(1)

Információelmélet

Shannon, C. E.

"A Mathematical Theory of Communication"

Magyar fordítás:

Claude E. Shannon – Warren Weaver:

"A kommunikáció matematikai elmélete (az információelmélet születése és távlatai)"

OMIKK, Bp. 1986.

Javasolt irodalom:

Györfi László – Győri Sándor – Vajda István:

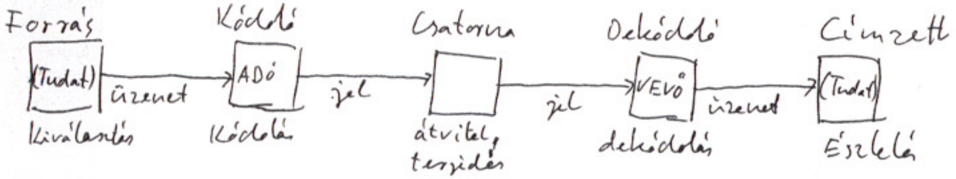
"Információ és kódelmélet,

Bp. Typotex Kiadó, 2000

Bevezetés: ullman:/

Shannon kommunikációs modellje

"A kommunikáció felöleli mindazokat az eljárásokat, amelyeken keresztül az egyik emberi elme a másikra hatni képes". (W. Weaver)



Szintek:

A) mennyiségi szint:

mennyi vihető át rekonstruálhatóan a hibamentesített üzenet az adott csatornán?

(mennyiségi jellemzők: entrópia és csatorna kapacitás)

B.) megértési szint (jelentés, szemantika)

C.) hatékonysági szint (a kívánt hatás elérése)

Diszkrét forrás, diszkrét, zajmentes csatorna

PÉLDA:

Legegyszerűbb csatornával kezdjük (logikai, formális lehet

Bináris, szimmetrikus csatorna,

jelei: 0 és 1,

átviteli idő: mindkettőre 1/C. (Hasonlóak digitális csatornák.

Egy Forrás: igen-nem üzeneteket választ.

a.) Szabályos pénzérme feldobása C-szer, időegység alatt

2^C egyforma kimenetelű,

Kódolás: igen → 0
nem → 1

2^C különböző kódszó, egységnyi idő hosszú,

Megoldja az átvitelt.

$T = 2^k$ idő alatt kezelhető üzenetekre működik.

b.) Lottó húzás,

igen: 5 találat van a szelvényemen
nem: nincs ötös a szelvényemen.

$1024 = 2^{10}$ ~~idő~~ hét eredménysorozatát milyen rövid idő ~~alatt~~ alatt tudjuk csatornánk segítségével a címzetthez eljuttatni?

valószínűség

3 eset: Nincs ötös egyszer se: kódszó: 1 $P_0 = (1 - P(\text{ötös}))$

Egy ötös van, a k-adik: 00 $[\log_2 k]$ $P_1 = 2^{10} P(\text{ötös}) \cdot (1 - P(\text{ötös}))$

~~Tö~~ Egynél több ötös van: 01 [1 K bit] / a teljes sorozat $P_2 = 1 - P_1 - P_2$

Átviteli idő várható értéke: <

$$\left(P_0 + P_1(2+10) + P_2(2+1024)\right)C^{-1} \text{ sec} \leq$$

$= 2$

[(Poisson-közelítés: binomiálisban $np=\lambda$-val

$$\lambda = 1024 \cdot \binom{90}{5}^{-1} = 2,33 \text{ E-5}$$

$P_0 = 0,999977$ $\quad (P_0 = e^{-\lambda})$

$P_1 = 2,33 \text{ E-5}$ $\quad (P_1 = \lambda e^{-\lambda})$

$P_2 =$ számítási hibánál kisebb]

PÉLDA vége.

Mennyiségek:

1.) Adott csatornához a kapacitás – a bináris szimmetrikus csatornával való ekvivalencia alapján:

Egységnyi idő alatt „átlag" ugyanannyi különböző jelsorozat vihető át, mint egy C bit/sec sebességű bináris csatornán.

Ekkor a kapacitás C.

2.) A forrás szimbólumok sorozatát választja, a választás bizonytalanságát jellemezzük mennyiséggel; ez lesz a Shannon-entrópia.

N hosszú sorozat, $p_1, p_2, \dots, p_{n^N}$ a választások valószínűsége.
n lehetséges szimbólum: $\sum_{i=1}^{n^N} p_i = 1$.

Shannon entrópia-formula (levezetése következik majd)

$$H(p_1, \dots, p_m) = -\sum_{j=1}^{m} p_j \log_2 p_j .$$

A forrás (N hosszú) egy szimbólumra jutó bizonytalansága

$$H = \frac{1}{N} \sum_{i=1}^{n^N} p_i \log_2 p_i^{-1} .$$

A forrás sebessége legyen V szimbólum/sec.

Shannon csatorna-kódolási alaptétele: C, H, V-re

Nem lehet veszteségmentesen működtetni a kommunikációs (hírközlési) rendszert, ha

$$V > \frac{C}{H} ;$$

tetszőleges $\varepsilon > 0$-hoz létezik kódolás, hogy

$$V < \frac{C}{H} - \varepsilon$$

sebességgel működtetve a forrást, minden üzenet veszteségmentesen átvihető.

Ezt bizonyítjuk majd formálisan a legelemibb esetre.

A csatorna kapacitása

Diszkrét, zajmentes csatorna

$s_1, s_2, \dots, s_m$ jelek,

$t_1, t_2, \dots, t_m$ – átviteli idő jelenként.

szimmetrikus csatorna: $t_1 = t_2 = \dots = t_m = t$

Szimmetrikus bináris csatorna: $m = 2$. (jelei: bitek <u>bi</u>nary digi<u>t</u>)

Kapacitása: másodpercenként átvihető bitek száma.

Nyilván $C = t^{-1}$, ha t^{-1} egész szám.

Ha nem, és nem összemérhető t és 1, akkor T idő alatt $N(T)$ jelsorozatunk lehet, és

$$2^{T(t^{-1}-\varepsilon)} < N(T) < 2^{T(t^{-1}+\varepsilon)}$$

teljesül tetszőleges ε-ra, ha T elég nagy.

A logaritmusokat (a továbbiakban 2-es alapú) véve és T-vel osztva:

$$t^{-1} - \varepsilon < \frac{1}{T} \log_2 N(T) < t^{-1} + \varepsilon ,$$

azaz a határérték $c = t^{-1}$ lesz.

Intuitíve:
Két csatorna ekvivalens, ha T hosszú sorozataik száma minden (határértékben) T-re megegyezik. (A T hosszú kódszavak kölcsönösen egyértelműen megfeleltethetők egymásnak.)

Pontos definíció:
Legyen $N(T)$ egy csatorna T hosszú (T-nél nem hosszabb, maximális hosszú) jelsorozatainak száma.
A csatorna kapacitása:

$$C = \lim_{T \to \infty} \frac{\log_2 N(T)}{T} .$$

Következmény: minden C kapacitású csatorna ekvivalens a C kapacitású bináris ~~bo~~ csatornával.
Ezért C dimenziója: bit/sec.

Példa:
1.) Szimmetrikus csatorna, m szimbólum, t egy szimbólum átviteli ideje:

$$N(T) \cong m^{T/t} ,$$

$$C = \lim_{T \to \infty} \frac{\log_2 m^{T/t}}{T} = t^{-1} \log_2 m$$

2.) Egyszerű, nem szimmetrikus csatorna: $t_1, t_2, \ldots, t_m$ átviteli idők, nincs megszorítás a jelsorozatra

$$N(T) = N(T - t_1) + N(T - t_2) + \ldots + N(T - t_m)$$

Karakterisztikus polinom: $1 - x^{-t_1} - x^{-t_2} \ldots - x^{-t_m} = 0$, x_0: legnagyobb megoldása:

$N(T) = K x_0^T$ a megoldás. (kis gyökök elhanyagolható adalékot adnak)

$c = \log x_0$

A forrás bizonytalanságának mérőszáma: a Shannon-entrópia

A hírközlő rendszer minden lehetséges üzenet-választásra (sokszor ismétlődő választásra) működjön. Nem az a fontos az átvitelben, hogy mit választunk, hanem az, hogy milyen bizonytalan a választás. ~~(Az átvitel invariáns a szimbólumok választásával~~

(A szimbólumokat a forrás „átnevezheti", ettől független a hírközlés mennyiségi viselkedése.)

