# Ograniczenia informacji wzajemnej a estymacja ukrytego rzędu Markowa

Łukasz Dębowski
ldebowsk@ipipan.waw.pl

Instytut Podstaw Informatyki PAN

Seminarium ZAMS, 23 listopada 2020

# Hipoteza Hilberga

- $H(X)$ – entropia Shannona.
- $I(X;Y) := H(X) + H(Y) - H(X,Y)$ – info. wzajemna.
- $K(w)$ – złożoność Kołmogorowa.
- $J(u;w) := K(u) + K(w) - K(u,w)$ – alg. info. wzajemna.
- Hipoteza Hilberga: Dla tekstów w języku naturalnym

$$J(x_1^n; x_{n+1}^{2n}) \propto n^\beta, \quad \beta > 0. \tag{1}$$

- Jeżeli dla procesu stacjonarnego $(X_i)_{i\in\mathbb{Z}}$ zachodzi

$$I(X_1^n; X_{n+1}^{2n}) \propto n^\beta, \quad \beta > 0, \tag{2}$$

  to proces nie może być ukrytym procesem Markowa o skończonej liczbie stanów ukrytych.

Chcemy pokazać, że to samo zachodzi dla (1)
– również, gdy p-stwa przejścia są nieobliczalne.

# Procesy unifilarne (ogólne)

## Definicja

Proces $(X_i)_{i\in\mathbb{N}}$ nad przeliczalnym alfabetem $\mathbb{X}$ nazywa się unifilarnym względem $(Y_i)_{i\in\mathbb{N}}$, gdy $(Y_i)_{i\in\mathbb{N}}$ jest procesem nad przeliczalnym alfabetem $\mathbb{Y}$ takim, że

1. $P(X_i|Y_1^i, X_1^{i-1}) = \varepsilon(X_i|Y_i)$ – ukryta markowskość,
2. $Y_{i+1} = \tau(Y_i, X_i)$ – determinizm automatowy

dla pewnych $\varepsilon : \mathbb{X} \times \mathbb{Y} \to [0, 1]$ oraz $\tau : \mathbb{Y} \times \mathbb{X} \to \mathbb{Y}$.

*Przykład 1: Procesy Markowa $n$-tego bądź zmiennego rzędu.*

*Przykład 2: Ciekawa jest też klasa procesów unifilarnych, gdzie $\mathbb{X} = \{0, 1\}$, $\mathbb{Y} = \{0, 1\}^*$, a funkcje $\varepsilon$ i $\tau$ są obliczalne. To m.in. procesy symulujące obliczenia maszyny Turinga z szumem.*

*Przykład 3: Położenie $Y_i := P(X_i^\infty|X_{-\infty}^{i-1})$ spełnia warunki 1. i 2., ale niekoniecznie nad przeliczalnym alfabetem – również dla nieunifilarnych procesów o skończonej liczbie stanów ukrytych.*

# Modele unifilarne (skończone, HMM)

Zakładamy, że alfabet $\mathbb{X}$ jest skończony.

Unifilarne rozkłady o $k = 1, 2, 3, \ldots$ stanach ukrytych:

$$\mathbb{P}(x_1^n, y_1^n | k, \pi, \tau, \varepsilon) := \pi(y_1)\varepsilon(x_1|y_1) \prod_{i=2}^{n} 1\{y_i = \tau(y_{i-1}, x_{i-1})\} \varepsilon(x_i|y_i), \tag{3}$$

gdzie

- $\pi : \{1, .., k\} \to [0, 1]$, $\sum_y \pi(y) = 1$,
- $\tau : \{1, .., k\} \times \mathbb{X} \to \{1, .., k\}$,
- $\varepsilon : \mathbb{X} \times \{1, .., k\} \to [0, 1]$, $\sum_x \varepsilon(x|y) = 1$.

