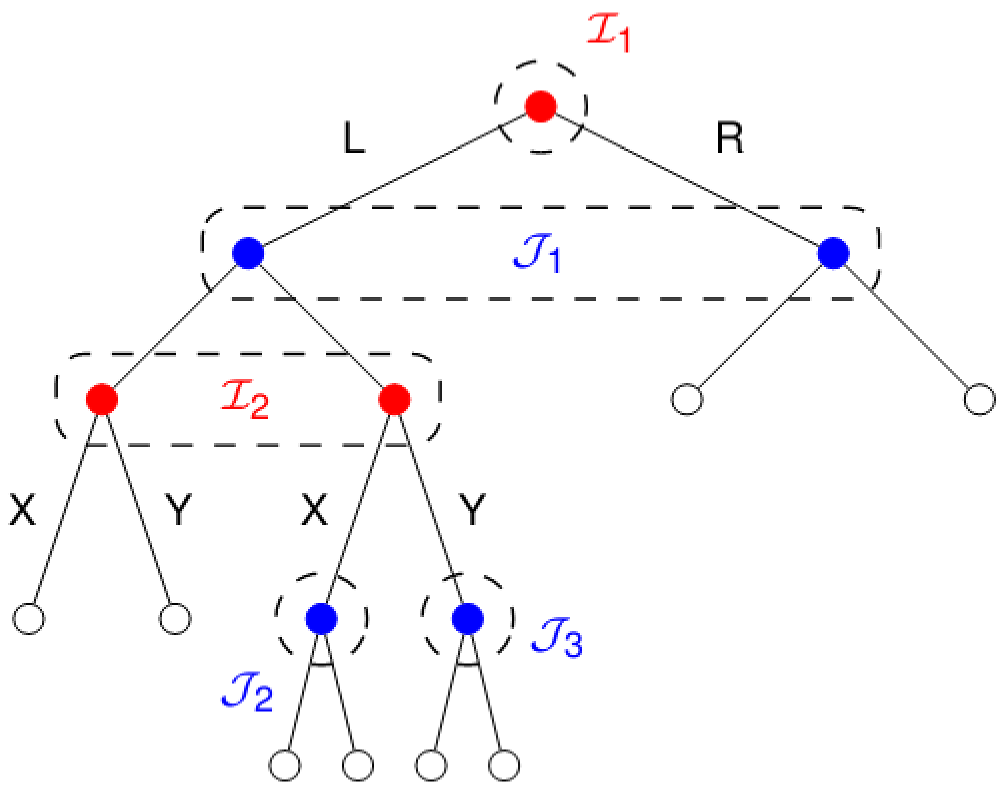
- are the finite sets of all pure strategies for players P_1 and P_2, respectively. As we saw in earlier, this game can be represented in matrix form by an $m \times n$ matrix A_{ext}
- For games in extensive form, a ***mixed strategy*** corresponds to selecting a ***pure strategy*** randomly according to a previously selected probability distribution before the game starts, and then playing that policy throughout the game:

Example

$$\Gamma_1 = \{\gamma_1, \gamma_2, \gamma_3, \gamma_4\}$$
$$= \{\mathrm{LX}, \mathrm{LY}, \mathrm{RX}, \mathrm{RY}\}$$

- Where LX means

$\gamma_1(I_1) = L \quad \gamma_1(I_2) = X$



Mixed strategies

- Consider the set of pure strategies for a player i:

$$\Gamma_i = \{\gamma_1, \gamma_2, \dots, \gamma_m\}$$

- A mixed strategy $y \in R^m$ for Player 1 (and equivalently $z \in R^n$ for Player 2) corresponds to randomly selecting a pure strategy from the set of pure strategies Γ, according to the probabilities matrix $\mathrm{y}_1, \mathrm{y}_2, \dots, \mathrm{y}_m$