A forrás N hosszú választásait n szimbólumból n^N elemű valószínűségeloszlás jellemzi. Ennek meghatározása empirikus alapon, közelítőleg történhet, majd nézzük később az alapeseteket.

Feladatunk: ~~adott~~
Keressük a diszkrét v. eloszlások halmazán a $H(p_1, p_2, \ldots, p_n)$ függvényt, ahol $n \geq 1$, egész, $\sum_{i=1}^{n} p_i = 1$, $p_i \neq 0$, $i = 1, \ldots, n$,

Természetes elvárások:

0. H nem függ p_i-k permutációjától; (Ez jelenti azt, hogy eloszlást nézünk, és nem a p_i-k sorozatát!)

Shannon axiómái:

1. Folytonosság: H minden argumentumában legyen folytonos.
2. Az $A(n) = H(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$ jelöléssel
$A(n) < A(m)$, ha $n < m$.
3. Elágazási szabály; A megfigyelés lépcsőzése:
$H(p_1, \ldots, p_{n-2}, p_{n-1}, p_n) = H(p_1, \ldots, p_{n-2}, p_{n-1} + p_n) + (p_{n-1} + p_n) H\left(\frac{p_{n-1}}{p_{n-1} + p_n}, \ldots\right.$

3°. Az elágazási szabály általános alakja: (következik 3-ból)

Legyen $q_j = \sum_{i=1}^{n_j} p_{j,i}$, és $\sum_{j=1}^{m} q_j = 1$. Akkor

$$H(p_{1,1},\dots,p_{1,n_1},p_{2,1},\dots,p_{2,n_2},\dots,p_{m,1},\dots,p_{m,n_m}) =$$
$$= H(q_1,q_2,\dots,q_m) + \sum_{j=1}^{m} q_j H\left(\frac{p_{j,1}}{q_j},\dots,\frac{p_{j,n_j}}{q_j}\right).$$

1. Tétel: Az 1 – 3 axiómákat kielégítő függvények csak a

$$K\sum_{i=1}^{n} p_i \log_2 p_i$$

alakú függvények lehetnek, ahol $K < 0$ konstans.

Bizonyítás:

(i) Racionális $p_i = \frac{q_i}{m}$-re, $i=1,\dots,n$, és $\sum_{i=1}^{n} q_i = m$ esetén fejezzük ki az elágazási szabály segítségével $H(p_1,\dots,p_n)$-et $A(m)$ és $A(q_i)$-k segítségével.

$$H(p_1,\dots,p_n) = H\left(\frac{1}{m},\frac{1}{m},\dots,\frac{1}{m}\right) - \sum_{i=1}^{n} p_i H\left(\frac{1}{q_i},\dots,\frac{1}{q_i}\right) =$$
$$= A(m) - \sum_{i=1}^{n} p_i A(q_i) = -\sum_{i=1}^{n} p_i \left(A(q_i) - A(m)\right).$$

(ii) Lemma: $A(n) = c\log_2 n$, ahol $c > 0$, tetszőleges.

Az elágazási szabállyal

$$A(nm) = H(\underbrace{\tfrac{1}{nm},\dots,\tfrac{1}{nm}}_{m},\dots,\underbrace{\tfrac{1}{nm},\dots,\tfrac{1}{nm}}_{m}) =$$
$$= H\left(\frac{1}{n},\dots,\frac{1}{n}\right) + n\cdot\frac{1}{n} H\left(\frac{1}{m},\dots,\frac{1}{m}\right) = A(n) + A(m).$$

Ebből $A(n^m) = m A(n)$ adódik. (1)

Legyen $s > 0$ rögzített egész szám. Tetszőleges t-re és n-re létezik m, hogy

(2) $$s^m < t^n \le s^{m+1}.$$

Kihasználva az A függvény monotonitását,

$$A(s^m) < A(t^n) \le A(s^{m+1}).$$

Felhasználva (1)-et:

$$mA(s) < nA(t) \le (m+1)A(s),$$ ahonnan a

(3) $$\frac{m}{n} < \frac{A(t)}{A(s)} \le \frac{m+1}{n},$$ minden n-re.

Hasonlóan kapjuk (2)-ből a $\log_2$ függvénnyel

(4) $$\frac{m}{n} < \frac{\log_2 t}{\log_2 s} \le \frac{m+1}{n}.$$

A (3) és (4) egyenlőtlenségekből $n \to \infty$ során

$$\frac{A(t)}{A(s)} = \frac{\log_2 t}{\log_2 s}$$ adódik, tehát

$$c = \frac{A(s)}{\log_2 s}$$ választással

$$A(t) = c \log_2 t$$

bizonyítja a lemma állítását.

(iii) A (ii) lemma és a racionális elemű eloszlásokra (i) alatt kapott kifejtés együttesen

$$H(p_1,\dots,p_n) = -\sum_{i=1}^{n} p_i\left(A(q_i) - A(m)\right) = -c\sum_{i=1}^{n} p_i \log_2\frac{q_i}{m} =$$
$$= -c\sum_{i=1}^{n} p_i \log_2 p_i.$$

Ezzel a tételt racionális eloszlásokra igazoltuk. Tetszőleges eloszlásra a folytonosságból és $\log_2$ folytonosságából következik a tétel.

A K konstans megválasztása a $H(\frac{1}{2},\frac{1}{2})$ egységként való választásából

$$K\left(\frac{1}{2}\log_2\frac{1}{2}+\frac{1}{2}\log_2\frac{1}{2}\right)=K(-1)=1$$

miatt $K=-1$.

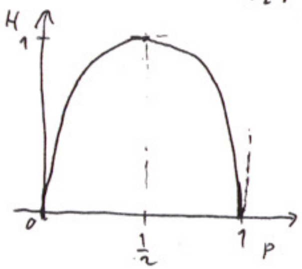
Def Tehát a $(p_1,\dots,p_n)$, $p_i>0$, $\sum_{i=1}^n p_i=1$ eloszlás bizonytalansága, a Shannon-entrópia:

$$H(p_1,\dots,p_n)=-\sum_{i=1}^n p_i\log_2 p_i.\ \square$$

Tulajdonságok:

1. Az n elemű halmazon egyenletes eloszlás entrópiája: $H(\frac{1}{n},\dots,\frac{1}{n})=A(n)=\log_2 n$.
2. A két elemű eloszlások $(p, 1-p)$ jellemzése: $H(p,1-p)=-p\log_2 p-(1-p)\log_2(1-p)$,

alakja:



$H(p,1-p)\le H(\frac{1}{2},\frac{1}{2})=1$, egyenlőség csak $p=\frac{1}{2}$-re.
$H(p,1-p)$ szigorúan monoton növekvő $[0,\frac{1}{2}]$-en, és csökkenő $[\frac{1}{2},1]$-en.

3. $H(p_1,\dots,p_n)\le H(\frac{1}{n},\dots,\frac{1}{n})=A(n)=\log_2 n$.
Egyenlőség csak $p_i=\frac{1}{n}$, $i=1,\dots,n$ teljesül.
Biz: Tegyük fel, van két nem egyenlő, $p_{n-1}\ne p_n$. Az elágazási szabállyal és a 2. tulajdonság felhasználásával:

$$H(p_1,\dots,p_{n-1},p_n)=H(p_1,\dots,p_{n-2},p_{n-1}+p_n)+(p_{n-1}+p_n)H\left(\frac{p_{n-1}}{p_{n-1}+p_n},\frac{p_n}{p_{n-1}+p_n}\right)<$$

$$<H(p_1,\dots,p_{n-2},p_{n-1}+p_n)+(p_{n-1}+p_n)H\left(\frac{1}{2},\frac{1}{2}\right).$$

Az eloszlás kiegyenlítése felé változás növeli az entrópiát.

4. Valószínűségi változók entrópiája

Def: eloszlásuk entrópiája.

ξ, η valószínűségi változók.

$P(\xi=x_i)$, $i=1,\dots,n$ a ξ eloszlása. (összeg = 1)

$P(\eta=y_j)$, $j=1,\dots,m$ az η eloszlása. — " —

$P(\xi=x_i,\eta=y_j)$, $i=1,\dots,n$; $j=1,\dots,m$ a — " — (ξ,η) együttes eloszlása.