Rozkład brzegowy:

$$\mathbb{P}(x_1^n | k, \pi, \tau, \varepsilon) := \sum_{y_1^n} \mathbb{P}(x_1^n, y_1^n | k, \pi, \tau, \varepsilon). \tag{4}$$

# Trzy pojęcia z teorii MDL

Maksymalna wiarogodność – ML:

$$\hat{\mathbb{P}}(x_1^n|k) := \max_{y,\tau,\varepsilon} \mathbb{P}(x_1^n|k, y, \tau, \varepsilon). \tag{5}$$

Unormowana maksymalna wiarogodność – NML (Shtarkov):

$$\mathbb{P}(x_1^n|k) := \frac{\hat{\mathbb{P}}(x_1^n|k)}{\sum_{z_1^n \in \mathbb{X}^n} \hat{\mathbb{P}}(z_1^n|k)} \le \hat{\mathbb{P}}(x_1^n|k). \tag{6}$$

Rozkład uniwersalny (Ryabko):

$$\mathbb{P}(x_1^n) := \sum_{k=1}^{\infty} w_k \mathbb{P}(x_1^n|k), \quad w_k := \frac{1}{k} - \frac{1}{k+1}. \tag{7}$$

# Proste ograniczenia

Rząd maksymalizujący:

$$\mathbb{G}(x_1^n) := \arg\max_{k \geq 1} \mathbb{P}(x_1^n | k). \tag{8}$$

Ograniczenie dwustronne rozkładu Ryabki przez NML:

$$\mathbb{P}(x_1^n | \mathbb{G}(x_1^n)) \geq \mathbb{P}(x_1^n) \geq w_k \mathbb{P}(x_1^n | k). \tag{9}$$

„Superaddytywność" pseudorozkładu ML:

$$\hat{\mathbb{P}}(x_1^n | k) \hat{\mathbb{P}}(x_{n+1}^{n+m} | k) \geq \hat{\mathbb{P}}(x_1^{n+m} | k). \tag{10}$$

## Złożoność statystyczna

Używamy logarytmu binarnego $\log x := \log_2 x$.

Złożoność statystyczna (klasy rozkładów):

$$\mathbb{C}(n|k) := \log \frac{\hat{\mathbb{P}}(x_1^n|k)}{\mathbb{P}(x_1^n|k)} = \log \sum_{z_1^n \in \mathbb{X}^n} \hat{\mathbb{P}}(z_1^n|k) \le n \log |\mathbb{X}| . \quad (11)$$

Zbiór różnych parametrów maksymalnej wiarogodności:

$$\mathcal{P}_{kn} := \left\{ (y, \tau, \varepsilon) : \exists_{x_1^n} \mathbb{P}(x_1^n|k, y, \tau, \varepsilon) = \hat{\mathbb{P}}(x_1^n|k) \right\} . \quad (12)$$

Zauważmy, że

$$\mathbb{C}(n|k) = \log \sum_{x_1^n \in \mathbb{X}^n} \hat{\mathbb{P}}(x_1^n|k) \le \log |\mathcal{P}_{kn}| . \quad (13)$$

# Złożoność ukrytych modeli Markowa

Dla ustalonego $(y, \tau)$, $\mathbb{P}(x_1^n|k, y, \tau, \varepsilon)$ jest maksymalizowane dla

$$\varepsilon(b|a) = \frac{\sum_{i=1}^{n} 1\{(y_i, x_i) = (a, b)\}}{\sum_{i=1}^{n} 1\{y_i = a\}}, \tag{14}$$

gdzie $y_1 = y$ oraz $y_i = \tau(y_{i-1}, x_{i-1})$ dla $i \geq 2$.

Różnych rozkładów $\varepsilon$ jest nie więcej niż $(n+1)^{k|\mathbb{X}|}$.

Różnych kombinacji $(y, \tau)$ jest nie więcej niż $k \cdot k^{k|\mathbb{X}|}$.

Zatem

$$\mathbb{C}(n|k) \leq \log |\mathcal{P}_{kn}| \leq [k\,|\mathbb{X}| + 1] \log[k(n+1)]. \tag{15}$$

# Uniwersalność rozkładu Ryabki

Zakładamy, że alfabet $\mathbb{X}$ jest skończony.