where $$\sum_{i=1}^{m} y_i = 1,\ y_i \geq 0,\ i = 1, \dots, m$$

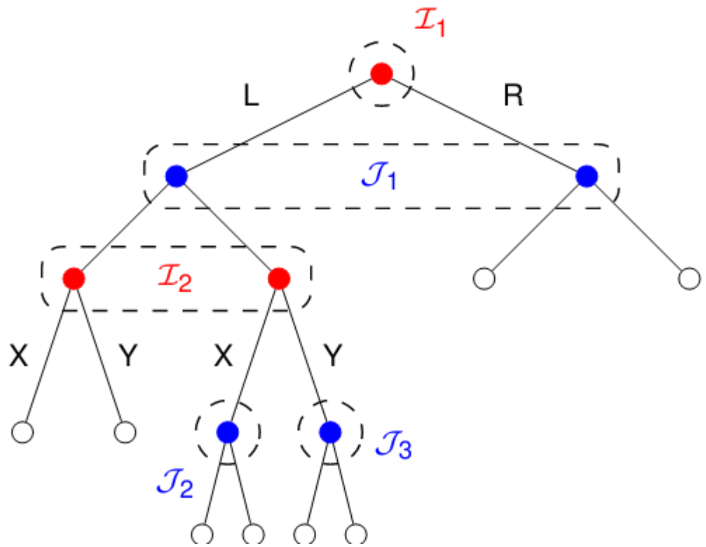
Exactly the same definition as in games in matrix form

Example

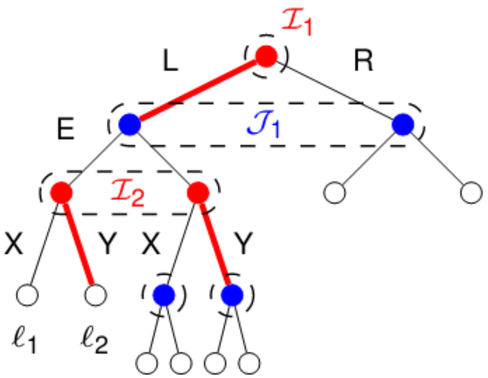
$\Gamma_1 = \{LX, LY, RX, RY\}$

- For example

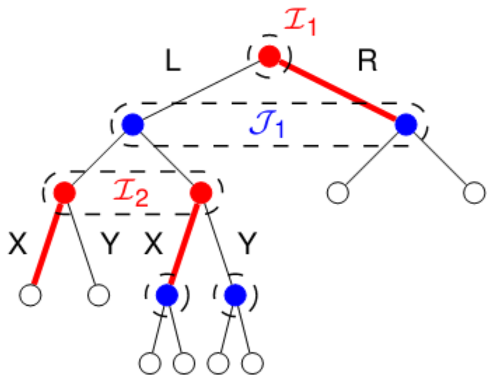
  $Y = [0, 0.5, 0.5, 0]^T$



50%



50%



Mixed strategies

- Given this definition, a number of results follow naturally.
- Expected outcome of a game

$$J(\mathrm{y},\mathrm{z}) = \sum_{\gamma\in\Gamma_1}\sum_{\sigma\in\Gamma_2} J(\gamma_i,\sigma_j)\,\mathrm{Pr}ob(P_1\ selects\ i\ ,P_2\ selects\ \ j\,)$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n} J(\gamma_i,\sigma_j)y_i z_j$$

If A_{ext} is the equivalent matrix form, then

$$J(\mathrm{y},\mathrm{z}) = y^T A_{ext} z$$

Mixed strategies

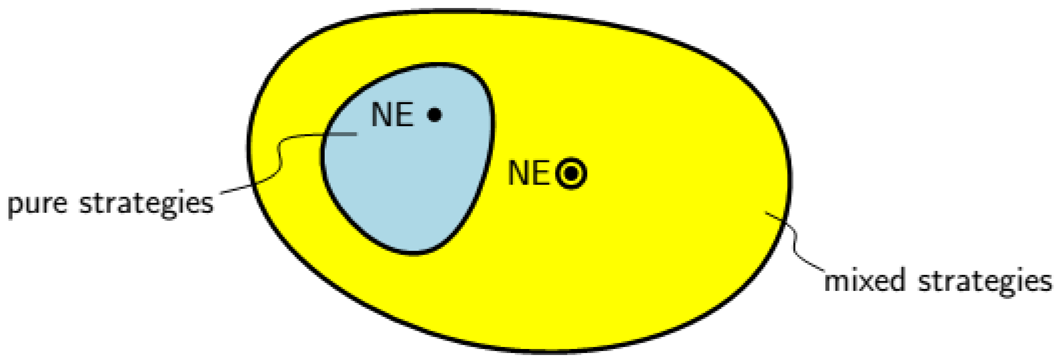
Definition: (Mixed Nash equilibrium):