A teljes valószínűség tétele alapján

$$P(\xi=x_i)=\sum_{j=1}^m P(\xi=x_i,\eta=y_j),\quad i=1,\dots,n.$$

és

$$P(\eta=y_j)=\sum_{i=1}^n P(\xi=x_i,\eta=y_j).$$

A feltételes valószínűség definíciója szerint

$$P(\xi=x_i,\eta=y_j)=P(\eta=y_j)\,P(\xi=x_i\mid\eta=y_j).$$

A ξ feltételes eloszlása $\eta=y_j$ feltétel szerint

$$P(\xi=x_i\mid\eta=y_j)=\frac{P(\xi=x_i,\eta=y_j)}{P(\eta=y_j)}=\frac{P(\xi=x_i,\eta=y_j)}{\sum_{i=1}^n P(\xi=x_i,\eta=y_j)}$$

Ennek megfelelően az egyes eloszlások entrópiáját véve:

$$H(\xi)=-\sum_{i=1}^n P(\xi=x_i)\log_2 P(\xi=x_i),$$

$$H(\xi,\eta)=\sum_{i=1}^n\sum_{j=1}^m P(\xi=x_i,\eta=y_j)\log_2 P(\xi=x_i,\eta=y_j),\ \text{és}$$

$$H(\xi\mid\eta=y_j)=-\sum_{j=1}^n P(\xi=x_i\mid\eta=y_j)\log_2 P(\xi=x_i\mid\eta=y_j),$$

$j=1,2,\dots,m$

Független valószínűségi változók együttes entrópiája az egyes változók entrópiájának összege:

Biz: Ha ξ, η függetlenek, akkor

$$P(\xi=x_i,\ \eta=y_j)=P(\xi=x_i)P(\eta=y_j),\quad i=1,\dots,n,\ j=1,\dots,m.$$

$$H(\xi,\eta)=-\sum_{i=1}^{n}\sum_{j=1}^{m}P(\xi=x_i)P(\eta=y_j)\left[\log_2 P(\xi=x_i)+\log_2 P(\eta=y_j)\right]$$

$$=-\sum_{i=1}^{n}P(\xi=x_i)\sum_{j=1}^{m}P(\eta=y_j)\log_2 P(\eta=y_j)$$

$$-\sum_{j=1}^{m}P(\eta=y_j)\sum_{i=1}^{n}P(\xi=x_i)\log_2 P(\xi=x_i)=$$

$$=H(\eta)+H(\xi).$$

Az egyenlőség a függetlenség szükséges és elégséges feltétele:

2. <u>Tétel</u>. $H(\xi,\eta)\le H(\xi)+H(\eta)$, és az egyenlőség akkor és csak akkor teljesül, ha ξ és η függetlenek.

/Az entrópiák additivitása a függetlenség karakterisztikus tulajdonsága./

Bizonyítás a feltételes entrópia egy általánosabb tulajdonságából következik.

<u>A feltételes entrópia</u>: $H(\xi|\eta)$.

Kérdezhetjük, mennyi bizonytalanság marad átlagosan ξ-re nézve, ha lehetőségünk lesz η előzetes megfigyelésére.

A teljes bizonytalanságból, $H(\xi,\eta)$-ból ezek szerint $H(\eta)$ mennyiséget tisztázunk átlagosan ξ megfigyelése előtt, tehát marad

$$H(\xi,\eta)-H(\eta).$$

Másként gondolkodva, ha az elágazási szabály alapján – η megfigyeléséig víve az elágazást – az elágazás után megmaradó részt tekintjük ξ-nek η megfigyelése után megmaradó átlagos bizonytalanságának.

Az elágazási szabály szerint az (5) formulát víve

$$H(\xi,\eta)=H(\eta)+\sum_{j=1}^{m}P(\eta=y_j)H(\xi|\eta=y_j).$$

Tehát mindkét úton ugyanazt kapjuk:

Def. A ξ valószínűségi változó η szerinti feltételes entrópiája, jelölése $H(\xi|\eta)$:

$$H(\xi|\eta)=H(\xi,\eta)-H(\eta)=\sum_{j=1}^{m}P(\eta=y_j)H(\xi|\eta=y_j).$$

Kifejtve (5) szerint $H(\xi|\eta=y_j)$-t, (~~és~~ felhasználva $P(\xi=x_i|\eta=y_j)=P(\xi=x_i,\eta=y_j)/P(\eta=y_j)$)

$$H(\xi|\eta)=-\sum_{j=1}^{m}P(\eta=y_j)\sum_{i=1}^{n}\frac{P(\xi=x_i,\ \eta=y_j)}{P(\eta=y_j)}\log_2\frac{P(\xi=x_i,\ \eta=y_j)}{P(\eta=y_j)}=$$

$$= H(\xi|\eta) = +\sum_{j=1}^{m}\sum_{i=1}^{n} P(\xi=x_i,\eta=y_j)\log_2 \frac{P(\eta=y_j)}{P(\xi=x_i,\eta=y_j)}$$

/A fenti (6) formulát használjuk majd a következő lemma bizonyítása során./

A 2. Tétel átírása a feltételes entrópiára nézve (~~amin~~ az egyenlőtlenség mindkét oldalából $H(\eta)$ levonásával:

$$H(\xi,\eta)\le H(\xi)+H(\eta) \iff H(\xi|\eta)\le H(\xi),$$

Kapjuk a

2'. Tétel. $H(\xi|\eta)\le H(\xi)$, és az egyenlőség akkor és csak akkor igaz, ha ξ és η függetlenek.

A 2'. Tételnél általánosabb egyenlőtlenséget bizonyítunk:

3. Tétel.
$H(\xi|\eta)\le H(\xi|f(\eta))$, és az egyenlőség akkor és csak akkor teljesül, ha minden x, y, z-re, ha $f(y)=z$ fennáll

$$P(\xi=x|\eta=y)=P(\xi=x|f(\eta)=z).$$

Bizonyítás. (3. Tétel)

Szükségünk van $f(\eta)$ és $(\xi, f(\eta))$ eloszlására:

$$P(f(\eta)=z)=\sum_{f(y)=z}P(\eta=y),$$

$$P(\xi=x, f(\eta)=z)=\sum_{f(y)=z}P(\xi=x,\eta=y)$$

VEZETŐ INFORMATIKUSOK SZÖVETSÉGE

Írjuk fel $H(\xi|f(\eta))$-t (6) szerint:

$$H(\xi|f(\eta))=\sum_{x}\sum_{z} P(\xi=x,f(\eta)=z)\log_2\frac{P(f(\eta)=z)}{P(\xi=x,f(\eta)=z)}=$$

$$=\sum_{x}\sum_{z}\left(\sum_{f(y)=z}P(\xi=x,\eta=y)\right)\left(\log_2\frac{\sum_{f(y)=z}P(\eta=y)}{\sum_{f(y)=z}P(\xi=x,\eta=y)}\right)$$

Hasonlóan $H(\xi|\eta)$ felírását végezzük $x, z, f(y)=z$ bontásban:

$$H(\xi|\eta)=\sum_{x}\sum_{z}\sum_{f(y)=z}P(\xi=x,\eta=y)\log_2\frac{P(\eta=y)}{P(\xi=x,\eta=y)}.$$

Elegendő az x, z összegzésen belüli tagokat összehasonlítani. Az összehasonlítandó formulák alakja: $a_1,\dots,a_n>0$, $b_1,\dots,b_n>0$, $a=\sum_{i=1}^n a_i$, $b=\sum_{i=1}^n b_i$

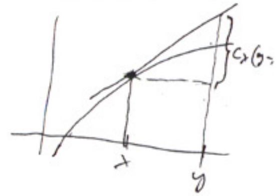
$$\sum_{i=1}^n a_i\log_2\frac{b_i}{a_i} \quad \text{és} \quad a\log_2\frac{b}{a}.$$

/ a_i – $P(\xi=x,\eta=y)$, b_i – $P(\eta=y)$, és $a=\sum_{f(y)=z}P(\xi=x,\eta=y)$, $b=\sum_{f(y)=z}P(\eta=y)$

Lemma:
$$\sum_{i=1}^n a_i\log_2\frac{b_i}{a_i}\le a\log_2\frac{b}{a},$$
egyenlőség csak $\frac{b_i}{a_i}=\frac{b}{a}$ esetén áll fenn.

Biz: a $\log_2 x$ konvexitása miatt

$\log_2 y\le\log_2 x + C_x(y-x)$, és egyenlőség csak $x=y$ esetén.



