Információs rendszerek elméleti alapjai

Információelmélet

Relatív bonyolultság

Kolmogorov-Chaitin-Solomonoff alaptétele, invariancia tétel (újabban)

<u>Tétel</u>: Létezik f_0: $N \to N$ optimális parciális rekurzív függvény, amelyhez bármely f: $N \to N$ parciális rekurzív függvény esetében létezik k_f konstans, hogy

$C_{f0}(x) \leq C_f(x) + k_f$, $\forall$ x-re

Következmény: (invariancia) f_0, g_0 optimális függvények,

$| C_{f0}(x) - C_{g0}(x) | < k_{f0,g0,}$ $\forall$ x-re

Tehát $C_{f0}(x)$ konstans erejéig egyértelmű

Relatív bonyolultság

Rögzítsük mostantól az f_0 optimális függvényt (Általában jó az univerzális Turing-gép megfelelő változata)

Definició: Az $x \in N$, Ω Kolmogorv bonyolultságán $C(x) = C_{f0}(x)$ függvényt értjük

Intuitív háttér

Invariancia tétel

Bizonyítása :

Tetszőleges f, x. Ha $C_f(x) = k$, akkor legyen

$f(p)=x$, és $l(p)=k$. $C_f(x)=\min\{l(p): f(p)=x\}$

Vegyük az $U(g,n)$ univerzális parciális rekurzív függvényt.

Legyen f, sorszáma n, azaz

$U(p,n)= f(y)$, A $z=n'p$ kódból

$f_0(z)=U(\pi_{12}(z), \pi_{11}(z))= U(p,n)= f(p)= x$

Ezért $C_{f0}(x) \leq l(z)=l(n'p)= l(p)+2l(l(n))+l(n)+1= C_f(x) +k_f$

Tehát f_0 optimális. QED.

Invariancia tétel
Tulajdonságok

Két konkrét tulajdonsággal lehet bizonyítást végezni:

a) Az Invariancia tétel alapján felülről becsülhető $C(x)$ konkrét $C_f(x)$ –ből.

b) $C(x) = k$, jelentése: létezik egy konkrét *k* hosszú *p* szó, amelyre $l(p)=k$.

1. Tulajdonság: $C(x) \leq l(x) + k_1$
   Bizonyítás: $f(p)=p$-vel , az a) tulajdonságból.
2. lim $C(x) = \propto$, $x \rightarrow \propto$.
   Bizonyítás: ha $\propto$ sokszor N korlát alatt maradna, az a b) tulajdonság miatt csak 2^N, kód használatát jelentené, ami lehetetlen

Invariancia tétel
Tulajdonságok

3. $C(x)$ totális (teljes) függvény, de nem kiszámítható.
   Bizonyítás: Indirekten: ha az volna, akkor legyen
   x_1: ahol $C(x)$ először nagyobb 2-nél,
   x_2 : ahol $C(x)$ először nagyobb 2^2-nél,
   .
   .
   x_k: $x_k > x_{k-1}$, és $C(x)$ először nagyobb 2^k-nál,
   $C(x_k) > 2^k$.
   .

Legyen $f(p)=x_p$.
$C_f(x_k) \leq log_2 k \leq log_2(log_2 C(x_k))$ azért, mert
$C_f(x_k)=\min\{l(p): f(p)=x_k\}=\min\{l(k): f(k)=x_k\}$, és $l(k)= log_2 k$, $log_2 C(x_k) > k$
Nyilván:
$C(x_k) \leq C_f(x_k)+k_f \leq log_2(log_2 C(x_k))+ k_f$
Csak véges sok x_k-ra teljesülhet, ellentmondás. QED.

Invariancia tétel
Tulajdonságok

Az a) és b) tulajdonságok alapján a $C(x)$ – el mégis lehet valamit kezdeni.

Tulajdonság:

$|\{x : l(x)=n,\ C(x) < n-\delta\}| < 2^{n-\delta}$

a lehetséges kódok száma b) szerint :

$1+2+4+\ldots+ 2^{n-\delta-1} = 2^{n-\delta}-1$

Jelentés: a szavak több mint a fele nem tömöríthető!

Invariancia tétel
Tulajdonságok

Tulajdonság:

i. lim C (x) =$\propto$, x$\rightarrow \propto$.; vagyis nem korlátos fv. (a b) miatt)

ii. Legyen m (x) = min C (y), ahol $y \geq x$, az alsó burkoló. m a legnagyobb monoton növő alsó korlátja C-nek. m (x) nem korlátos

iii. Bármely olyan végtelenbe tartó, monoton, parciális rekurzív $\phi(x)$ fv.-re , amelyre valamilyen x_0-tól ($\phi(x) \rightarrow \propto$) , $m(x)$ < $\phi(x)$ véges sok x kivételével. Más szavakkal, noha $m(x)$ a végtelenbe tart, sokkal lassabban teszi azt, mint bármelyik nem korlátos parciális függvény.

Bizonyítás:

(ii.) indirekt módon:

m (x) növekedési helyen levő x-nek legyen m (x) a kódja. A legrövidebb program kód, amely x –et előállítja. m (x) = i,

$x_i < x \leq x_{i+1}$., ilyen i indexek, x_i kódok/ programok találhatók

$f(p) = x$, $l(p) \geq i$, ha $x > x_i$

(*a)* miatt ez ellentmondásra vezet, ha $m(x)$ korlátos volna, akkor $C(x) \leq C_f(x) + c_f$ és lim C (x) =$\propto$ miatt)

(iii.) Lassabb kiszámítható fv. létezése, ugyan ehhez az ellentmondáshoz vezet. QED.

Invariancia tétel
Tulajdonságok

i. Bármely olyan végtelenbe tartó, monoton, parciális rekurzív $\phi(x)$ fv.-re , amelyre valamilyen x_0-tól ($\phi(x) \to \infty$) , $m(x) < \phi(x)$ véges sok x kivételével. Más szavakkal, noha $m(x)$ a végtelenbe tart, sokkal lassabban teszi azt, mint bármelyik nem korlátos parciális függvény.