- $(y^*, z^*) \in Y \times Z$ is a **mixed saddle point (Nash equilibrium)** if

$$J(y, z^*) \leq J(y^*, z^*) \leq J(y^*, z)$$

for any $y \in Y$, and any $z \in Z$

$J(y^*, z^*)$ is called the saddle point value.



Mixed Strategies

(Mixed saddle-point vs. security levels)

For every zero-sum game in extensive form with (finite) matrix representation A_{ext}

- *A mixed saddle-point equilibrium always exist and*

$$\underline{V}_m(A_{ext}) = \max_{y \in Y} \min_{z \in Z} y^T A_{ext}\, z = \min_{z \in Z} \max_{y \in Y} y^T A_{ext}\, z = \overline{V}_m(A_{ext})$$

(y^*, z^*) is a mixed saddle-point equilibrium

- $\underline{V}_m(A) = \overline{V}_m(A) = y^* A_{ext} z^*$ is the saddle point value

This condition holds for all matrices A_{ext}

➡ For any two **player zero-sum game** in extensive form with (finite) matrix there exists a saddle point equilibrium (Nash equilibrium)

Behavioral strategies

- **Behavioral strategies** are randomized strategies in which the randomization is done **over actions** as the game is played and not **over pure policies** before the game starts

- It is assumed that the random selections by **both players** at the **different information sets** are all done statistically **independently** and that the players try to **optimize** the resulting **expected outcome** J of the game

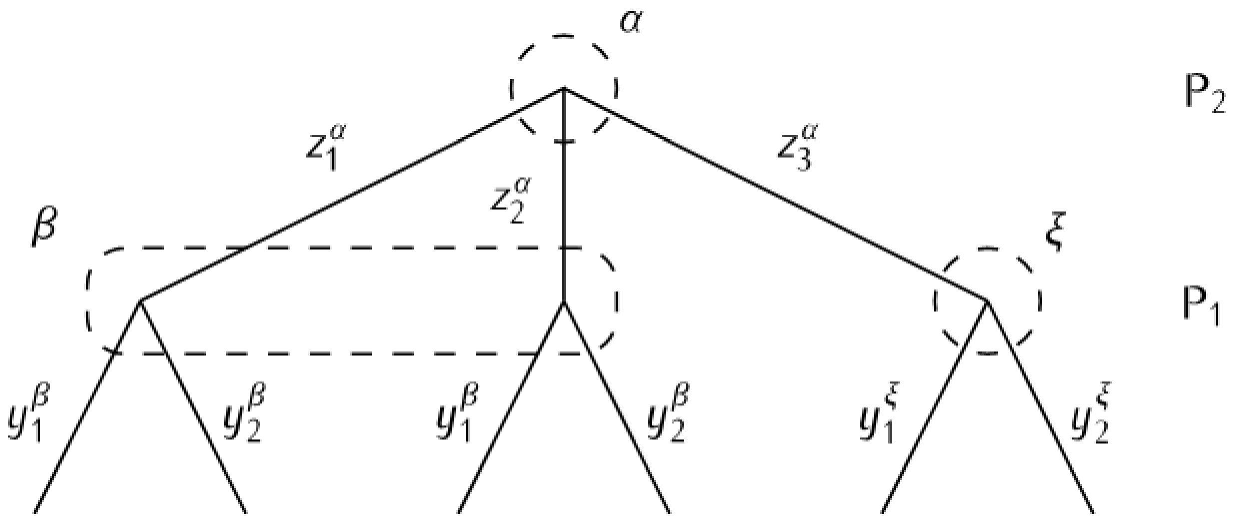
Behavioral strategies

- The following systematic procedure can be used to compute the expected outcome for an arbitrary game in extensive form:

1. Label every link of the tree describing the game in extensive form with the probability with which that action will be chosen by the behavioral policy of the corresponding player.
2. For each leaf multiply all the probabilities of the links that connect the root to that leaf. The resulting number is the probability for that outcome.
3. The expected reward is the sum over all leaves of the product of the outcome corresponding to that leaf with the probability for that outcome computed above.