Ebből $\log_2\frac{b_i}{a_i}\le\log_2\frac{b}{a}+C_{\frac{b}{a}}\left(\frac{b_i}{a_i}-\frac{b}{a}\right)$, = csak $\frac{b_i}{a_i}=\frac{b}{a}$-ra

szorozva a_i-vel és összegezve:

$$\sum_{i=1}^n a_i\log_2\frac{b_i}{a_i}\le\sum_{i=1}^n a_i\log_2\frac{b}{a} + C_{\frac{b}{a}}\left[\sum_{i=1}^n a_i\left(\frac{b_i}{a_i}-\frac{b}{a}\right)\right]=$$

$$=a\log_2\frac{b}{a} + C_{\frac{b}{a}}\left(\underbrace{\sum_{i=1}^n b_i}_{b} - \underbrace{\sum_{i=1}^n a_i\frac{b}{a}}_{b}\right)=a\log_2\frac{b}{a}.$$

A lemma miatt

$H(\xi|\eta) \le H(\xi|f(\eta))$, és

az egyenlőség csak akkor, ha minden z-re teljesül a lemma, azaz $f(y)=z$-n

$$\frac{P(\eta=y)}{P(\xi=x,\eta=y)} = \frac{\sum_{f(y)=z} P(\eta=y)}{\sum_{f(y)=z} P(\xi=x,\eta=y)}.$$

A 3. Tétel következménye: a 2. Tétel:

Az $f(y)\equiv c$ függvényre

$$H(\xi|\eta) \le H(\xi|f(\eta)) \overbrace{= H(\xi|f(\eta)=c)}^{P(f(\eta)=c)} = H(\xi)$$

Az egyenlőség ~~Ez~~ szüks. és elégs. feltétele:

$$\frac{P(\xi=x,\eta=y)}{P(\eta=y)} = \frac{P(\xi=x, f(\eta)=c)}{\underbrace{P(f(\eta)=c)}_{1}} = P(\xi=x).$$

Tehát függetlenek.

2'. Tétel:

$H(\xi_3|\xi_1,\xi_2) \le H(\xi_3|\xi_2)$, egyenlőség a.cs.a., ha $P(\xi_3=x_3|\xi_1=x_1,\xi_2=x_2) =$

$= P(\xi_3=x_3|\xi_2=x_2)$, minden x_1, x_2, x_3-ra.

Biz: $f(x_1,x_2)=x_2$ választással

$H(\xi_3|\xi_1,\xi_2) \le H(\xi_3|f(\xi_1,\xi_2)) = H(\xi_3|\xi_2)$,

egyenlőség a.cs.a., ha

$P(\xi_3=x_3|\xi_1=x_1,\xi_2=x_2) = P(\xi_3=x_3|f(\xi_1,\xi_2)=x_2)$

$= P(\xi_3=x_3|\xi_2=x_2)$.

A lemma következményei

$$\sum a_i \log_2 \frac{b_i}{a_i} \le a \log_2 \frac{b}{a}$$

1.) $(p_1,\dots,p_n)$, $(q_1,\dots,q_n)$

$$\sum p_i \log \frac{q_i}{p_i} \le 1\cdot \log_2 1 = 0$$

$$\left(-\sum p_i \log_2 p_i\right) - \left(-\sum p_i \log_2 q_i\right) \le 0$$

$$* \quad H(p_1,\dots,p_n) \le -\sum p_i \log q_i$$

2.) $b_i = 1$ $b = n$ $a_i = p_i$

$$\sum p_i \log p_i^{-1} \le 1 \log_2 \frac{n}{1} = \log_2 n$$

Jensen-egyenlőtlenség:

$h(x)$ konvex függvény,

$$h(M(\xi)) \le M(h(\xi)),$$

$h(x)$ szigorúan konvex, akkor az = teljesül: $P(\xi = M(\xi)) = 1$.

Általában:

$$H(\xi_t = x_t \mid \xi_1 = x_1, \dots, \xi_{t-1} = x_{t-1}) \le$$
$$\le H(\xi_t = x_t \mid \xi_{t-1} = x_{t-1}),$$ egyenlőség

(minden t-re) a.cs.a., ha

$$P(\xi_t = x_t \mid \xi_1 = x_1, \dots, \xi_{t-1} = x_{t-1}) =$$
$$P(\xi_t = x_t \mid \xi_{t-1} = x_{t-1}).$$ Markov-lánc tulajdonság.

A forrás közelítései

Megfigyeljük a forrást M hosszú üzenetet véve. Ebből közelítjük a forrás valószínűség-eloszlását statisztikákból. / Stacionárius forrásokat nézünk azaz időtől ne függjön az eloszlás; időeltolás-invariáns

Az üzenet (diszkrét) lehetséges szimbólumai:
$x_1, x_2, \dots, x_n$

0-ad rendű közelítés: csak n értékeit használjuk fel.

A legbizonytalanabb: az egyenletes eloszlás;

Egy szimbólum választás: ξ_i valószínűségi változó.

$$P(\xi_i = x_j) = \frac{1}{n}, \quad j = 1, \dots, n, \quad \text{minden } i\text{-re}.$$

N szimbólum egymás után: (n^N lehetőség

$$P(\xi_1 = x_{i_1}, \xi_2 = x_{i_2}, \dots, \xi_N = x_{i_N}) = n^{-N}.$$

/ független, azonos, egyenletes eloszlású változó

Entrópia: $H(\xi_i) = \log_2 n$

$$H(\xi_1, \dots, \xi_N) = N\,H(\xi) = N \log_2 n$$

$$\lim_{N\to\infty} \frac{1}{N} H(\xi_1, \dots, \xi_N) = \log_2 n = H(\xi)$$

Első-rendű közelítés:

Megfigyeljük M-ből:

m_i : az x_i előfordulásainak száma, $i = 1, \dots, n$.

Rögzítjük ξ_i eloszlásának ~~becs~~ közelítését:

$$P(\xi_i = x_j) = \frac{m_j}{M} = p_j, \quad j = 1, \dots, n,$$

minden i-re.

A vetületi eloszlások rögzítése mellett a legbizonytalanabb együttes eloszlás a független eloszlás.

Első rendű közelítés: független, azonos eloszlású valószínűségi változók.

$$P(\xi_1 = x_{i_1}, \dots, \xi_N = x_{i_N}) = \prod_{j=1}^{N} P(\xi_j = x_{i_j}) =$$

$$= \prod_{i=1}^{n} p_i^{\mu_i}, \quad \text{ahol}$$

μ_i = az x_i előfordulási száma $\{x_{i_1}, \dots, x_{i_N}\}$-ben

Entrópia:

$$H(\xi_i) = -\sum_{j=1}^{n} p_j \log_2 p_j, \quad \text{minden } i\text{-re}$$

Együttes entrópia (függetlenség miatt additív!)

$$H(\xi_1, \dots, \xi_N) = -N \sum_{j=1}^{n} p_j \log_2 p_j \quad ;$$

Egy szimbólumra jutó entrópia:

$$H(\xi) = \lim_{N \to \infty} H(\xi_1, \dots, \xi_N) = -\sum_{j=1}^{n} p_j \log_2 p_j$$

Másod-rendű közelítés:

Megfigyeljük az egymás utáni párok gyakoriságát:

m_{ij} : az $x_i x_j$ előfordulásainak száma

m_i : az x_i előfordulásainak száma

A $p_{j|i} = \dfrac{m_{ij}}{m_i}$ hányados

megfelel a

$$P(\xi_t = x_j \mid \xi_{t-1} = x_i) = \frac{P(\xi_t = x_j, \xi_{t-1} = x_i)}{P(\xi_{t-1} = x_i)}$$

feltételes valószínűségnek.

Ezzel rögzítjük az átmeneti-valószínűségeket.