Bizonyítás:

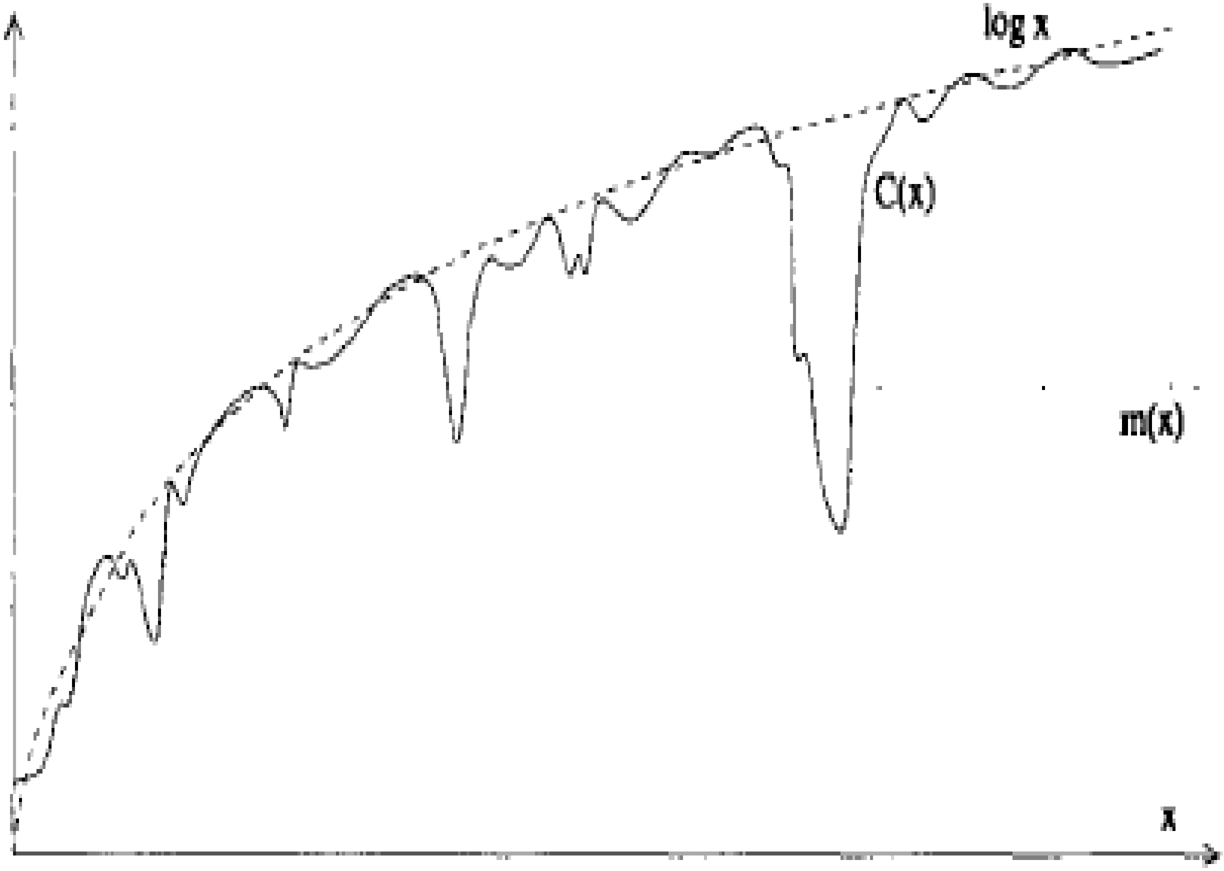
(ii.) indirekt módon:

$m(x)$ növekedési helyen levő x-nek legyen $m(x)$ a kódja. A legrövidebb program kód, amely x –et előállítja. $m(x) = i$, $x_i < x \leq x_{i+1}$., ilyen i indexek, x_i kódok/ programok találhatók $f(p) = x$, $l(p) \geq i$, ha $x > x_i$

(*a*) miatt ez ellentmondásra vezet, ha $m(x)$ korlátos volna, akkor $C(x) \leq C_f(x) + c_f$ és lim C (x) =∞ miatt)

(iii.) Lassabb kiszámítható fv. létezése, ugyan ehhez az ellentmondáshoz vezet. QED.

C(*x*) grafikonja mint fv. az egész számokon



Invariancia tétel
Tulajdonságok

6. Tulajdonság:
   kis változásra mennyit változik? (Folytonosság)
   $C(x+h) \leq C(x) + l(h) + 2\log l(h) + c$,
   $f_0(p)=x$, $l(p)= C(x)$ választással
   $z=h'p$-ből,
   $f(z)= f_0(\pi_{12}(z)) + \pi_{11}(z)= f_0(p)+h= x+h$,
   $C_f(x+h) \leq l(p) + l(h) + 2l(\, l(h)\,) + c$. *QED.*

A $g(x,h)=x+h$ helyett tetszőleges kétváltozós kiszámítható függvényt vehetünk.

Adatbázisokra értelmezve, nagy adatbázis kis módosítása, nem növeli a Kolmogorov bonyolultságot.

Megjegyzés: (a kétváltozós függvény ***kiszámítási idejét*** itt nem veszi figyelembe!)

Feltételes Kolmogorov bonyolultság

Kérdés: Mennyit segít *y* ismerete *x* kiszámításában? Mekkora a legrövidebb program kódjának a hossza, ami *y*-ból *x*-et kiszámítja? *p* kód, f_p: $N \to N$ parciális függvények $f_p(y)=x$.

A programok kódjai $p_1, p_2, \ldots, p_n$.

