# Wieloetapowy dylemat więźnia

15 stycznia 2025

# Dylemat więźnia

## Przypomnienie

Klasyczny dylemat więźnia to gra dwumacierzowa postaci

| | *lojalny* | *interesowny* |
|---|---|---|
| *lojalny* | $(L, L)$ | $(F, Z)$ |
| *interesowny* | $(Z, F)$ | $(I, I)$ |

gdy $F < I < L < Z$. Poszczególne wypĹ,aty nazwiemy następująco:

- F -Wypłata frajera;
- I -Wypłata dla dwóch interesownych graczy;
- L -Wypłata dla dwóch lojalnych graczy;
- Z -Wypłata dla zdrajcy;

# Dylemat więźnia

Dylemat więźnia w formie funkcyjnej

- Zbiór strategii gracza 1 i 2 to $\{l, i\}$, $l$-lojalność, $i$-interesowność;
- Wypłaty dla obu graczy są $u_i : \{l, i\}^2 \mapsto \mathbb{R}$ postaci

$$\begin{aligned} u_1(l,l) = u_2(l,l) = L; \\ u_1(i,i) = u_2(i,i) = I; \\ u_1(i,l) = u_2(l,i) = Z; \\ u_1(l,i) = u_2(i,l) = F; \end{aligned}$$

# Dylemat więźnia

- Brak porozumienia oznacza, że obaj gracze będą *interesowni*, bo ta strategia dominuje nad byciem *lojalnym*:
  1. Jeśli przeciwnik będzie *interesowny*, gracz też powinien być *interesowny*, bo inaczej otrzyma wypłatę frajera;
  2. Jeśli przeciwnik będzie *lojalnym*, również lepiej być *interesownym*, bo bycie zdrajcą oznacza większą wypłatę (w tej grze dobre imię nie istnieje);
- **Jeśli gracze nie oczekują, że kiedyś dojdzie do rewanżu, będą interesowni.**

A co jeśli dojdzie do rewanżu?

- Można być teraz interesownym, ale należy oczekiwać, że przeciwnik straci do nas zaufanie i wtedy nie zdobędziemy więcej niż $I$;
- Zatem można się zastanowić czy istnieje równowaga gdzie gracze są jednak lojalni gdy gra jest rozgrywana przez wiele rund i wynik jest poznawany po każdej z nich.

# Dylemat więźnia

*n*-etapowy dylemat więźnia, opis algorytmiczny.

1. Etap=1;
2. Jeśli Etap=n, koniec gry;
3. Gracze niezależnie od siebie wybierają strategię lojalność lub interesowność, a następnie poznają wynik gry i pośrednio strategię przeciwnika;
4. Etap$\rightarrow$Etap+1 i idź do kroku 2;

### Uwaga

Gracz dąży do maksymalizacji wypłat w każdym kroku, ale wtedy jego celem byłaby maksymalizacja wszystkich współrzędnych a więc funkcji wektorowej $\mathcal{U}_j : \{l, i\}^n \mapsto \mathbb{R}^n$, co zwykle jest niewygodne ($j = 1, 2$);

# Wieloetapowy dylemat więźnia

Alternatywnie można przyjąć dwa kryteria:

1. Maksymalizacja sumy użyteczności: $U_1 : \{l, i\}^n \mapsto \mathbb{R}$ i $U_2 : \{l, i\}^n \mapsto \mathbb{R}$:
$$U_j^n(x; y) = \sum_{t=1}^{n} u_j(x_t, y_t),$$
gdzie $(x_t, y_t) \in \{l, i\}^n$ dla wszystkich $t$, oraz $x = (x_1, x_2, \ldots, x_n)$ i $y = (y_1, y_2, \ldots, y_n)$.
2. Maksymalizacja dyskontowanej sumy użyteczności: $U_1^\delta : \{l, i\}^n \mapsto \mathbb{R}$ i $U_2^\delta : \{l, i\}^n \mapsto \mathbb{R}$:
$$U_j^{n,\delta}(x; y) = \sum_{t=1}^{n} u_j(x_t, y_t)\delta^{t-1},$$
gdzie $(x_t, y_t) \in \{l, i\}^n$ dla wszystkich $t$, oraz $x = (x_1, x_2, \ldots, x_n)$ i $y = (y_1, y_2, \ldots, y_n)$, natomiast $\delta \in (0, 1)$ jest nazywany *czynnikiem dyskonta*.