A 2' Tétel szerint a legbizonytalanabb eloszlás, ami ennek eleget tesz, az egylépéses homogén Markov-lánc:

Minden t-re:

$$P(\xi_t = x_{i_t} \mid \xi_1 = x_{i_1}, \dots, \xi_{t-1} = x_{i_{t-1}}) =$$

$$= P(\xi_t = x_{i_t} \mid \xi_{t-1} = x_{i_{t-1}}) = p_{i|j} \quad .$$

Az együttes eloszlás:

$$P(\xi_1 = x_{i_1}, \dots, \xi_N = x_{i_N}) = P(\xi_1 = x_{i_1}, \dots, \xi_{N-1} = x_{i_{N-1}}) \cdot$$

$$\cdot P(\xi_N = x_{i_N} \mid \xi_1 = x_{i_1}, \dots, \xi_{N-1} = x_{i_{N-1}}) =$$

$$= P(\xi_1 = x_{i_1}, \dots, \xi_{N-1} = x_{i_{N-1}}) P(\xi_N = x_{i_N} \mid \xi_{N-1} = x_{i_{N-1}}) =$$

$$= P(\xi_1 = x_{i_1}) P(\xi_2 = x_{i_2} \mid \xi_1 = x_{i_1}) \dots P(\xi_N = x_{i_N} \mid \xi_{N-1} = x_{i_{N-1}})$$

$$= P(\xi_1 = x_{i_1}) \prod_{i,j} p_{i|j}^{\mu_{ij}}, \quad \text{ahol } \mu_{ij} \text{ az } x_i x_j \text{ átmenetek száma.}$$

Markov lánc entrópiája: határértéke lesz.
Szükséges hozzá az ergodikus (határ)eloszlás.

(Ergodikus Markov-lánc: minden állapotból minden állapot elérhető (véges lépéssel pozitív valószínűséggel), és nem periodikus.)

$H(\xi_1,\dots,\xi_N)$ és $H(\xi_1,\dots,\xi_{N-1})$ viszonya:

$$H(\xi_1,\dots,\xi_N) = H(\xi_1,\dots,\xi_{N-1}) + H(\xi_N \mid \xi_1,\dots,\xi_{N-1}).$$

A Markov-lánc tulajdonság miatt

$$H(\xi_N \mid \xi_1,\dots,\xi_{N-1}) = \sum_{i_1,i_2,\dots,i_{N-1}} P(\xi_1 = x_{i_1},\dots,\xi_{N-1} = x_{i_{N-1}}) \ast H(\xi_N \mid \xi_1 = x_{i_1},\dots,\xi_{N-1} = x_{i_{N-1}}) =$$

áttérünk az ergodikus feltételes Markov-lánc

$H(\xi_N \mid \xi_{N-1})$

$$= \sum_{i_1,\dots,i_{N-1}} P(\xi_1 = x_{i_1},\dots,\xi_{N-1} = x_{i_{N-1}}) \Big(-\sum_{i=1}^{n} P(\xi_N = x_i \mid \xi_1 = x_{i_1},\dots,\xi_{N-1} = x_{i_{N-1}}) \ast \log_2 P(\xi_N = x_i \mid \xi_1 = x_{i_1},\dots,\xi_{N-1} = x_{i_{N-1}}) \Big) =$$

Markov-lánc, és ξ_{N-1} eloszlása az együttes eloszlásból:

$$= \sum_{j=1}^{n} P(\xi_{N-1} = x_j) \Big(-\sum_{i=1}^{n} P_{i|j} \log_2 P_{i|j} \Big) =$$

$$= \sum_{j=1}^{n} P(\xi_{N-1} = x_j)\, H(\xi \mid x_j).$$

Visszahelyettesítve $H(\xi_1,\dots,\xi_N)$-be, és alkalmazva ugyanezt visszafelé $N=2$-ig, kapjuk:

$$H(\xi_1,\dots,\xi_N) = \sum_{t=2}^{N} \sum_{j=1}^{n} P(\xi_{t-1} = x_j)\, H(\xi \mid x_j) =$$

$$= \sum_{j=1}^{n} H(\xi \mid x_j) \Big(\sum_{t=2}^{N} P(\xi_{t-1} = x_j) \Big)$$

Kihasználva, hogy ergodikus Markov-láncok létezik

$$\lim_{t\to\infty} P(\xi_t = x_i) = \pi_i \quad \text{határeloszlás}, \quad i = 1,\dots,n.$$

A π_i az átmenet-valószínűségi mátrix sajátvektora:

$$P(\xi_t = x_i) = \sum_{j=1}^{n} P(\xi_{t-1} = x_j)\, P_{i|j},$$

Mindkét oldalt határértékhez véve:

$$\pi_i = \sum_{j=1}^{n} \pi_j P_{i|j}.$$

Visszatérve az entrópiára:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=2}^{N} P(\xi_{t-1} = x_j) = \pi_j, \quad \text{ezért}$$

$$H(\xi) = \lim_{N\to\infty} \frac{1}{N} H(\xi_1,\dots,\xi_N) = \sum_{j=1}^{n} \pi_j H(\xi \mid x_j) \quad \blacksquare$$

Lehet 3-ad, 4-ed, stb. közelítés, ami két-, három- stb. lépéses Markov lánc.

Nyelvek esetén szokás/lehetőség szavakat tekinteni a jelek választásának. Lehet a szavak felett 0-ad, 1-ső, 2-od, .. stb. közelítés.

(A 2-od rendű alapján szimulálva értelmes mondattöredékek nagy gyakorisággal keletkeznek.)

Hosszú sorozatok jellemzése

Független, azonos eloszlású változók esetére bizonyítás
(Első-rendű közelítés)

$\vec{\xi}$ $\quad \xi_1, \xi_2, \dots, \xi_N$ – független, azonos eloszlású sorozat.
$P(\xi = x_i) = p_i$ a közös eloszlás.

Legyen az $\vec{x} = x_{i_1}, x_{i_2}, \dots, x_{i_N}$ megfigyelés olyan, hogy az x_i előfordulásainak száma $\sim N p_i$ legyen $i = 1, \dots, n$

Ekkor
$$P(\vec{\xi} = \vec{x}) \cong \prod_{i=1}^{n} p_i^{N p_i} = 2^{-N\left(-\sum_{i=1}^{n} p_i \log_2 p_i\right)} = 2^{-NH}.$$

/A feltevés a nagy-számok ~~tör~~ törvényeinek megfelelő tipikus gyakoriságnak tekinthető./

Igen hosszú (N elég nagy) sorozatokra az x_i gyakorisága, μ_i, tipikus esetben $(1-\delta) N p_i < \mu_i < (1+\delta) N p_i$ közé esik.

Ebből az ilyen feltételt kielégítő sorozatok valószínűsége $2^{\sum_{i=1}^{n} \mu_i \log_2 p_i}$ becslésével alulról is felülről becsülhető.

A pontos megfogalmazás:

Tétel. Tetszőleges $\varepsilon > 0$ és $\varrho > 0$-hoz létezik N_0, hogy $N > N_0$ esetén az N hosszú sorozatok két halmazba sorolhatók:

(1) A lényegtelen halmaz. Ebbe legfeljebb ε valószínűséggel esik a választás.

(2) A tipikus halmaz. Ebbe legalább $1-\varepsilon$ valószínűséggel esik a választás. Tetszőleges $\vec{x}$ tipikus sorozatra
$$\left| \frac{\log_2 P(\vec{\xi} = \vec{x})}{N} - H \right| < \varrho$$

$$\left| \frac{\log_2 P^{-1}(\vec{\xi} = \vec{x})}{N} - H \right| < \varrho . \quad \blacksquare$$

Kifejezve $P(\vec{\xi} = \vec{x})$ becslése:
$$2^{-N(H+\varrho)} < P(\vec{\xi} = \vec{x}) < 2^{-N(H-\varrho)} .$$

Bizonyítás: a Csebisev-egyenlőtlenségen ~~vezethető~~ alapul.

A μ_i, $i = 1, \dots, n$ gyakoriságokat kell becsülni.

A μ_i (N, p_i) paraméterű binomiális eloszlású változó.

$M(\mu_i) = N p_i$, és $\quad$ (μ_i felírható N db. független, 0-1 értékű valószínűségi változó összegeként.)
$D^2(\mu_i) = N(1-p_i) p_i$.

A Csebisev-egyenlőtlenség η valószínűségi változóra:
$$P\left((\eta - M(\eta))^2 > \gamma^2 \right) \le \frac{D^2(\eta)}{\gamma^2} .$$

Ezt alkalmazva μ_i-re:
$$P\left((\mu_i - N p_i)^2 > \gamma^2 \right) \le \frac{N(1-p_i) p_i}{\gamma^2}$$

Ebből $\gamma = \varrho' N$ választással
$$P\left((\mu_i - N p_i)^2 > (\varrho' N)^2 \right) \le \frac{N(1-p_i) p_i}{N^2 \varrho'^2} =$$
$$= N^{-1} K < \varepsilon' ,$$ ha N elég nagy ε'-höz viszonyítva.