Legáltalánosabban rekurzívan felsorolható halmaz, egy $h(i)=p_i$ a felsorolás, akkor

$$h(i) = p_i$$

$$f(i,y) \to (h(i))(y) = p_i(y) = f_{p_i}(y) = x$$

$$f(p,y) = \begin{cases} f_{p_i}(y), \text{ha valamilyen } i\text{-re } p = h(i) = p_i \\ \text{nincs értelmezve , ha nincs } i, \text{ amire } h(i) = p. \end{cases}$$

Molnár Bálint, Benczúr András

Feltételes Kolmogorov bonyolultság

Definíció: f_p: $N^2 \rightarrow N$ szerinti feltételes bonyolultsága x-nek az y feltétel ismeretében:

$$C_f(x|y)=\begin{cases} \min_p \{l(p): f(p,y)=x\} \\ \infty, \text{ha nem létezik ilyen } p \end{cases}$$

Tétel : Létezik optimális $f_0^{(2)} : N^2 \rightarrow N$

parciálisa n rekurzív függvény,

amire bármely $f \;\; : N^2 \rightarrow N$

parciálisa n rekurzív függvényhe z létezik

k_f konstans, hogy :

$$C_{f_0}(x|y) \leq C_f(x|y) + k_f \;, \forall x, y \in N - \text{re}$$

Feltételes Kolmogorov bonyolultság

optimális parciálisan rekurzív függvények, feltételes bonyolultsága csak egy konstansban tér el egymástól

$f_0^{(2)}, g_0^{(2)} : N^2 \to N$ optimális

$$\left| C_{f_0^{(2)}}\left(x \middle| y\right) - C_{g_0^{(2)}}\left(x \middle| y\right) \right| \leq K_{f_0^{(2)}, g_0^{(2)}}$$

rögzítjük az $f_0^{(2)}$ optimális függvényt

$x \in N$, feltételes Kolmogorov bonyolultsága

$$C\left(x \middle| y\right) = C_{f_0^{(2)}}\left(x \middle| y\right)$$

Feltételes Kolmogorov bonyolultság

Bizonyítás:

- Ugyanúgy mint az egyszerű (nem feltételes bonyolultság) esetében
- U((n,p), y) kétváltozós univerzális parciális függvények segítségével

Feltételes Kolmogorov bonyolultság

Tulajdonságai:

1) Alaptétel: $C(x|y)$ konkrét $C_f(x|y)$ ismeretében csak felülről becsülhető.
2) $C(x|y)=K$ jelentése: $\exists\ p{:}f_0(p,y)=x$, és $l(p)=K$
3) $C(x) \leq C(x|y)+C(y)+2\log_2 C(y)+K$

Programok és C(x)

LISP (Erlang)-funkcionális, lista kezelő

Fortran

$\lambda_1, \lambda_2, \ldots,$ - Lisp programok

$\pi_1, \pi_2, \ldots..,$ Fortran programok

Felsorolás – minden program tekinthető parciális rekurzív függvény

$C_{LISP}(x)$ és $C_{Fortran}(x)$, kijelölve egy referencia gépet. Ezek egy additív konstans erejéig különböznek.

Mindegyik felsorolás tartalmaz egy *U* univerzális gépet (interpretert)

Programok és C(x)

LISP programok felsorolása tartalmaz egy LISP interpretert, amelyik minden LISP programot végre tud hajtani.

De nyilván létezik λ_p LISP program, amelyik Fortran interpreter, amelyik minden Fortran programot végre tud hajtani.

$C_{\text{LISP}}(x) \leq C_{\text{Fortran}}(x)+2l(\lambda_p)$, hasonlóan

$C_{\text{Fortran}}(x) \leq C_{\text{LISP}}(x)+2l(\pi_L)$

$| C_{\text{LISP}}(x)- C_{\text{Fortran}}(x) | \leq 2l(\lambda_p)+ 2l(\pi_L)$

Feltételes Kolmogorov bonyolultság

4) Halmaz szerinti feltételes bonyolultság (véges halmaz elemeinek a halmaz szerinti feltételes bonyolultsága)
$\mathcal{A}=\{a_1, \ldots, a_n\}$
$C(x|\mathcal{A})=?$, $x\in \mathcal{A}$,

$<\mathcal{A}>=\overline{a}_1, \ldots, \overline{a}_n \Rightarrow C(x|\mathcal{A})=C(x|<\mathcal{A}>)$
[$\mathcal{A}$ kódolása önkorlátozó kóddal]
általános eset:
adott $\mathcal{A}$, $|\mathcal{A}|=N$, kódja p
[Mennyire tömören ábrázolható?]

Feltételes Kolmogorov bonyolultság

4) $|\{x| \; x \in \mathcal{A}$ és $C(x| \mathcal{A}) < log_2 N - \delta\}|$?, keressük : p kódot: $f_0(p, <\mathcal{A}>) = x \in \mathcal{A}$

$0, 1, 2, log_2 N - \delta - 1$ <-[kódhossz]

$1+2+2^2+\ldots+2^{log_2 N - \delta - 1} = 2^{log_2 N - \delta} - 1$

$$\frac{|\{x| \; x \in \mathcal{A} \wedge C(x|\mathcal{A}) < log_2 N\text{-}\delta\}|}{|\{x| \; x \in \mathcal{A}\}|} = \frac{2^{log_2 N\text{-}\delta}}{2^{log_2 N}} \leq 2^{-\delta}$$

Feltételes Kolmogorov bonyolultság

- $\mathcal{A}$ elemeinek többségére az $\mathcal{A}$ szerinti feltételes bonyolultság nem lehet érdemben kisebb mint $log_2 N$
  - $log_2 N$-nél δ-val rövidebb kódok száma, amit f_0 használhat $\leq N\, 2^{-\delta}$
  - egyenletes kódhossz választása indokolt
  - Szükséges a megfelelő halmaz megtalálása
- A halmaz ismeretében sem lehet tömörebben ábrázolni a legtöbb elemet az egyenletes kódolásnál.
- Keresem a lehető a legkisebb halmazt, ami lefedi a jelenséget és egyenletesen kódolom