Entropie warunkowe:

$$h_k^P := H(X_0|X_{-k}^{-1}) = \mathbf{E}\left[-\log P(X_0|X_{-k}^{-1})\right].$$

Intensywność entropii:

$$h^P := \lim_{n\to\infty} \frac{H(X_1^n)}{n} = \inf_k h_k^P = H(X_0|X_{-\infty}^{-1}).$$

## Twierdzenie

Dla stacjonarnego ergodycznego procesu $(X_i)_{i\in\mathbb{Z}}$,

$$\lim_{n\to\infty} \frac{1}{n}\left[-\log \mathbb{P}(X_1^n)\right] = h^P \text{ prawie na pewno,} \tag{16}$$

$$\lim_{n\to\infty} \frac{1}{n}\mathbf{E}\left[-\log \mathbb{P}(X_1^n)\right] = h^P. \tag{17}$$

# Dowód

Dla $\tau(X_1^k, X_{k+1}) = X_2^{k+1}$ i $\varepsilon(X_{k+1}|X_1^k) = P(X_{k+1}|X_1^k)$ mamy

$$\hat{\mathbb{P}}(X_1^n|\,|\mathbb{X}|^k) \geq \mathbb{P}(X_1^n|\,|\mathbb{X}|^k, X_{-k+1}^0, \tau, \varepsilon) = \prod_{i=1}^n P(X_i|X_{i-k}^{i-1}).$$

Zatem z górnego ograniczenia (15) na złożoność statystyczną i z twierdzenia ergodycznego Birkhoffa,

$$\limsup_{n\to\infty} \frac{1}{n}\left[-\log \mathbb{P}(X_1^n|\,|\mathbb{X}|^k)\right] = \limsup_{n\to\infty} \frac{1}{n}\left[-\log \hat{\mathbb{P}}(X_1^n|\,|\mathbb{X}|^k) + \mathbb{C}(n|\,|\mathbb{X}|^k)\right] \leq h_k^P \text{ p.n.p.}$$

Zatem z górnego ograniczenia (9) i lematu Barrona, otrzymujemy (16). Obserwując $\mathbb{P}(X_1^n) \geq w_n \mathbb{P}(X_1^n|n) = w_n |\mathbb{X}|^{-n}$, stąd otrzymujemy (17) ze zbieżności majoryzowanej.

# Zgodność i nieobciążoność estymatora rzędu

Zakładamy, że alfabet $\mathbb{X}$ jest skończony.

Unifilarny rząd procesu:

$$M^P := \inf \left\{ k : \exists_{\pi,\tau,\varepsilon} \forall_{n \ge 1, x_1^n} P(X_1^n = x_1^n) = \mathbb{P}(x_1^n | k, \pi, \tau, \varepsilon) \right\}.$$

Estymator rzędu:

$$\mathbb{M}(x_1^n) := \min \left\{ k : \hat{\mathbb{P}}(x_1^n | k) \ge w_n \mathbb{P}(x_1^n) \right\}, \quad w_n := \frac{1}{n} - \frac{1}{n+1}.$$

Mamy $\mathbb{M}(x_1^n) \le n$, gdyż $\hat{\mathbb{P}}(x_1^n | k) = 1$ dla $k \ge n$.

**Twierdzenie**

Dla stacjonarnego ergodycznego procesu $(X_i)_{i \in \mathbb{Z}}$,

$$\lim_{n \to \infty} \mathbb{M}(X_1^n) = M^P \text{ prawie na pewno,} \tag{18}$$

$$\lim_{n \to \infty} \mathbf{E}\, \mathbb{M}(X_1^n) = M^P. \tag{19}$$

# Dowód braku przeszacowania

Z nierówności $\hat{\mathbb{P}}(X_1^n|M^P) \geq P(X_1^n|Y_1)$ i lematu Barrona otrzymujemy

$$P\left(\mathbb{M}(X_1^n) > M^P\right) \leq P\left(\hat{\mathbb{P}}(X_1^n|M^P) < w_n\mathbb{P}(X_1^n)\right) \leq P\left(\frac{w_n\mathbb{P}(X_1^n)}{P(X_1^n|Y_1)} > 1\right) \leq w_n. \tag{20}$$

Ponieważ $\sum_{n=1}^{\infty} w_n = 1$, niemożliwość przeszacowania wynika z lematu Borela-Cantellego.