Behavioral strategies

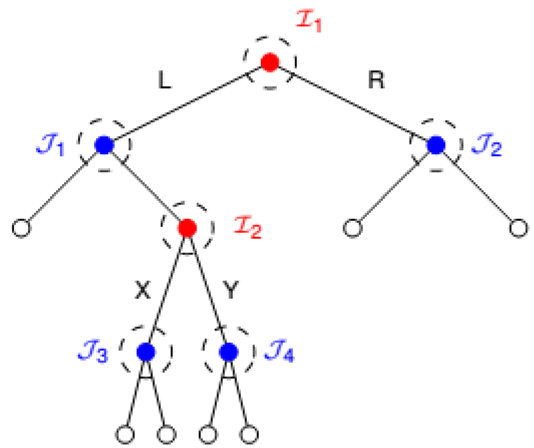
- The following systematic procedure can be used to compute the expected outcome for an arbitrary game in extensive form:



$$1(z_1^\alpha y_1^\beta) \quad +6(z_1^\alpha y_2^\beta) \quad +3(z_2^\alpha y_1^\beta) \quad +2(z_2^\alpha y_2^\beta) \quad +0(z_3^\alpha y_1^\xi) \quad +7(z_3^\alpha y_2^\xi)$$

Behavioral strategy: Example

- **Behavioral strategy**
- in I_1, 20% L, 80% R,
- in I_2, 50% X , 50%Y .



$$\gamma^b(I_1) = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} \qquad \gamma^b(I_2) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

Corresponding mixed strategy?

pure strategies: $\Gamma_1 = \{LX, LY, RX, RY\}$

What is P(LX)?

