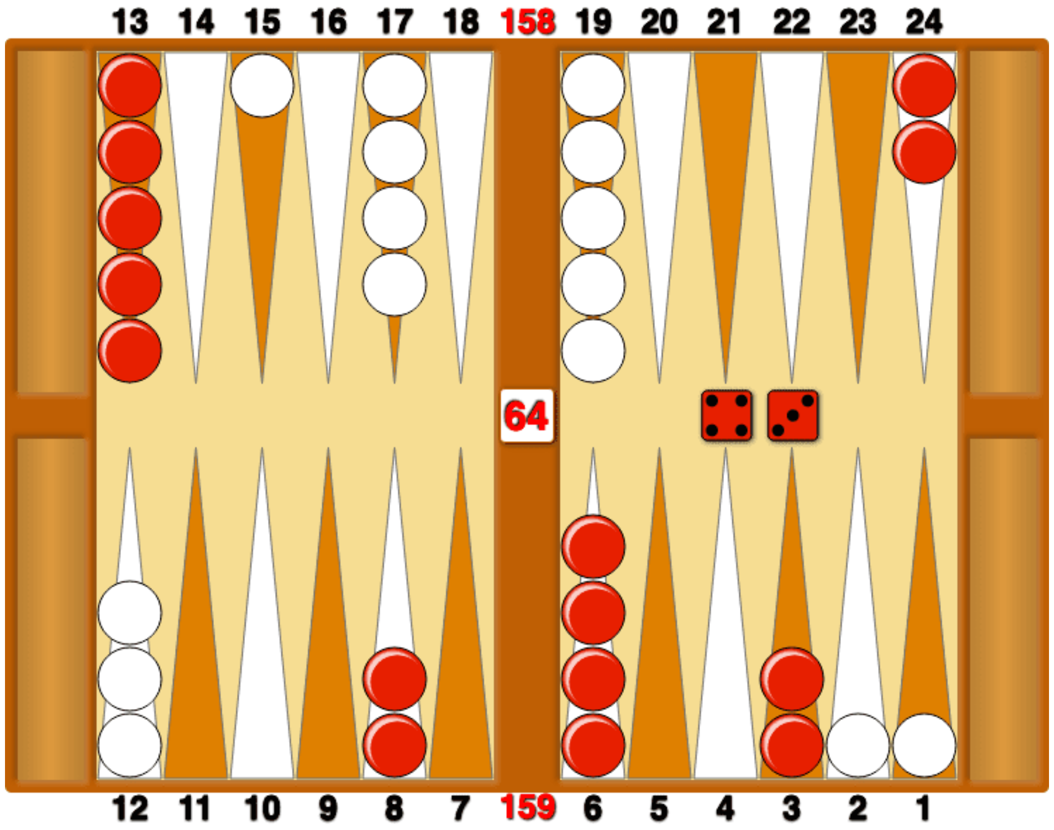
БАГАТОКРОКОВІ АНТАГОНІСТИЧНІ СТОХАСТИЧНІ ІГРИ

ВІТАЛІЙ ЧУМАК

Марківські ігри з неповною інформацією для 2 гравців

- **S** – множина станів
- **A, B** – множини дій
- $\mu_0 \in P(s)$ – початковий розподіл
- $p_{ij}(a, b)$ – ймовірність переходу,
- $r_i : A \times B \to R$ – функція нагороди
- $\beta \in R_+$ – коефіцієнт дисконтування



Пошук оптимальних стратегій

1) Визначити $n = 0,\ \mu = \mu_0$
2) Знайти оптимальне рішення $(z^*,\ x^*)$ для $D_{N-n}(\mu)$
3) Запам'ятати i_n та вибрати дію a_n
4) Запам'ятати b_n
5) $n+=1,\ \mu = \Phi\ (\mu,\ f^*,\ a_n,\ b_n)$
6) Якщо $n==N$ – зупинитись, інакше – Крок 1.

$$D_c = \begin{cases} z \ \to \max \\ \sum_{i,a} T(i,a,d)x_{ia} - z \ \geq 0, d \ \in D \\ \sum_{a \in A} x_{ia} = \ \mu(i), \qquad i \in S \\ z \in R, x_{ia} \geq 0 \end{cases}$$

$$\Phi(\mu, f, a, b)(j) = \frac{\sum_i \mu(i) f(i,a) p_{ij}(a,b)}{\sum_i \mu(i) f(i,a)}, j \in S$$