# Dowód braku niedoszacowania (1/4)

Ponieważ rozkład Ryabki jest uniwersalny w sensie (16), a $-\log w_n = o(n)$, wystarczy dowieść

$$\liminf_{n\to\infty} \frac{1}{n} \left[-\log \hat{\mathbb{P}}(X_1^n|k)\right] > h^P \text{ p.n.p. dla } k < M^P. \qquad (21)$$

W dalszej kolejności pokażemy, że lewa strona równa jest pewnej entropii warunkowej $h^P_{[k]}$, która jest większa niż $h^P$ dla $k < M^P$.

## Dowód braku niedoszacowania (2/4)

Zauważmy, że dla skończonego zbioru $\mathcal{M}$,

$$\liminf_{n\to\infty} \min_{m\in\mathcal{M}} a_{nm} = \min_{m\in\mathcal{M}} \liminf_{n\to\infty} a_{nm}. \tag{22}$$

W naszym przypadku, zbiór par $(y, \tau)$ dla ustalonego $k$ jest skończony. Zatem

$$\begin{aligned} &\liminf_{n\to\infty} \min_{y,\tau,\varepsilon} \frac{1}{n} \left[-\log \mathbb{P}(X_1^n|k, y, \tau, \varepsilon)\right] = \\ &\min_{y,\tau} \liminf_{n\to\infty} \min_{\varepsilon} \frac{1}{n} \left[-\log \mathbb{P}(X_1^n|k, y, \tau, \varepsilon)\right]. \end{aligned} \tag{23}$$

# Dowód braku niedoszacowania (3/4)

Z teorii procesów asymptotycznie średnio stacjonarnych – Kieffer i Rahe (1981) oraz Gray i Kieffer (1980) – istnieją procesy stacjonarne $(\bar{Y}_{i+1}^{y,\tau})_{i\in\mathbb{Z}}$ takie, że zachodzi rekursja $\bar{Y}_{i+1}^{y,\tau} = \tau(\bar{Y}_i^{y,\tau}, X_i)$ oraz

$$\begin{aligned} &\min_{y,\tau} \liminf_{n\to\infty} \min_{\varepsilon} \frac{1}{n} \left[-\log \mathbb{P}(X_1^n|k,y,\tau,\varepsilon)\right] \\ &= \min_{y,\tau} \liminf_{n\to\infty} \min_{\varepsilon} \frac{1}{n} \sum_{i=1}^{n} \left[-\log \varepsilon(X_i|Y_i^{y,\tau})\right] \\ &= h_{[k]}^P := \min_{y,\tau} \mathbf{E} \left[-\log P(X_i|\bar{Y}_i^{y,\tau})\right] \text{ p.n.p.} \end{aligned} \tag{24}$$

Niech $y$ i $\tau$ będą minimalizatorami. Oznaczmy $\bar{Y}_i := \bar{Y}_i^{y,\tau}$.

Co się dzieje, gdy $h_{[k]}^P := H(X_i|\bar{Y}_i) = h^P$?

# Dowód braku niedoszacowania (4/4)

Ponieważ $\bar{Y}_{i+1} = \tau(\bar{Y}_i, X_i)$, mamy

$$H(X_i|\bar{Y}_i) - H(X_i|X_1^{i-1}, \bar{Y}_1) = I(X_i; X_1^{i-1}, \bar{Y}_1^{i-1}|\bar{Y}_i). \quad (25)$$

Zatem ze stacjonarności $(X_i, \bar{Y}_i^{y,\tau})_{i\in\mathbb{Z}}$,

$$\sum_{i=1}^{n} H(X_i|\bar{Y}_i) = H(X_1^n|\bar{Y}_1) + \sum_{i=0}^{n-1} I(X_j; X_{j-i}^{j-1}, \bar{Y}_{j-i}^{j-1}|\bar{Y}_j). \quad (26)$$