Példa

Személyek adataira relációs Adatbázis

$\mathcal{R}$(csnév, unév)

csnév: 20 byte = 160 bit

unév: 20 byte = 160 bit

$\mathcal{A}$: az összes lehetséges reláció előfordulás

– Lehetséges különböző sorok száma: $2^{160} \times 2^{160} = 2^{320}$
– Lehetséges előfordulások száma, az összes lehetséges részhalmaz, a hatványhalmaz, aminek a számossága:

Példa

$$|\mathcal{A}| = 2^{2^{320}}$$

$$\{I | I \in \mathcal{A}\}, C(I|\mathcal{A}) = log_2 2^{2^{320}} = 2^{320}$$

Sorok számára van egy korlát a való életben

< 10 milliárd (10^9)

Példa

csnév: első 256 leggyakoribb (80 %)

unév: első 256 leggyakoribb (80 %)

4 szegmens készítése:

1. Mindkét név gyakori: 64% - 2 byte
2. Családnév gyakori : 16% - 22 byte
3. Utónév gyakori : 16% - 22 byte
4. Egyik név sem gyakori : 4% - 4 byte

(Jól kell részhalmazokra szétszedni).

Paraméteres halmazsereg

(pl. adatbázis (reláció) sémának rögzítek egy paraméterét és az előfordulás a halmazsereg).

- *a* paraméter kód,

a -> B_a , véges halmaz

$h(a)= B_a$ – *h* egy totális rekurzív fv. (meg tudom mondani, hogy mi van a halmazban és mi nincs benne).

Paraméteres halmazsereg

Legyen $\mathcal{A} \subseteq N \times N$, rekurzívan felsorolható részhalmaz,

$\mathcal{B}_a = \{x \mid (x,a) \in \mathcal{A}\}$

Paraméteres halmazsereg

Megszorítás: minden a-ra $|\mathcal{B}_a| = m_a$ véges

a alapján $\mathcal{B}_a$ elemei felsorolhatók algoritmikusan

m_a-t nem tudom megmondani(kiszámítani) (mert $\mathcal{A}$ rekurzívan felsorolható, de a komplementer halmazáról nem tudunk semmit, ezért nem eldönthető)

Paraméteres halmazsereg

<u>Tétel</u>: létezik

$\exists\, k_{\mathcal{A}}$, hogy

$$C(x|\mathcal{B}_a) = C(x|a) \leq log_2\, m_a + k_{\mathcal{A}}$$

$\forall x \in \mathcal{B}_a -$ ra, és $\forall \mathcal{A}$ - ra

A halmaz elemeinek a tömöríthetősége mindig eléri a $\log_2$ (elemszám) –ot, alulról felülről is becsülni tudjuk

Paraméteres halmazsereg

Bizonyítás:

$\mathcal{B}_a$ elemei : $\mathcal{A}$-hoz $\exists$ felsoroló függvény

$f(t,a)$ konstruálása

- Futtatjuk $\mathcal{A}$ felsorolását,

Paraméteres halmazsereg

$\Rightarrow$ Első $x \in \mathcal{B}_a \Rightarrow f(0,a) = x_1$

$\Rightarrow$ Megyünk tovább

$\Rightarrow$ $f(1,a) = x_2, \ldots, f(n,a) = x_n$

$\Rightarrow$ $\mathcal{B}_a$ $\forall$ eleme felsorolásra került $\Rightarrow$ $f(t,a)$ ezután nincs értelmezve

$\Rightarrow$

Paraméteres halmazsereg

Az f függvény, $f(t,a)$ az a - ban értelmezve van $t = 0,1,\dots,m_a - 1$ helyeken és másutt sehol.

Legyen $x \in \mathcal{B}_a$, akkor nyilván létezik t, hogy $f(t,a) = x$, és $t \le m_a$,

Ebből

$$C(x|a) \le C_f(x|a) + k_f \le l(t) + k_f \le l(t) + k_f \le \underbrace{log_2 m_a}_{\text{felsőbecslés}} + k_f.\text{QED.}$$

Vagyis :

$$C(x|\mathcal{B}_a) = C(x|a) \le log_2 m_a + k_{\mathcal{A}}$$

Jelentés*:* (adatbázis) séma+paraméter jó megválasztása, Ezen belül már nem érdemes más kódolást használni, csak a homogén hosszúságút*,* azaz $log_2\ m_a$ hosszút*.*

Érdekes halmazsereg

$\mathcal{D}_k = \{x | C(x) \leq k\}$ k paraméterb en ez egy halmazsere g

Meg tudjuk - e konstruáln i a felsoroló fv. - t?

f_0 optimális fv. hely és lépésszám bejárása

$f_0 : 0$

$f_0 : 1$

$\frac{hely}{lépés}$	0	1	2	3
1	1	1	2	3
1	1	2	3	
2	2	3		
3	3			

ha $x \in \mathcal{D}_k$: valamelyik k

hosszú értékre kapunk

$f_0(p) = x$ eredményt

Érdekes halmazsereg

$\mathcal{D}_k^{\star} = \{x \mid C(x) = k\}$ nem kiszámítható halmaz

$$\left(\begin{array}{l}\text{ui. akkor a } C(x) \text{ is kiszámítható volna;} \\ \text{általában nem lehet bebizonyítani, hogy} \\ \text{a Kolmogorov bonyolultsága valaminek } k\end{array}\right)$$

$\mathcal{D}_k' = \{x \mid C(x) < k\}$ ez sem kiszámítható halmaz

Nem tudom megmondani, hogy

nincs-e jobb kód nála

Összehasonlítás a Shannon entrópiával

Minél „nagyobb/jelentősebb” az adott, vizsgált jelenség, annál szorosabb a kapcsolat a két entrópia fogalom között.