$$P(\mathrm{LX}) = P(u_1 = L, u_2 = X) = P(u_1 = L)P(u_2 = X) = 0.2.0.5 = 0.1$$

$$y = [\mathrm{P(LX), P(LY), P(RX), P(RY)}]^T = [0.1, 0.1, 0.4, 0.4]^T$$

Behavioral ,mixed, pure strategies

- **Pure strategy**

A pure strategy γ maps each information set into an action

$\gamma_1(I_1) = L \quad \gamma_1(I_2) = X$

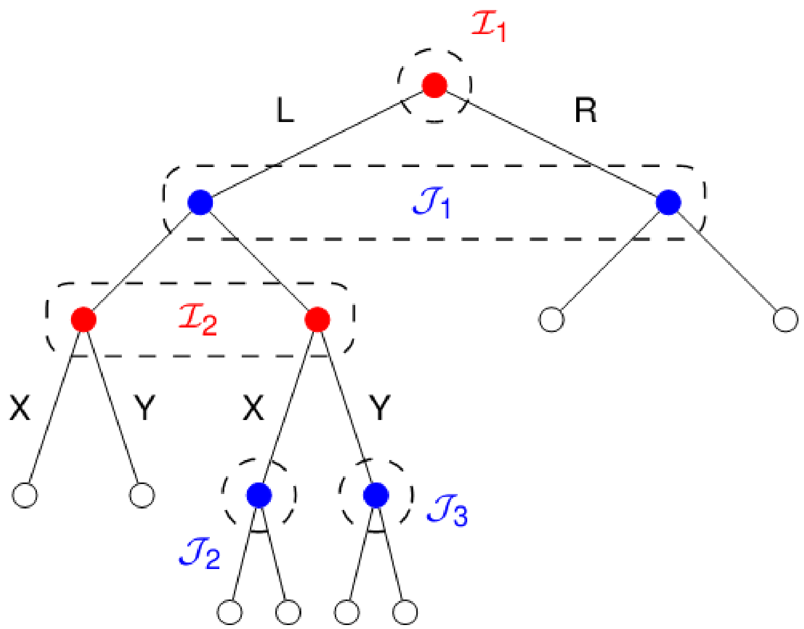
- **Mixed strategy**

A mixture of pure strategies from

$\Gamma_1 = \{LX, LY, RX, RY\}$

For example 50% LX , 50% LY

$Y = [0.5, 0.5, 0, 0]^T$

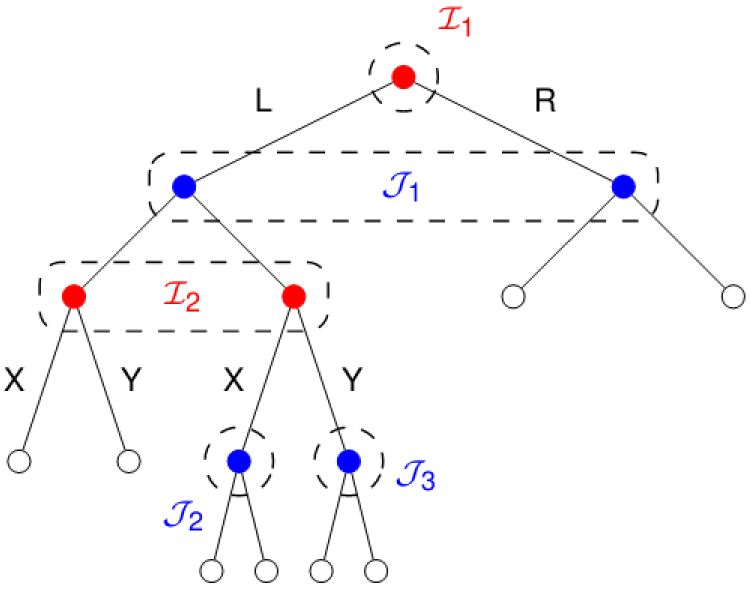


Behavioral ,mixed, pure strategies

- **Behavioral strategy**
- A random action at each IS.
- For example, in I_1, 50% L, 50% R,

while in I_2, 50% X , 50% Y .

$$\gamma^b(I_1) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \quad \gamma^b(I_2) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$



Behavioral strategies

- **Pure strategies**

A map that assigns an action to each information set

$$I_i \mapsto u_i = \gamma(I_i) \in U_i$$

- **Mixed strategies**

A probability distribution over the pure strategies $\gamma_i \in \Gamma$

$$y \in R^m, \qquad m = |\Gamma|, \qquad y_i = P(\gamma_i)$$

- **Behavioral strategies**

A map that assigns a probability distribution over the available actions to each information set

$$I_i \mapsto \gamma^b(I_i) \in y_i \subset U_i$$

Behavioral strategies vs mixed strategies

- The total number of distinct pure policies for a given player P_i is given by

(# actions of 1st IS) x (# actions of 2nd IS) x • • • x (# actions of last IS) .

and therefore mixed policies are probability distributions over these many pure actions, which means that we have

(# actions of 1st IS) x (# actions of 2nd IS) x • • • x (# actions of last IS) **-1**

degrees of freedom in selecting a mixed policy.

- As for behavioral policies, for each information set we need to select a probability distribution over the actions of that information set. Such probability distribution has as many degrees of freedom as the number of actions minus one. Therefore, the total number of degrees of freedom available for the selection of behavioral policies is given by

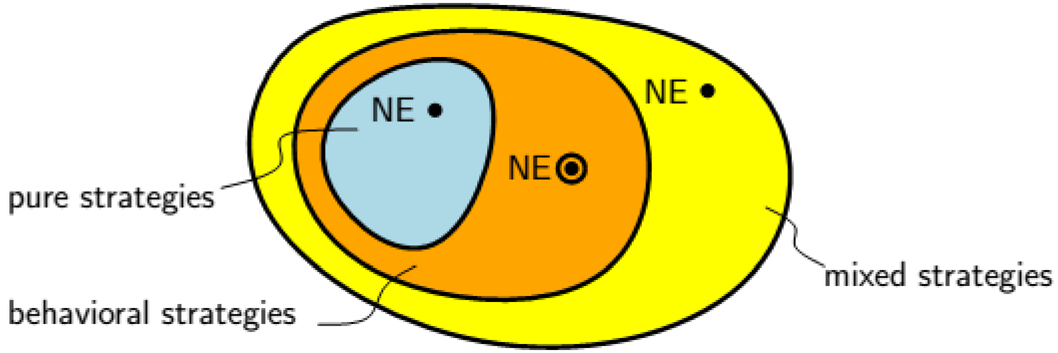
(# actions of 1st IS -1) + (# actions of 2nd IS-1) + • • • + (# actions of last IS-1),