Dzieląc przez $n$ i kładąc $n \to \infty$ otrzymujemy

$$h_{[k]}^P := H(X_i|\bar{Y}_i) = h^P + I(X_j; X_{-\infty}^{j-1}, \bar{Y}_{-\infty}^{j-1}|\bar{Y}_j). \quad (27)$$

A zatem jeżeli $h_{[k]}^P = h^P$, to $I(X_i; X_{-\infty}^{i-1}, \bar{Y}_{-\infty}^{j-1}|\bar{Y}_i) = 0$.
Czyli $(X_i)_{i\in\mathbb{N}}$ jest unifilarnym procesem o $\leq k$ stanach ukrytych o rozkładzie danym przez $\bar{Y}_i$.

# Dowód braku obciążenia

Ponieważ $\mathbb{M}(x_1^n) \le n$ oraz $P\left(\mathbb{M}(X_1^n) > M^P\right) \le w_n$, mamy

$$\mathbf{E}\,\mathbb{M}(X_1^n) \le M^P + nP(\mathbb{M}(x_1^n) > M^P) = M^P + \frac{1}{n+1}. \quad (28)$$

Z drugiej strony ze zgodności i lematu Fatou wynika

$$M^P = \mathbf{E} \liminf_{n\to\infty} \mathbb{M}(X_1^n) \le \liminf_{n\to\infty} \mathbf{E}\,\mathbb{M}(X_1^n). \quad (29)$$

# Tempo wzrostu potęgowego

Wykładnik Hilberga:

$$\operatorname*{hilb}_{n\to\infty} s(n) := \limsup_{n\to\infty} \frac{\log\max\{1, s(n)\}}{\log n}. \tag{30}$$

Mamy $\operatorname{hilb}_{n\to\infty} n^{\beta} = \beta$ for $\beta \geq 0$.

## Twierdzenie

Dla funkcji $\mathfrak{S} : \mathbb{N} \to \mathbb{R}$, połóżmy $\mathfrak{J}(n) := 2\mathfrak{S}(n) - \mathfrak{S}(2n)$.
Jeżeli $\lim_{n\to\infty} \mathfrak{S}(n)/n = \mathfrak{s}$ dla $\mathfrak{s} \in \mathbb{R}$, to

$$\operatorname*{hilb}_{n\to\infty} (\mathfrak{S}(n) - n\mathfrak{s}) \leq \operatorname*{hilb}_{n\to\infty} \mathfrak{J}(n) \tag{31}$$

z równością, gdy $\mathfrak{S}(n) - n\mathfrak{s} \geq 0$.

# Wykładniki Hilberga dla procesu

$(X_i)_{i\in\mathbb{Z}}$ – proces stacjonarny nad skończonym alfabetem $\mathbb{X}$.

Entropia nadwyżkowa:

$$E^P := \lim_{n\to\infty}\left[H(X_1^n) - nh^P\right] = \lim_{n\to\infty} I(X_1^n; X_{n+1}^{2n}).$$

Z nierówności przetwarzania informacji $E^P \le \log M^P$.

Możemy zdefiniować wykładniki Hilberga $\beta_1^P \le \beta_2^P \le \beta_3^P$, gdzie

$$\beta_1^P := \operatorname*{hilb}_{n\to\infty}\left[H(X_1^n) - nh^P\right] = \operatorname*{hilb}_{n\to\infty} I(X_1^n; X_{n+1}^{2n}) \le 1,$$

$$\beta_2^P := \operatorname*{hilb}_{n\to\infty} \mathbf{E}\left[K(X_1^n) - nh^P\right] = \operatorname*{hilb}_{n\to\infty} \mathbf{E}\, J(X_1^n; X_{n+1}^{2n}) \le 1,$$

$$\beta_3^P := \operatorname*{hilb}_{n\to\infty} \mathbf{E}\left[-\log \mathbb{P}(X_1^n) - nh^P\right]$$

$$= \operatorname*{hilb}_{n\to\infty} \mathbf{E}\left[-\log \mathbb{P}(X_1^n) - \log \mathbb{P}(X_{n+1}^{2n}) + \log \mathbb{P}(X_1^{2n})\right] \le 1.$$

# Ograniczenie górne informacji wzajemnej

Zakładamy, że alfabet $\mathbb{X}$ jest skończony.