Tehát

(*) $$P\left(-\varrho' N < (\mu_i - N p_i) < \varrho' N \right) \ge 1 - \varepsilon' ,$$ minden i-re, ha N elég nagy.

Ebből a lényegtelen halmaz: legalább egy i-re a $|M_i - Np_i| > \delta' N$ teljesül, és

$$P(\exists i: |M_i - Np_i| > \delta' N) < \sum_{i=1}^{n} P(|M_i - Np_i| > \delta' N) \le n\varepsilon'.$$

Tehát $\varepsilon' = \frac{\varepsilon}{n}$ választással kaptuk a lényegtelen halmazt.

A lényeges halmaz ennek komplementere, tehát minden $\vec{x}$ elemén $\vec{\xi} = \vec{x}$ bekövetkezésekor tartozó M_i értékekre:

$$N(p_i - \delta') < M_i < N(p_i + \delta'), \quad i = 1, \dots, n$$

teljesül.

Ebből $P(\vec{\xi} = \vec{x}) = 2^{\sum M_i \log_2 p_i}$ megbecsülhető mindkét oldalról:

$$2^{N\left(\sum_{i=1}^{n} p_i \log_2 p_i + \delta' \sum_{i=1}^{n} \log_2 p_i\right)} < P(\vec{\xi} = \vec{x}) < 2^{N\left(\sum_{i=1}^{n} p_i \log_2 p_i - \delta' \sum_{i=1}^{n} \log_2 p_i\right)}$$

Reciprokot véve (megfordul az egyenlőtlenség), majd logaritmust és N-nel osztva:

$$H - \delta' \sum_{i=1}^{n} \log_2 p_i > \frac{\log_2 P^{-1}(\vec{\xi} = \vec{x})}{N} > H + \delta' \sum_{i=1}^{n} \log_2 p_i .$$

A $\delta = -\delta' \sum_{i=1}^{n} \log_2 p_i$ választással kapjuk a 4. Tétel bizonyítását.

Következmény: $N > N_0$ esetén a tipikus halmaz M_N elemszámára alsó és felső becslést kapunk a (valószínűségek reciproka segítségével):

$$(1-\varepsilon)\, 2^{N(H-\delta)} < M_N < 2^{N(H+\delta)} .$$

A forrás kódolása csatorna jelekké

Forrás: $\xi_1, \xi_2, \dots, \xi_N, \dots$ v. változó sorozat

Kódolás eredménye:

$$\eta_1\, \eta_2 \dots \eta_{M(\xi_1, \dots, \xi_N)}$$ – kódsorozat,

Ha a kódolás $f(x_{i_1}, \dots, x_{i_N}) = (y_{i_1} \dots y_{i_{M(x_{i_1}, \dots, x_{i_N})}})$ függvény, akkor

$$H(\xi_1 \dots \xi_N) \ge H(f(\xi_1 \dots \xi_N)) = H(\eta_1, \dots, \eta_{M(\xi_1, \dots, \xi_N)})$$

Ha nincs veszteség – minden visszaállítható, azaz létezik f^{-1} inverz, amin

$$H(\eta_1, \dots, \eta_M) \ge H(f^{-1}(\eta_1 \dots \eta_M)) = H(\xi_1, \dots, \xi_N).$$

Ebből veszteségmentes esetben

$$H(\xi_1 \dots \xi_N) = H(\eta_1, \dots, \eta_{M(\xi_1, \dots, \xi_N)}).$$

Ha minden T ideig megy, a csatorna bemenetén

$$N(T) = 2^{T(C \pm \delta)}$$

jelsorozat jelenik meg. Ezen a egyenletes eloszlás a maximális entrópiája, $T(C \pm \delta)$.

összetettebb eset: a kódoló egy véges állapotú automata. Az automata állapota ilyenkor belső, nem ismert érték:

$$f(x_1,\dots,x_N) = \left(y_1,\dots,y_{M(x_1,\dots,x_N)},\ \alpha(x_1,\dots,x_N)\right),$$

$$H(\xi_1,\dots,\xi_N) \geq H(f(\xi_1,\dots,\xi_N)) = H\left(\eta_1,\dots,\eta_{M(\xi_1,\dots,\xi_N)},\ \alpha(\xi_1,\dots,\xi_N)\right)$$
$$\geq H\left(\eta_1,\dots,\eta_{M(\xi_1,\dots,\xi_N)}\right).$$

A csatorna kimenetén ebben az esetben is TC entrópia a maximális.

Az alaptétel bizonyítása:

Forrás: H bit/szimbólum entrópia, V szimbólum/sec sebesség
Csatorna: C bit/sec

Tétel: a) Nem lehet a forrást C/H-nál nagyobb sebességgel működtetni, hogy a csatornán minden veszteségmentesen átvihető legyen.

b) Tetszőleges $\delta > 0$-hoz létezik kódolás, hogy $C/H - \delta$ sebességgel működtetve a forrást, minden veszteségmentesen átvihető legyen.

Bizonyítás:

a) Legyen a sebesség V.
T idő alatt a forrás entrópia: $[TV(H-\delta),\ TV(H+\delta)]$ közé esik.
A csatorna-jelek száma: $2^{T(C-\delta)} < N(T) < 2^{T(C+\delta)}$, a csatorna-kimeneten az entrópia legfeljebb $T(C+\delta)$ lehet.
Tehát $TV(H-\delta) \leq T(C+\delta)$, amiből

$$V \leq \frac{C+\delta}{H-\delta} \to \frac{C}{H}.$$

b) N hosszú üzenetben az átlagos átviteli idő. Kétféle hosszúságú kód: tipikusra rövidebb, lényegtelenre hosszabb.
(Független, azonos eloszlás!)

A görög betűk tetszőleges kis mennyiséget jelölnek, ha N vagy T elég nagy.

N-hez a hosszú sorozatok jellemzése szerint

(i) A tipikus halmaz $1-\varepsilon$-nál nagyobb valószínűségű.
Elemeinek száma: $\leq 2^{N(H+\delta)}$

(ii) Lényegtelen halmaz: legfeljebb ε valószínűségű, elemeinek száma $\leq 2^{N\log_2 n}$.

Kódhossz időben

(i)-hez: T_1 idő, amire

$$2^{T_1(C-\delta)} > 2^{N(H+\delta)}.$$

Ekkor $T_1 > N\left(\frac{H+\delta}{C-\delta}\right) = N\left(\frac{H}{C} + \delta_1\right)$.

(ii)-höz: egy speciális T_1 hosszú kód, majd T_2 idő, amire

$$2^{T_2(C-\delta)} > 2^{N\log_2 n},$$ amiből

$$T_2 > N\left(\frac{\log_2 n}{C-\delta}\right) = N\left(\frac{\log_2 n}{C} + \delta_2\right).$$

Az átviteli idő várhatóértéke:

$$(1-\varepsilon)T_1 + \varepsilon(T_1+T_2) = N\left(\frac{H}{C}+\delta_1\right) + \varepsilon N\left(\frac{\log_2 n}{C}+\delta_2\right).$$

Egy szimbólumra jutó idő:

$$\frac{H}{C} + \delta_1 + \varepsilon\left(\frac{\log_2 n}{C} + \delta_2\right) = \frac{H}{C} + \delta_3.$$

A másodpercenkénti szimbólumok száma ennek reciproka, $V = \left(\frac{H}{C}+\delta_3\right)^{-1} = \frac{C}{H} - \delta.$ ∎

Nevezetes kódolások

Kódolási feladat:

Adott a $P(\xi = x_i) = p_i$, $i = 1, \dots, n$ eloszlás.

Kódoljuk x_i-t a véges bináris sorozatokkal /(0,1) abc/

$x_i - w_i$ a kódja.

Invertálható kódolás: az $x_{i_1} x_{i_2} \dots x_{i_k}$ a $w_{i_1} w_{i_2} \dots w_{i_k}$ kódból visszaállítható legyen.

Elégséges feltétel: prefix-mentes kódolás, (kódfa!)

w_i nem kezdőszelete (prefixe) w_j-nek, ha $i \neq j$.

$l(w_i)$ – az w_i hossza.