# Strategie w grze $n$ etapowej

## Zbiór strategii

Strategie dla obu graczy:

- Strategia gracza 1 w $n$-etapowym dylemacie więźnia jest postaci $\xi := (\xi_1, \xi_2, \ldots, \xi_n)$, gdzie $\xi_1 \in \{l, i\}$, $\xi_2 : \{l, i\} \mapsto \{l, i\}$, oraz dla dowolnego $t = 2, \ldots, n$, $\xi_t : \{i, l\}^{t-1} \mapsto \{i, l\}$;
- Strategia gracza 2 w $n$-etapowym dylemacie więźnia jest postaci $\eta := (\eta_1, \eta_2, \ldots, \eta_n)$, gdzie $\eta_1 \in \{l, i\}$, $\eta_2 : \{l, i\} \mapsto \{l, i\}$, oraz dla dowolnego $t = 2, \ldots, n$, $\eta_t : \{i, l\}^{t-1} \mapsto \{i, l\}$;

# Strategie graczy

Przykłady strategii.

- Gracz stosuje *naiwną* strategię, a więc jest lojalny na każdym kroku;
- Gracz stosuje *bezwzględną* strategię, a więc jest interesowny na każdym kroku;
- Gracz stosuje strategię *wet za wet*, a więc:
  1. zaczyna od bycia lojalnym;
  2. w kolejnym kroku wybiera strategię taką jaką przeciwnik wybrał w poprzednim kroku;
- Gracz stosuje strategię *pamiętliwą*, a więc:
  1. zaczyna od bycia lojalnym;
  2. jeśli przeciwnik choć raz wybrał strategię interesowną, ten do końca gry wybiera strategie interesowne;

## Matematyczny opis strategii

Niech $y = (y_1, y_2, ..., y_n)$ będzie dowolnym wyborem gracza 2.

- Gracz 1 będzie naiwny jeśli: $\xi_1 = l$, oraz dla każdego $t \geq 2$, $\xi_t(y_1, y_2, ..., y_{t-1}) = l$;
- Gracz 1 będzie bezwzględny: $\xi_1 = i$, oraz dla każdego $t \geq 2$, $\xi_t(y_1, y_2, ..., y_{t-1}) = i$;
- Gracz 1 stosuje wet za wet jeśli $\xi_1 = l$, oraz dla każdego $t \geq 2$, $\xi_t(y_1, \ldots, y_{t-1}) = y_{t-1}$;
- Gracz 1 jest pamiętliwy jeśli $\xi_1 = l$, oraz dla każdego $t \geq 2$, $\xi_t(l, l, \ldots, l) = l$ oraz $\xi_t(y_1, y_2, ..., y_n) = i$ w pozostałych przypadkach;

# $n$-osobowy dylemat więźnia

Czy teraz gracze będą lojalni? Niech przeciwnik zastosuje jakąkolwiek strategię. Jak postąpić optymalnie?

1. W ostatnim etapie strategia i tak $i$ dominuje nad $l$, więc wtedy optymalny wybór to $i$, bo wie że przeciwnik go nie ukarze;
2. Trudno oczekiwać, żeby przeciwnik karał lub wynagradzał w ostatnim okresie, gdy wybór $i$ i tak dominuje $l$;
3. Ponieważ gracze spodziewają się zachowania przeciwnika, zatem gra kończy się w etapie $n-1$;
4. W ten sposób można spodziewać się, że przeciwnik w etapie $n-1$ również zastosuje $i$, czyli gra kończy się na etapie $n-2$, później $n-3$ i dojdziemy do wniosku, że gra jest jednoetapowa;
5. Konsekwentnie jedyną równowagą Nasha jest układ strategii gdy obaj gracze są bezwzględni.

# Dylemat więźnia z nieskończonym horyzonetem

- Rozpatrzmy dylemat więźnia ktŕy odbywa się z nieskończonym horyzoncie czasowym, a więc dylemat wiçźnia powtarza się nieskończenie wiele razy;
- Funkcja użyteczności ma postać dyskontowanej funkcji użyteczności:

$$U_j^\delta := \sum_{t=1}^{\infty} u_j(x_t, y_t)\delta^{t-1},$$

gdzie $(x_t)_{t=1}^{\infty}$ i $(y_t)_{t=1}^{\infty}$ są wyborami obu graczy;

- Tak jak w przypadku gry $n$-etapowej, strategie są postaci $\xi_1 \in \{i, l\}$, $\eta_1 \in \{i, l\}$, oraz dla $t > 1$, $\xi : \{i, l\}^{t-1} \mapsto \{i, l\}$ oraz $\eta_t : \{i, l\}^{t-1} \mapsto \{i, l\}$.