**Twierdzenie**

Dla procesu stacjonarnego $(X_i)_{i\in\mathbb{Z}}$,

$$\beta_3^P \le \beta_{\mathbb{M}}^P := \operatorname*{hilb}_{n\to\infty} \mathbf{E}\, \mathbb{M}(X_1^n). \tag{32}$$

**Wniosek**

Dla unifilarnych procesów o skończonej liczbie stanów ukrytych zachodzi $\beta_1^P = \beta_2^P = \beta_3^P = \beta_{\mathbb{M}}^P = 0$.

**Wniosek**

Dla nieunifilarnych procesów o skończonej liczbie stanów ukrytych zachodzi $\beta_1^P = \beta_2^P = 0$.

# Dowód

Notacja: $\lfloor r := -\log r$.

Kładąc $k := \mathbb{M}(X_1^{2n}) \le 2n$, z nierówności (9) i (10) mamy

$$\begin{aligned}
&\lfloor \mathbb{P}(X_1^n) + \lfloor \mathbb{P}(X_{n+1}^{2n}) - \lfloor \mathbb{P}(X_1^{2n}) \\
&\le \lfloor w_k \mathbb{P}(X_1^n|k) + \lfloor w_k \mathbb{P}(X_{n+1}^{2n}|k) - \lfloor \hat{\mathbb{P}}(X_1^{2n}|k)/w_{2n} \\
&\le 2\mathbb{C}(n|k) + \lfloor w_k^2 w_{2n} + \lfloor \hat{\mathbb{P}}(X_1^n|k) + \lfloor \hat{\mathbb{P}}(X_{n+1}^{2n}|k) - \lfloor \hat{\mathbb{P}}(X_1^{2n}|k) \\
&\le 2\mathbb{C}(n|k) + \lfloor w_k^2 w_{2n}. \qquad (33)
\end{aligned}$$

Zatem z nierówności (15) otrzymujemy

$$\beta_3^P \le \operatorname*{hilb}_{n\to\infty} \mathbf{E}\, \mathbb{C}(n|\mathbb{M}(X_1^{2n})) \le \operatorname*{hilb}_{n\to\infty} \mathbf{E}\, \mathbb{M}(X_1^n). \qquad (34)$$

# Ograniczenie dolne informacji wzajemnej

Zakładamy, że alfabet $\mathbb{X}$ jest skończony.

**Twierdzenie (Dębowski, 2018)**

Niech $g : \mathbb{N} \times \mathbb{X}^* \to \{0, 1, 2\}$ – funkcja obliczalna, zaś $z = (z_k)_{k\in\mathbb{N}}$ – sekwencja algorytmicznie losowa, tzn. $K(z_1^n) \geq n - c$ dla $c > 0$. Połóżmy

$$U_g(x_1^n|z) := \min\left\{k \geq 1 : g(k, x_1^n) \neq z_k\right\}. \tag{35}$$

Dla procesu stacjonarnego $(X_i)_{i\in\mathbb{Z}}$,

$$\beta^P_{g,z} := \operatorname*{hilb}_{n\to\infty} \mathbf{E}\, U_g(X_1^n|z) \leq \beta^P_2. \tag{36}$$

*Przykład: Wynik nieskończonego rzutu uczciwą monetą jest sekwencją algorytmicznie losową prawie na pewno. (lemat Barrona)*

# Procesy perygraficzne

**Definicja (Debowski, 2018)**

Proces stacjonary $(X_i)_{i\in\mathbb{Z}}$ nazywa się perygraficznym, gdy $\beta^P_{g,z} > 0$ dla pewnej f-cji obliczalnej $g : \mathbb{N} \times \mathbb{X} \to \{0, 1, 2\}$ i sekwencji algorytmicznie losowej $z = (z_k)_{k\in\mathbb{N}}$.