Hosszú sorozat , *N* hosszú üzenetek :

H entrópia (Shannon) [stacionáris, ergodikus forrás]

Két halmazba sorolható:

Lényeges, tipikus halmaz: ebbe 1-ε val.szeg.-gel. esik a választás

$2^{N(H-\delta)} < M(N) < 2^{N(H+\delta)}$.

M(N)- tipikus halmaz elemszáma.

Összehasonlítás a Shannon entrópiával

Mi jellemzi a tipikus halmazt?

Kolmogorov: $\overline{x}$ üzenet tipikus $\Rightarrow$ statisztika jellemzi , $\mathcal{A}_N$ halmazba esik, (ez lesz a paraméteres halmaz sereg).

$\mathcal{A}_N$={tipikus N hosszú sorozatok}

$\overline{x} \in \mathcal{A}_N$: $C(\overline{x} | \mathcal{A}_N) = C(\overline{x} | N) \leq \log_2 M(N) + k_{forrás}$

Összehasonlítás a Shannon entrópiával

$$P_{TIP}(\bar{x}) = \begin{cases} 0, \text{ ha } x \in TIP \\ 1, \text{ ha nem } x \notin TIP \end{cases}$$

Kiszámítha tó predikátum

A paraméter halmaz :

$$\mathcal{A} = \left\{ N, \left\{ \bar{x} \middle| l(\bar{x}) = N \right\} \wedge P_{TIP}(\bar{x}) = 0 \right\}$$

$$\mathcal{A}_{\mathcal{N}} = \{ \text{tipikus } N \text{ hosszú sorozatok} \}$$

$$log_2 M(N) = log_2 |\mathcal{A}_{\mathcal{N}}|$$

$$C(\bar{x} | \mathcal{A}_{\mathcal{N}}) \leq log_2 |\mathcal{A}_{\mathcal{N}}| + k_{\mathcal{A}} = log_2 M(N) + k_{\mathcal{A}}$$

Összehasonlítás a Shannon entrópiával

Behelyettesítve tetszőleges

δ -ra az $\mathcal{A}_{\mathcal{N}}(-\delta)$ tipikus elemére

az alsó becslés, a teljesre a felső teljesül:

$$log_2\left|\mathcal{A}_{\mathcal{N}}\right| - \varepsilon = N(H-\delta) - \varepsilon \le C\left(x\middle|\mathcal{A}_{\mathcal{N}}\right) \le N(H+\delta) + k_{forrás}$$

Egy szimbólumra

$$\underbrace{H}_{\text{Shannon}} \sim (H-\delta) - \frac{\varepsilon}{N} \le \frac{\overbrace{C\left(x\middle|\mathcal{A}_{\mathcal{N}}\right)}^{\text{Kolmogorov}}}{N} \le (H+\delta) + \frac{k_{forrás}}{N} \sim \underbrace{H}_{\text{Shannon}}$$

N hosszú üzenetek kódját osztjuk N-nel, Shannon entrópia H

Összehasonlítás a Shannon entrópiával

Egy szimbólumra jutó (Shannon) entrópia a tipikus halmaz elemein a feltételes Kolmogorov entrópia egy szimbólumra jutó részével

Ha nagyon nagy jelenségeket veszünk, akkor a tipikus halmaz Kolmogorov bonyolultsága visszaadja a Shannon entrópiát.

Összehasonlítás a Shannon entrópiával

Mögötte meghúzódó jelenség: a tipikus halmazbeli üzenetek egyenletes eloszlással jellemezhetők.

Összes közül melyek a tipikus halmazbeli üzenetek és melyek nem

0 – 1 sorozatok véletlenszerűségét vizsgáljuk. Lehet-e pénzfeldobás eredménye?

Összehasonlítás a Shannon entrópiával

0 0 1 1 0 1 0 0 . . . 1 . . .

0 0 0 1 0 1 0 0 . . . .

0 0 0 0 0 1 0 0 . . . .

1000 pénzfeldobás

101 – véletlenszerű

Egyeseket *(1)* egyesével 0-ra cserélem:

0000….0….0000 nem véletlen

Van-e határvonal a véletlen és a nem véletlen között?

Összehasonlítás a Shannon entrópiával

Szabályszerűség

Ami igazán véletlen az nem tömöríthető

(Kezdőszelet bonyolultsága a hosszal nő)

Mennyire tipikus

Véletlenség defektusa jellemezhető: δ

$\{x \mid x \in \mathcal{A}$ és $C(x \mid \mathcal{A}) < \log_2 |\mathcal{A}| - \delta\}$