# Dylemat więźnia z nieskończonym horyzontem czasowym

Interpretacja dylematu więźnia z $\infty$ horyzontem czasowym:

1. Etap=1
2. Gracze rozgrywają dylemat więźnia i poznają swoje wyniki, a więc i wybór przeciwnika;
3. Niezależnie od dotychczasowego przebiegu gry, z prawdopodobieństwem $\delta$ gra jest kontynuowana a wtedy Etap$\rightarrow$Etap+1 i powrót do punktu 2, a z prawdopodobieństwem $1-\delta$ gra zostaje przerwana.

# Równowaga Nasha w dylemacie więźnia z $\infty$ horyzontem

**Twierdzenia**

W grze z nieskończonym horyzontem profil strategiczny, gdzie obaj gracze stosują wet za wet jest równowagą Nasha jeśli zachodzą warunki

$$Z - L \leq \delta(L - F),$$

oraz

$$Z - L + \delta(I - L) \leq \delta^2(L - F).$$

# Równowaga Nasha w dylemacie więźnia z $\infty$ horyzontem

## Dowód

Załóżmy, że gracz 1 stosuje strategię wet za wet.

- Jeśli gracz 2 stosuje również wet za wet to będzie cały czas lojalny, a gracz 1 będzie lojalny w stosunku do niego. Mamy więc
$$U_2^\delta(wet, wet) := U_2^\delta((l, l, \ldots), (l, l, \ldots)) = \sum_{t=1}^{\infty} u_2(l, l)\delta^{t-1} = \frac{L}{1-\delta}.$$
- Rozpatrzmy alternatywne strategie dla 2, nazwijmy ją $\eta$ i pokażemy, ė gracz 2 nie wyjdzie na nich dobrze.

# Równowaga Nasha w dylemacie więźnia z $\infty$ horyzontem

## Dowód

- Na początek załóżmy, że gracz 2 tylko w kroku $\tau$ złamał zasadę lojalności, w a więc wybiera *l* we wszystkich krokach oprócz $\tau$, natomiast w $\tau$ stosuje *i*;
- Wtedy przeciwnik stosuje *l* we wszystkich krokach oprócz kroku $\tau + 1$, nattomiast w kroku $\tau + 1$ stosuje *i*;
- Zatem wyniki gdy są następujące:
  1. W etapach $1, 2, ..., \tau - 1$ oraz od okresu $\tau + 2$ i po tym okresie obaj otrzymują L;
  2. W etapie $\tau$ gracz 1 otrzyma *F*, a gracz 2 otrzyma *Z*;
  3. W etapie $\tau + 1$ gracz 1 otrzyma *Z*, a gracz 2 otrzyma *F*;
- Zatem gracz 2 otrzyma

$$U_2^{\delta}(wet, \eta) = \sum_{t=1}^{\tau-1} L\delta^{t-1} + Z\delta^{\tau-1} + F\delta^{\tau} + \sum_{t=\tau+2}^{\infty} L\delta^{t-1}.$$

### Dowód

Można przekształcić wyrażenie

$$
\begin{aligned}
U_2^\delta(wet,\eta) &= \sum_{t=1}^{\tau-1} L\delta^{t-1} + Z\delta^{\tau-1} + F\delta^{\tau} + \sum_{t=\tau+2}^{\infty} L\delta^{t-1} \\
&= \sum_{t=1}^{\infty} L\delta^{t-1} + (Z-L)\delta^{\tau-1} + (F-L)\delta^{\tau} \\
&= \frac{L}{1-\delta} + \delta^{\tau-1}(Z-L+\delta(F-L)), \\
&= U_2^\delta(wet,wet) + \delta^{\tau-1}(Z-L+\delta(F-L))
\end{aligned}
$$

Zatem $U_2^\delta(wet,\eta) \leq U_2^\delta(wet,wet)$ o ile $\delta \geq \frac{Z-L}{L-F}$.