Behavioral strategies vs mixed strategies

- By simply counting degrees of freedom we have seen that the set of mixed strategies is far **richer** than the set of behavioral strategies. In fact, this set is sufficiently rich so that every game in extensive form has saddle-point equilibria in mixed policies. It turns out that for **large classes of games**, the set of behavioral policies is already sufficiently rich so that one can already find saddle-point equilibria in **behavioral policies.**
- Moreover, since the number of degrees of freedom for behavioral policies is much lower, finding such equilibria is computationally much **simpler.**

Behavioral strategies vs mixed strategies

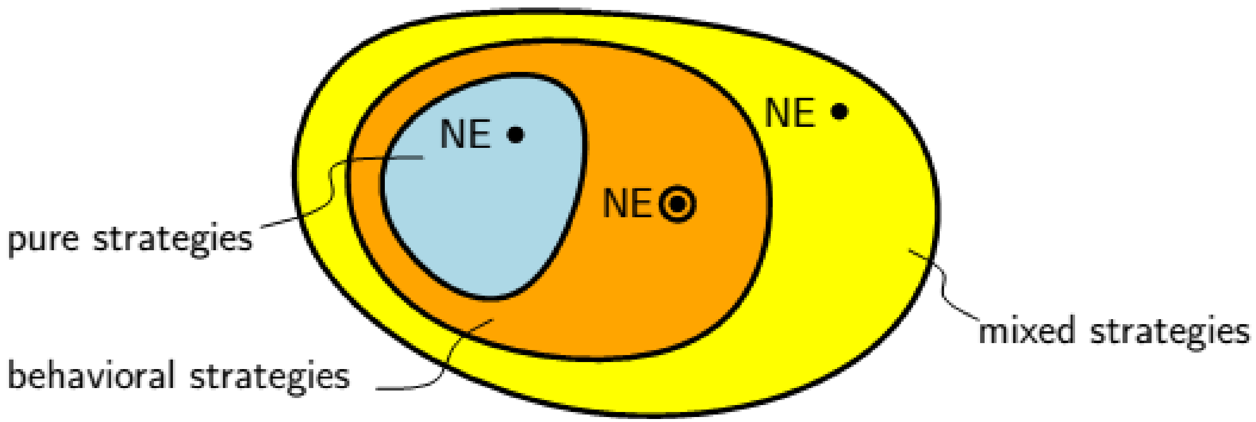
- We have defined the smaller set of behavioral strategies.



- In the following, we will see that, for **feedback games**,
  - It is not restrictive to consider only behavioral strategies.
  - The set of behavioral strategies is **computationally tractable**
    - There is an algorithm to find saddle point behavioral strategies
    - The low search space dimension makes the algorithm efficient

Kuhn's theorem

- Given **a feedback game** in extensive form, for any mixed strategy y for Player 1, there exists a behavioral strategy γ^b for Player 1 such that, for any **mixed strategy z played by Player 2**, y and γ^b give the **same probability distribution over the leaves of the tree**, and therefore the same **outcome**.



Behavioral strategies

Definition: (Behavioral Nash equilibrium):

- $(\gamma^{b*}, \sigma^{b*}) \in \Gamma_1^b \times \Gamma_2^b$ is a **behavioral saddle point (Nash equilibrium)** if

$$J(\gamma^{b}, \sigma^{b*}) \le J(\gamma^{b*}, \sigma^{b*}) \le J(\gamma^{b*}, \sigma^{b})$$

for any $\gamma^b \in \Gamma_1^b$, and any $\sigma^b \in \Gamma_2^b$

$J(\gamma^{b*}, \sigma^{b*})$ is called the saddle point value.