A kódolás jósága: a kódhossz várható-értéke:

$$\sum_{i=1}^{n} l(w_i) p_i \, .$$

A csatorna-alaptétel következménye: invertálható kódolásra

$$H(p_1, \dots, p_n) \le \sum_{i=1}^{n} l(w_i) p_i \, .$$

Már láttuk a $q_i = 2^{-l(w_i)}$ eloszlásra a lemma $H(p_1, \dots, p_n) \le -\sum p_i \log_2 q_i$

Állítás: Létezik prefix-mentes kódolás, amin

$$\sum_{i=1}^{n} l(w_i) p_i < H(p_1, \dots, p_n) + 1 \, .$$

Nevezetes kódok:

Shannon – Fano kód:

Feltesszük: $p_1 \ge p_2 \ge \dots \ge p_n$.

Legyen $Q_1 = 0$, $Q_i = \sum_{j=1}^{i-1} p_j$, $i = 2, \dots, n$.

w_i megadása: fejtsük ki binárisan Q_i-t olyan l_i hosszban, amin

$$2^{-l_i} \le p_i < 2^{-l_i+1} \quad \text{legyen.}$$

Ebből $\quad l_i \ge \underbrace{-\log_2 p_i > l_i - 1}$ és p_i-vel szorozva, Σ

A kódhossz várható-értéke:

$$\sum_{i=1}^{n} p_i l_i < \sum_{i=1}^{n} p_i(-\log_2 p_i + 1) = H(p_1, \dots, p_n) + 1.$$

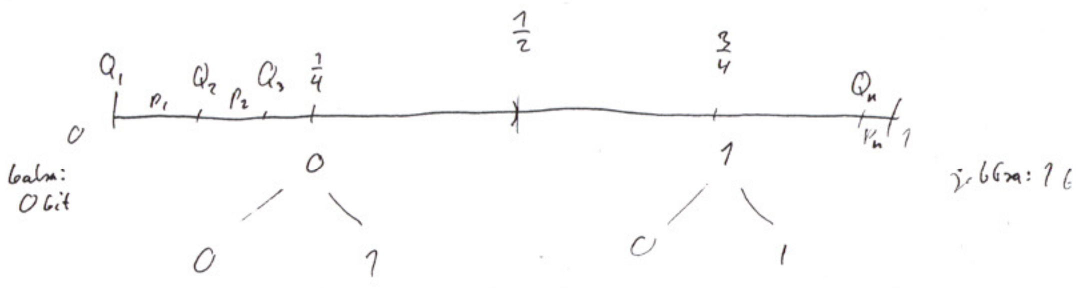
Bizonyítandó a prefix-mentesség:

- A kódszavak, mint bináris törtek
  - monoton növekvő értékek
  - hosszuk is monoton növekszik.
- Egy kódszó csak a nagyobb indexű prefixe lehet, de akkor már a rákövetkezőnek is prefixe, hiszen az kevesebbel nőtt értékben, mint a többi rákövetkező.
- Elég w_i és w_{i+1}-re bizonyítani, hogy w_i nem prefixe w_{i+1}-nek.

  w_i hossza, l_i olyan, hogy p_i első értékes jegye – vagyis 1 bitje éppen az l_i-edik.

  tehát $Q_{i+1} = Q_i + p_i$ miatt w_i-hez utolsó értékes jegyénél ~~ta~~ nagyobb értéket adunk, ezért az első l_i bitben lesz változás.

Szemléltetés:



addig felezünk, amíg az intervallumban csak egy Q_i pont marad.

2. Gilbert – Moore kód

szemléltetése hasonló, nem kell rendezni.

A felezőpontokra a felezős kódolás.
A felezőpont értéke 1 bittel hosszabb kiírását igényli,
ezért $\sum_{i=1}^{n} p_i l_i < H(p_1,\dots,p_n)+2$ teljesül.

3. Optimális kód: Huffman-kód:

Rekurzív felépítés: az eloszlás elemszáma szerint:

- $(q_1,\dots,q_s)$-re ismerjük.
- $(p_1,\dots,p_s,p_{s+1})$-re legyen p_s, p_{s+1} a két legkisebb.

p_s és p_{s+1} kódja legyen azonos hosszú,
p_s : $w0$, p_s : $w1$ alakú.

A $(p_1,\dots,p_{s-1}, p_s+p_{s+1}=q)$ eloszláshoz készítsük el a Huffman-kódot, és legyen w a q kódja.

Állítás: a prefixmentes kódok között a Huffman-kód optimális.

Bizonyítás: Vegyünk tetszőleges $w_1, w_2, \dots, w_n$ prefix
(Indukcióval) kódot $p_1, p_2, \dots, p_n$-hez.

Legyen p_{n-1} és p_n a két legkisebb érték.
Ha nem a leghosszabb kódok tartoznak hozzájuk, cseréljük fel. Ezzel a várható érték nem nő.
Átrendezhetjük a leghosszabb kódokat (párban vannak), hogy p_{n-1} és p_n kódja $w0$ és $w1$ legyen.
Összevonva $p_{n-1}+p_n = q - t$, a feladatot eggyel rövidebb eloszlásra vezettük.

Egy szimbólumra jutó kódhossz várható értéke
$1/N$-re közelíti H-t:

A $\xi_1,\dots,\xi_N$ együttes eloszlásához a Huffman kód, vagy Shannon-Fano kóddal elérhető
$$H(\xi_1,\dots,\xi_N)+1$$
várható értékű kód.

Mivel $H(\xi_1,\dots,\xi_N) = N(H+\delta)$,

az egy szimbólumra jutó kódhossz
$$\frac{1}{N}\left(H(\xi_1,\dots,\xi_N)+1\right) \le H+\delta+\frac{1}{N}.$$

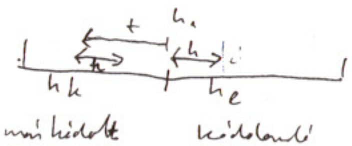
Univerzális forráskódolás

adaptív kód: nincs előre adott kódoló-rendszer, amit konstans kezdeti kódként át kell küldeni. (mint a Shannon-Fano v. Huffman-kód)

Lempel-Ziv-kódok

LZ77-algoritmus.

Csúszóablakban nézi a forrásszimbólumokat



- Minden szimbólumnak van kódja, $\log_2 n$ hosszú.

A kódolandó karaktersort keressük visszafelé a keresőpufferben. a leghosszabb egyezést megnézi(?), a kód: $\langle t, h, c \rangle$ hármas, ahol

- t a visszafelé lépés lépésszáma $\log_2 h_k$
- h az egyezés hossza $\log_2 h_k$
- c az egyezés utáni első karakter kódja. $\log_2 n$

(addig küld a kódolás, c-t is belevéve.)

$h_k, h_e \to \infty$ esetén aszimptotikusan optimális kódoló tetszőleges stacionárius és ergodikus forrásra.

LZ78 algoritmus:

szótárépítő, kezdetben minden szimbólumnak van kódja.

Módszer: az új kódolandó szövegrészben megkeresi a leghosszabb egyező, már kódolt szövegrészt, kódja i, átküldi az $\langle i, c \rangle$ párt, ahol c a következő szimbólum kódja.

Felveszi az átküldött szöveget a következő szabad indexre.

A szótár végtelenül(?) fel növekszik - ezért vagy végül törlik - fix szótár vagy elölről kezdi(?)

LZW (T. Welch) (c helyett(?) "a" kódja)

módosítás: $\langle i, c \rangle$ pár helyett c-t nem küldi, csak szótár-kódokat küld. Ezért az $\langle i, c \rangle$-vel megfelelő új szó szótárba kerül, viszont a kódolás "a"-val kezdődik.

Ezt használják a Unix compress

GIF (Graphics Interchange Format)

szótár mérete paraméterként kezelhető.

- bejegyzés hossza szerint (kitöltődés)

Visszaszámított(?): már tömörített szöveg újra-tömörítése akár hosszabb lehet!

Kombinatorikus entrópia:

n elem, kódolni mindet,

kódok összesített száma: $\sim n\cdot\log_2 n$.

$n_1,\dots,n_k$, $\sum n_i = n$, ismert(?) ..., szimbólumonként kódoljuk:

$n_i \cdot \log_2 n_i$ - az i-edik szim.

$$n\cdot\log_2 n - \sum_{i=1}^{k} n_i\log_2 n_i = n\left(-\sum_{i=1}^{k} \frac{n_i}{n}(\log_2 n_i - \log_2 n)\right) = n\,H\left(\frac{n_1}{n},\dots,\frac{n_k}{n}\right)$$

Kolmogorov bonyolultság – kiszámítható függvények

Bevezetés

A Shannon-entrópia szerepe:
ami majd jövőben bekövetkezhet, (üzenet) azt hogyan tudjuk legrövidebben, várható-értékben leírni.
Ez a jövőn való felkészülés – ami várható, is ami várható, azt különböztessük meg, átlagosan a legrövidebben.