Модифікація моделі для M гравців

- **M** – кількість гравців
- **S** – множина станів гри
- **A** – множина дій, $A = \{A^i(s), s \in S, i \in \{1, \ldots M\}\}$
- $\boldsymbol{\mu_0} \in P(S)$ – початковий розподіл
- $\mathbf{V = \{v_1, v_2\}}$ – коаліції гравців, де $v_i = \{j_1, j_i, \ldots, j_{M/2}\}, i \in \{1, \ldots M\}$
- $\mathbf{p_{ij}(a)}$ – імовірність переходу з вектором дій $a = (A^1(i), A^2(i), \ldots, A^M(i)), a = a^1 \cup a^2$
- $\mathbf{r_i(a)}$ – функція нагороди
- $\boldsymbol{\beta} \in R$ – коефіцієнт дисконтування

Пошук оптимальних стратегій

1) Визначити $n = 0,\ \mu = \mu_0$
2) Знайти оптимальне рішення $(z^*,\ x^*)$ для $D_{N-n}(\mu)$
3) Запам'ятати i_n та вибрати вектор дій a^1
4) Запам'ятати a^2
5) $n{+}{=}1,\ \mu = \Phi\ (\mu,\ f_n^*,\ a)$
6) Якщо $n{=}{=}N$ – зупинитись, інакше – Крок 1.

$$D_c = \begin{cases} z\ \to \max \\ \sum\limits_{i,a^1} T(i,a^1,d)x_{ia^1} - z\ \geq 0, d\ \in D \\ \sum\limits_{a^1\,\in\, a} x_{ia^1} = \ \mu(i), \qquad i \in S \\ z \in R, x_{ia^1} \geq 0 \end{cases}$$

$$\Phi(\mu, f, a)(j) = \frac{\sum_i \mu(i) f(i,a^1) p_{ij}(a)}{\sum_i \mu(i) f(i,a^1)}\ , j \in S$$

Параметри конкретної гри

- $M = 4$
- $N = 2$
- $S = [1, 2]$
- $A = [(0,1); (0,1); (0,1); (0,1)]$

- $\mu_0(1) = \mu_0(2) = 0.5$
- $\beta = 1$
- $p_{i,j}(a) = \delta_{i,j} = \begin{cases} 1, \text{якщо } i == j \\ 0, \text{якщо } i \neq j \end{cases}$
- $r_1(a) : r_1([1,0,0,0]) = -2$
- $r_2(a) : r_2([1,0,0,0]) = 2$

Результат роботи алгоритму

Початковий стан - 1

Етап 1	[0, 0]	[0, 1]	[1, 0]	[1, 1]
[0, 0]	0.0	0.0	0.0	0.0
[0, 1]	0.0	0.0	0.0	0.5
[1, 0]	0.0	0.0	0.0	0.5
[1, 1]	0.0	0.0	0.0	1.0

Етап 2	[0, 0]	[0, 1]	[1, 0]	[1, 1]
[0, 0]	0.0	0.0	0.0	0.0
[0, 1]	0.0	0.0	0.0	0.5
[1, 0]	0.0	0.0	0.0	0.5
[1, 1]	0.0	0.0	0.0	1.0

$r_1(a) : r_1([1,0,0,0]) = -2$

$r_2(a) : r_2([1,0,0,0]) = 2$

Початковий стан - 2

Етап 1	[0, 0]	[0, 1]	[1, 0]	[1, 1]
[0, 0]	1.0	0.0	0.0	0.0
[0, 1]	0.5	0.0	0.0	0.0
[1, 0]	0.5	0.0	0.0	0.0
[1, 1]	0.0	0.0	0.0	0.0

Етап 2	[0, 0]	[0, 1]	[1, 0]	[1, 1]
[0, 0]	1.0	0.0	0.0	0.0
[0, 1]	0.5	0.0	0.0	0.0
[1, 0]	0.5	0.0	0.0	0.0
[1, 1]	0.0	0.0	0.0	0.0

Результати перебору

- 512 варіантів розвитку гри
- Найбільший виграш – 8
- $r_1(a)$: $r_1([1,0,0,0]) = -2$
- $r_2(a)$: $r_2([1,0,0,0]) = 2$

Стани гри	[1, 1]	
Дії першої коаліції	[1, 1]	[1, 1]
Дії другої коаліції	[1, 1]	[1, 1]
Виграш	8	

Стани гри	[2, 2]	
Дії першої коаліції	[0, 0]	[0, 0]
Дії другої коаліції	[0, 0]	[0, 0]
Виграш	8	

Висновки

- Розглянуто Марківську модель гри для двох гравців
- Модифіковано для багатьох гравців
- Реалізовано алгоритм пошуку оптимальних стратегій
- Представлено результати та порівняно з перебором всіх варіантів