A véletlen defektusa legalább δ

Példa

Fekete – fehér színezések

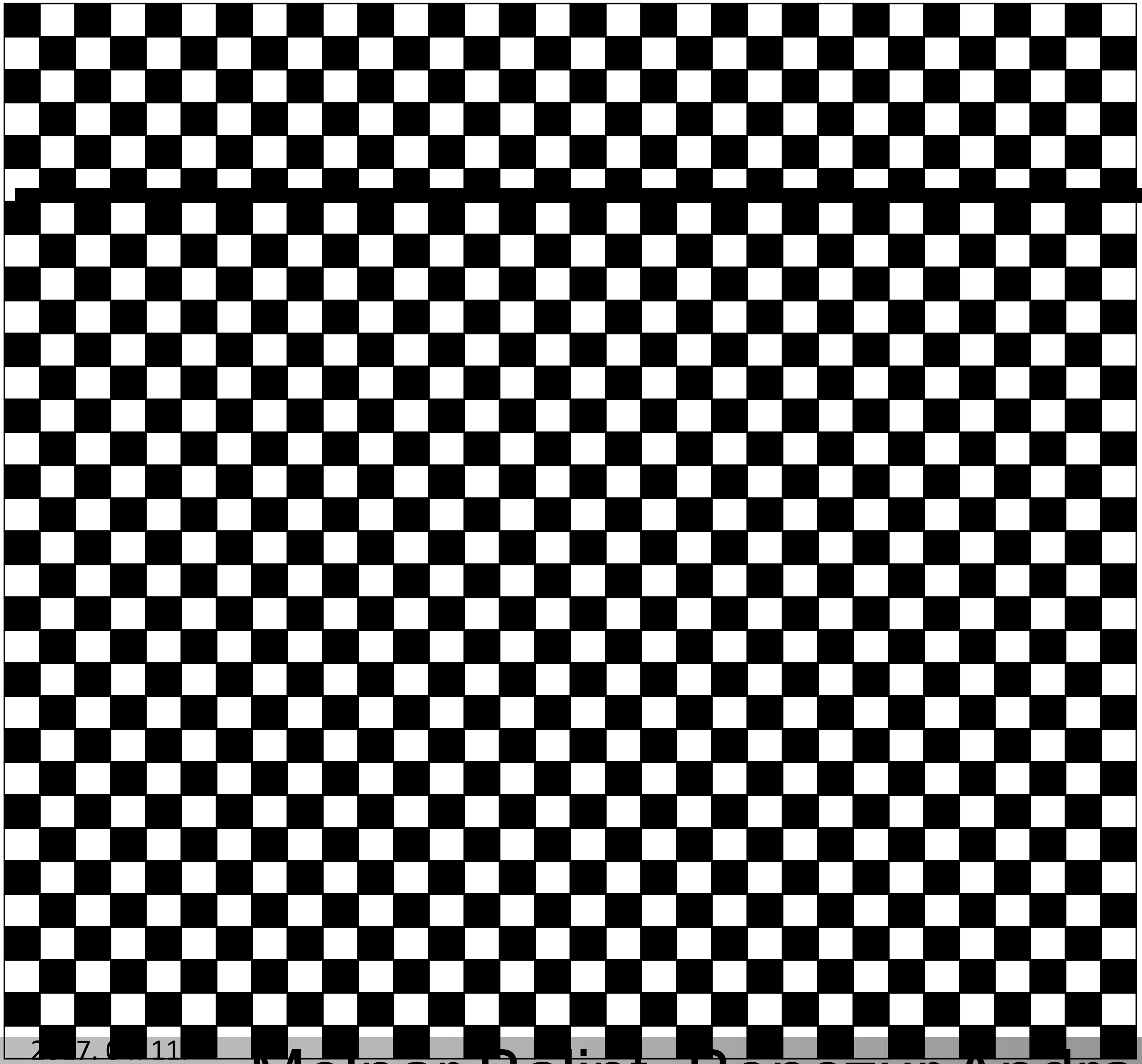
Sakktábla

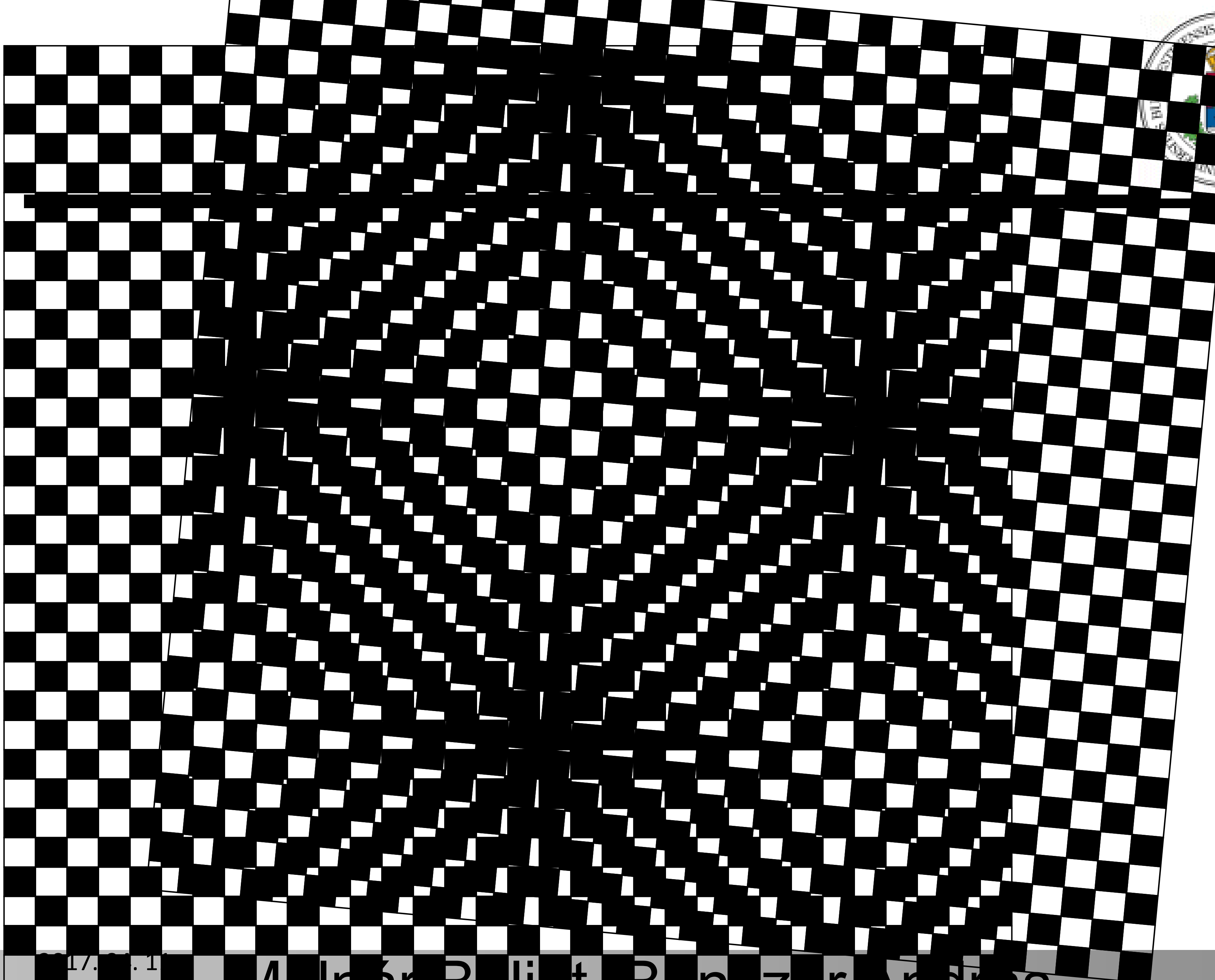