Dla nieskończonego alfabetu $\mathbb{X}$ prostym przykładem procesu perygraficznego jest proces Santa Fe

$$X_i = (K_i, z_{K_i}), \tag{37}$$

gdzie $K_i$ jest procesem IID o rozkładzie Zipfa $P(K_i = k) \propto k^{-\alpha}$ dla $\alpha > 1$. Wówczas $\beta^P_{g,z} = 1/\alpha$.

*Uwaga: Procesy perygraficzne są nieobliczalne, ale mogą być unifilarne. Obliczalne procesy unifilarne mogą mieć perygraficzne składowe ergodyczne prawie na pewno. (Ciekawe, kiedy tak jest.)*

# Procesy wyroczniowe

*W pętli generujemy losowe słowo binarne* $\boldsymbol{y}$, *symbol* $\mathbf{2}$, *a potem bit* $\boldsymbol{z_{\phi(y)}}$ *sekwencji algorytmicznie losowej i powtarzamy...*

## Definicja

Niech $\psi : \mathbb{N} \to \{0,1\}^*$ będzie rozwinięciem binarnym bez początkowej cyfry $\mathbf{1}$, tzn. $\psi(1) = \lambda$, $\psi(2) = 0$, $\psi(3) = 1$, $\psi(4) = 00$, ... Połóżmy $\phi = \psi^{-1}$. Niech $z = (z_k)_{k\in\mathbb{N}}$ będzie sekwencją algorytmicznie losową.

Proces Oracle($\theta$) z $\theta \in [0,1]$ to następujący proces unifilarny:

- $\mathbb{X} = \{0,1,2\}$,
- $\mathbb{Y} = \{a,b\} \times \{0,1\}^*$,
- $\varepsilon(x|ay) = \frac{\theta}{2}$, $\tau(ay,x) = ayx$ dla $x \in \{0,1\}$, $y \in \{0,1\}^*$,
- $\varepsilon(2|ay) = (1-\theta)$, $\tau(ay,2) = by$ dla $y \in \{0,1\}^*$,
- $\varepsilon(z_{\phi(y)}|by) = 1$, $\tau(by, z_{\phi(y)}) = a$ dla $y \in \{0,1\}^*$.

# Ograniczenia informacji wzajemnej są ścisłe!

Aby wyznaczyć wykładnik Hilberga $\beta_{g,z}$, użyjemy predyktora

$$g(k, x_1^n) := \begin{cases} 0 & \text{if } 2\_\psi(k)20 \sqsubseteq x_1^n \text{ and } 2\_\psi(k)21 \not\sqsubseteq x_1^n, \\ 1 & \text{if } 2\_\psi(k)21 \sqsubseteq x_1^n \text{ and } 2\_\psi(k)20 \not\sqsubseteq x_1^n, \\ 2 & \text{else.} \end{cases} \tag{38}$$

W powyższym symbol '_' dopasowuje się do dowolnego symbolu.

## Twierdzenie

Dla predyktora (38) i stacjonarnego procesu Oracle($\theta$),

$$\beta^P_{g,z} = \beta^P_2 = \beta^P_3 = \beta^P_{\mathbb{M}} = \frac{1}{1 - \log \theta}. \tag{39}$$

# Program badawczy na przyszłość

**Procesy BDHM (Binary Denumerable Hidden Markov):**

Klasa ukrytych procesów Markowa o $\mathbb{X} = \{\mathbf{0}, \mathbf{1}\}$ i $\mathbb{Y} = \{\mathbf{0}, \mathbf{1}\}^*$.

Warto zbadać interakcje warunków takich jak:

1. unifilarność – determinizm automatowy – minimalność.
2. (mocna nie)ergodyczność – perygraficzność – obliczalność.
3. stacjonarność rozkładu stanów ukrytych/symboli emitowanych – średnia stacjonarna – asymptotycznie średnia stacjonarność.
4. rozstrzygalność i złożoność czasowa w/w charakteryzacji.

Zbliżone podejścia: Computational Mechanics z Santa Fe Institute.