## Dowód

- Załóżmy, że gdy gracz 2 wyłamał się $n$ krotnie i nazwijmy strategię $\eta^n$ to zajdzie

$$U_2^\delta(wet, \eta^n) \leq U_2^\delta(wet, wet);$$

- Niech teraz gracz 2 wyłamie się w $n+1$ krokach, mamy dwa przypadki
  - Gracz 2 po raz $n$ wyłamał się w kroku $\tau$ a po raz $n+1$ w kroku $\tau' > \tau + 1$;
  - Gracz 2 po raz $n$ wyłamał się w kroku $\tau$ a po raz $n+1$ w kroku $\tau' = \tau + 1$;

# Równowaga Nasha w dylemacie więźnia z $\infty$ horyzontem

## Dowód

Niech $\tau' > \tau + 1$, czyli gracz 2 ostatni raz wyłamuje się w czasie $\tau + 2$ lub później:

- Wtedy

$$U_2^\delta(wet, \eta^{n+1}) = \sum_{t=1}^{\tau'-1} u_2(wet, \eta_t^n)\delta^{t-1} + \delta^{\tau'} Z + \delta^{\tau'+1} F + \sum_{t=\tau'+2} u_2(wet, \eta_t^n),$$

  bo wypłaty między strategiami $\eta^n$ i $\eta^{n+1}$ różnią się tylko w okresach $\tau'$ i $\tau' + 1$;

- Stąd i z założenia indukcyjnego

$$\begin{aligned}
U_2^\delta(wet, \eta^{n+1}) &= \sum_{t=1}^{\infty} u_2(wet, \eta_t^n)\delta^{t-1} + \delta^{\tau'}(Z - L) + \delta^{\tau'+1}(F - L), \\
&= \sum_{t=1}^{\infty} u_2(wet, \eta_t^n)\delta^{t-1} + \delta^{\tau'} [Z - L + \delta(F - L)] \\
&= U_2^\delta(wet, \eta^n) + \delta^{\tau'} [Z - L + \delta(F - L)] \\
&\leq U_2^\delta(wet, wet)
\end{aligned}$$

  gdzie ostatnia linijka wynika z założenia indukcyjnego i z $Z - L \leq \delta(L - F)$), zatem $n + 1$ krotne wyłamanie się również się nie opłaca.

**Dowód**

Niech $\tau' = \tau + 1$, czyli gracz 2 ostatni raz wyłamuje się zaraz po przedostatnim razie:

- Wtedy
$$U_2^\delta(wet, \eta^{n+1}) = \sum_{t=1}^{\tau'-1} u_2(wet, \eta_t^n)\delta^{t-1} + \delta^{\tau'} Z + \delta^{\tau'+1} I + \delta^{\tau'+2} F + \sum_{t=\tau'+3} u_2(wet, \eta_t^n),$$
bo wypłaty między strategiami $\eta^n$ i $\eta^{n+1}$ różnią się tylko w okresach $\tau'$ i $\tau' + 1$ i $\tau' + 2$;
- Stąd i z założenia indukcyjnego
$$\begin{aligned} U_2^\delta(wet, \eta^{n+1}) &= \sum_{t=1}^{\infty} u_2(wet, \eta_t^n)\delta^{t-1} + \delta^{\tau'}(Z - L) + \delta^{\tau'+1}(I - L) + \delta^{\tau'+2}(F - L), \\ &= \sum_{t=1}^{\infty} u_2(wet, \eta_t^n)\delta^{t-1} + \delta^{\tau'}\left[Z - L + \delta(I - L) - \delta^2(L - F)\right] \\ &= U_2^\delta(wet, \eta^n) + \delta^{\tau'}\left[Z - L + \delta(I - L) - \delta^2(L - F)\right] \\ &\leq U_2^\delta(wet, wet) \end{aligned}$$
gdzie ostatnia linijka wynika z założenia indukcyjnego i z $Z - L + \delta(I - L) \leq \delta^2(L - F)$, zatem $n + 1$ krotne wyłamanie się również się nie opłaca.

# Równowaga Nasha w dylemacie więźnia z $\infty$ horyzontem

**Dowód**

Ostatecznie
$$U_2^\delta(wet, wet) \geq U_2^\delta(wet, \eta^n)$$
dla wszystkich $n$. Gdy rozpatrzymy dowolną strategię $\eta$, przyjmijmy, $\eta^n$ taką strategię, że gracz postępuje jak $\eta$ do momentu gdy wyłamie się po raz $n$-ty.