Homogén illetve mákos sűrűség

Fele fekete- fele fehér

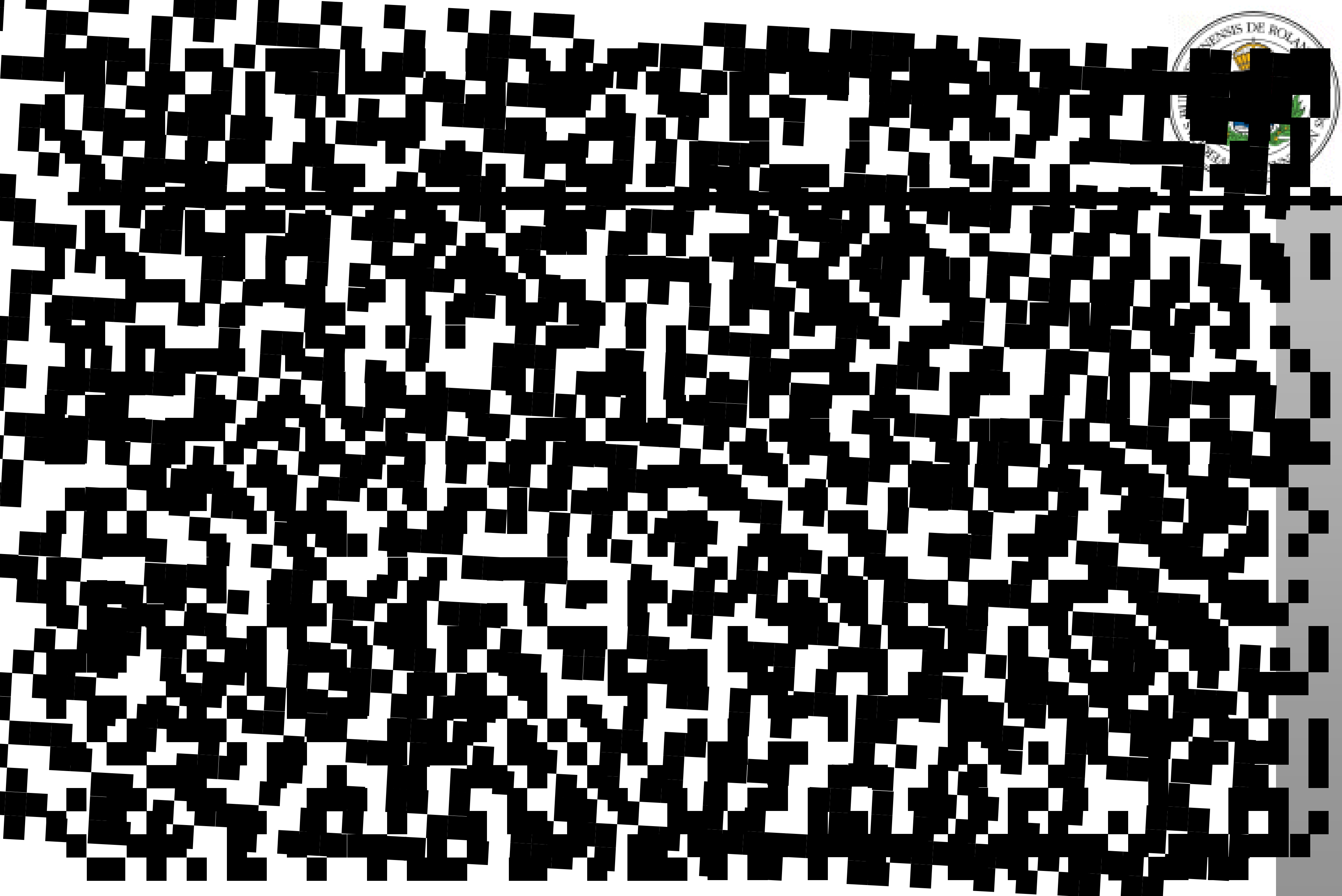
Egymásra helyezve a fóliát érdekes jelenség

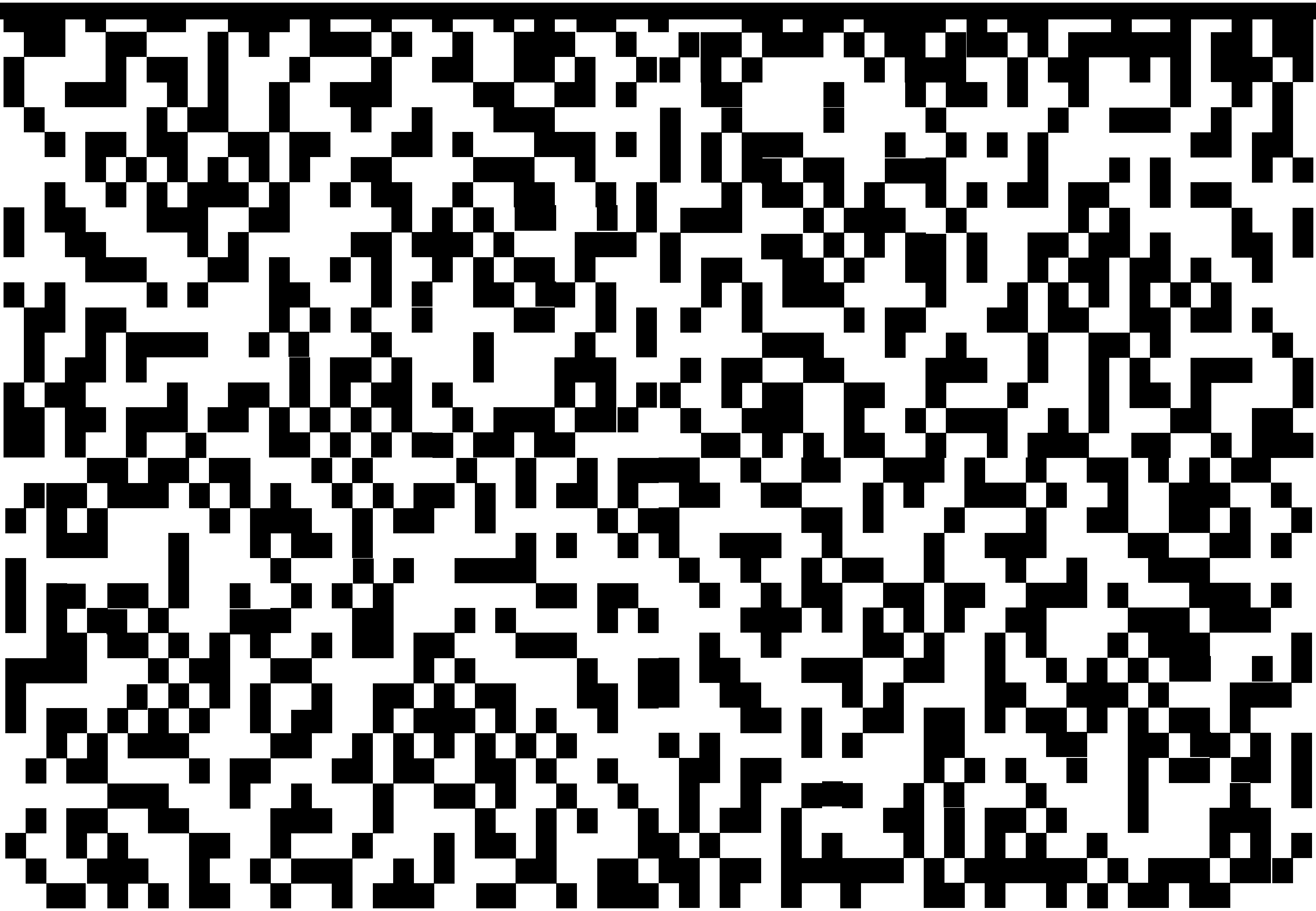
Saját görbére illeszkedő transzformáció felerősödik…





Molnár Bálint, Benczúr András





A jellemző szabályosságok írják le a tipikus halmazt, elemein a legbizonytalanabb, egyenletes eloszlást vesszük. Ha megtaláltuk a jó homogén halmazt, lehet, hogy a felhasználás, üzenet szempontjából a konkrét elem már érdektelen.
Példa: szürke színezés. Az 1/2 szürkeség - Fuzzy halmaz - bármelyik tipikus színezéssel szemléltethető.
Két színezés és sajátosságaik: a szórásnégyzet második tagja

A véges világban - tehát a digitalizált világban - a matematika végtelen ideális konstrukciói - a folytonosság, végtelen, univerzális algoritmusok - csak segítenek, végtelennel közelítik a véges modellt, amit utána vissza kell véges közelítésbe hozni.

$$np(2-p)(1-p)^2 + 2(n-k)p(1-p)^3$$

Matematikai kitérő –

$$C_f(x) = \min\{l(p): f(p) = x\}$$

$$\infty \text{ ha nincs } p, \text{ amire } f(p) = x$$

$$\text{Megjegyzés}: C_f(x) \text{ nem mindig kiszámítható}$$

A kibertérben mivel azonosíthatók

p -?

f -?

$C_f(x)$

$U(.,\ .)$

Intuitív kitérő –

- Egy bizonyos objektumot egy véges bináris stringgel kívánunk leírni
- Több leírás is létezhet, de egy leírás egy objektumot ír le
- „Egyszerű” objektum rövid leírás
- „Bonyolult” objektum hosszú leírás
- Richard-Berry paradoxon:
  - Definiáljunk egy természetes számot: „a legkisebb olyan természetes szám, amelyet nem lehet leírni <u>húsznál</u> kevesebb szóval”.

Intuitív kitérő –

– Ha ilyen szám létezik, akkor ezt a számot tizenkét szóval leírtuk, ellentmondva a definíciójának.
– Ha ilyen szám nem létezik, akkor az összes természetes számot le lehetne írni húsznál kevesebb szóval.
– A „leírás” fogalmát pontosítani kell !

D: $Y \rightarrow X$, $\boldsymbol{D(y)=x}$, *Y* leírások, *X* az objektumok,

x legyen véges mint leírás, $\Rightarrow |X|=|Y|=\omega$, $\aleph_0$(megszámlálható).

- $y \in \{0,1\}^*$, üzenet továbbítás költsége távközlésben az üzenet hossz $l(y)$.
- Legkisebb költség $\rightarrow$ legkisebb hossz

Intuitív kitérő –

– Algoritmikus információ elmélet vívmányai (innovációi)

1. Formális értelemben effektív *leírás* definíciójára szorítkozva, ez a definíció lefedi mind az intuitív mind a matematika és logikai szempontjából elfogadható nézőpontokat.
2. Erre az effektív definícióra szorítkozás maga után vonja, hogy létezik egy olyan univerzális *leíró* módszer, amely minorálja az összes egyéb leíró módszerből származtatható leírás hosszát vagy komplexitását.
3. Egy objektum leírásának hossza az objektum elválaszthatatlan tulajdonsága, amely független a partikuláris leíró vagy formalizáló módszertől.
4. A zavaró Richard–Berry paradoxon nem fog eltűnni, hanem Gödel nem-teljességi tételével kapcsolatban elő fog kerülni, vagyis nem minden igaz matematikai állítás bizonyítható matematikában.