A fordított feladat: ami a múltban bekövetkezett, azt hogyan tudjuk a legrövidebben jellemezni?

A múlt: egy hosszú adatsorként jelenjen meg. Hogyan lehet legtömörebben leírni? A leírásból algoritmus (számítógép) állítsa vissza!

Kolmogorov kérdése: Hogyan lehet a „Háború és békét" legrövidebben kódolni, hogy abból számítógép visszaállítsa?

(A kérdést alapvetően a véletlenszám-generátorok jóságának ellenőrzése, mérése vetette fel. Mennyire jó egy véletlen szám-tábla? Nem jó, ha ~~nem~~ lehet tömöríteni, amiből újra kiszámolható.)

Mit jelent a tömörítés? x tömörítése (kódja) p egy f függvény szerint, ha $f(p) = x$.

x alakja	$f(p)$	x-kód p hossza ($l(p)$)	x hossza (bináris)
$x \in N$	$f_0(p)=p$	$l(p) = l(x)$	$L(x) = L(p) \approx \log_2 x$
$x = 2^k$	$f_1(p)=2^p$	$l(p) = \log\log x$	$l(x) = 2^{l(p)}$
$x = 2^{2^k}$	$f_2(p)=2^{2^p}$	$l(p) = \log\log\log x$	$l(x) = 2^{2^{l(p)}}$
$x = 2^{2^{\cdot^{\cdot^{2^k}}}}$	$f_3(p)=2^{2^{\cdot^{\cdot^{2^p}}}}$	$l(p) = \log \underbrace{\log \ldots \log}_{p} x$	

Ezt a négy függvényt használva x kódolásához, észrevehetjük, melyik szerint (ha lehet) legrövidebb a kód. Ennek módszere (i,p) pár használható, tudni kell, melyik a dekódoló függvény.

Hogyan kódolnánk az 1025-öt?

például $f_4(p) = f_1(p) + 1$. Ez jobb, mint az $f_0, \ldots, f_3$ közül egyedül használható f_0 szerinti kód.

Általában, nagyon sok dekódolót választva a fenti módszernél tudunk egy továbbit választani, ami speciális számokra tömörebb kódot használ.

Milyen függvények választhatók a kódoláshoz?

Kiszámítható függvények.

Rekurzívan felsorolhatóak – ezért az (i,p) pár segítségével konstruált módszer működik.

A pontos definícióhoz, tulajdonságokhoz, használatához és tulajdonságainak vizsgálatához szükséges a kiszámítható függvények világának felidézése.

Parciális rekurzív függvények – kiszámíthatóság

Felépítésük: intuitív kiszámítható alapfüggvényekből kiszámítható módon újabb függvényeket képezünk.

$N = \{0, 1, 2, \dots\}$,

A függvények $N^k \to N$ típusúak.

Alapfüggvények (bázis függvények)

- zéró függvény: $Z : N \to N$, $Z(x) = 0$
- rákövetkező (successor): $S : N \to N$, $S(x) = x + 1$
- vetítő (projekció): $P_{n,i} : N^n \to N$, $P_{n,i}(x_1, x_2, \dots, x_i, \dots, x_n) = x_i$.

Függvényképzések:

Kompozíció: az $f : N^n \to N$ függvény a
$h : N^k \to N$ és
$g_1, \dots, g_k : N^n \to N$ -be képező
függvényekből kompozícióval keletkezik, ha
$$f(x_1, \dots, x_n) = h(g_1(x_1, \dots x_n), g_2(x_1, \dots, x_n), \dots, g_k(x_1, \dots, x_n)).$$
(kiszámítási menete egyszerűen megadható).

Primitív rekurzió: az $f : N^{n+1} \to N$ függvény a
$g : N^n \to N$ és
$h : N^{n+2} \to N$ függvényekből
primitív rekurzióval keletkezik, ha
$$f(x_1, \dots, x_n, 0) = g(x_1, \dots, x_n) \quad \text{(kezdő érték)}$$
$$f(x_1, \dots, x_n, y+1) = h(x_1, \dots, x_n, y, f(x_1, \dots, x_n, y)).$$

Zártság a primitív rekurzióra nézve –
primitív rekurzívan zárt függvényosztály C, ha Z, S, $P_{n,i}$ benne van, és a kompozíció és a primitív rekurzió nem vezet ki belőle.

Primitív rekurzív függvények: generáló sorozattal megadhatók.

$g_1, g_2, \dots, g_n = f$ az f generáló sorozata, ha

- vagy – g_i alapfüggvény
- vagy – g_i a nála kisebb indexűekből kompozícióval keletkezik
- vagy – g_i a nála kisebb indexűekből primitív rekurzióval keletkezett.

Tétel: A primitív rekurzív függvények osztálya a primitív rekurzív függvényosztályok metszete. ~~(Zárt, ...~~
(Két zárt osztály metszete is zárt)

Biz: a) Zárt osztály a primitív rekurzív függvények osztálya:
– alapfüggvény nyilván magának generálója
– generáló sorozatok „átindexelő" konkatenálása a kompozíció, ill. primitív rekurzió szerint generáló sorozat lesz.

b) Ha C zárt, akkor a $g_1, g_2, \dots, g_n = f$ benne van. n szerinti indukcióval egyszerűen követhető.

Megjegyzés: a primitív rekurzió helyett az iterációt is választhatjuk:
$F : N^2 \to N$ az $f : N \to N$ -ből iterációval keletkezik, ha
~~$F(x_1, \dots, x_n, 0) =$~~
$$F(x, 0) = x$$
$$F(x, y+1) = f(F(x, y)).$$

Példák ~~...~~ kiszámítható, de nem primitív rekurzív függvényre:

1.P. Tétel: Létezik kiszámítható univerzális függvény, $U : N^2 \to N$, hogy bármely $f : N \to N$ primitív rekurzív függvényhez létezik k, hogy $U(x,k) = f(x)$, és az $U(x,k)$ minden k-ra primitív rekurzív.

/Bizonyítás vázlata: a generáló sorozatokat fel tudjuk írni véges ABC feletti szó formájában. A generáló sorozatok grammatikáját egyszerű, könnyű eljárást adni, hogy egy ~~sorozat~~ szó generáló sorozatot jelent-e.
A szavak hossz-lexikografikus sorrend szerinti felsorolásával megkapható az első $f_1 : N \to N$ generáló sorozat, majd a második, stb.
Az $U(x,k)$-t úgy definiáljuk, hogy kikeressük a k-adik generáló sorozatot, majd használjuk a generálás által adott számítást x-re.) /

Állítás: a $h(x) = U(x,x)+1$ nem lehet primitív rekurzív.

Biz: Ha az lenne, legyen a sorszáma k.

$$h(k) = \begin{cases} U(k,k) \\ U(k,k)+1 \end{cases}$$ ellentmondás.

2.P. Az Ackerman függvény (ek):

$A : N^2 \to N$; $A(0,y) = y+1$; $A(x+1,0) = A(x,1)$, és $A(x+1, y+1) = A(x, A(x+1,y))$.

$a_m : N \to N$, $a_m(x) = A(m,x)$;

a_{m+1} az $A(m+1,x+1) = A(m, A(m+1,x))$,

vagyis $a_{m+1}(x+1) = a_m(A_{m+1}(x))$,

iteráció

Az a_m függvények rendkívül felgyorsulnak.

a_3 más exponenciális,

a_4 x magasságú torony-hatvány

a_5 : olyan toronyhatvány, aminek a magassága x-magasságú toronyhatvány.

$\vdots$

Az $f : N^k \to N$ az r-Ackerman szinten belül van, ha $X = \max(x_1, \dots, x_n)$ jelöléssel $f(x_1, \dots, x_n) \le A_r(X)$, (véges sok helytől eltekintve)

Állítás: Legyen f primitív rekurzív. Akkor f valamilyen r-re az A-szinten belül van.

(Biz. vázlat: a függvényképzések csak véges szinttel emelik a függvény A-szintjét.)

Következmény: az $f(x) = A_x(x)$ függvény nem primitív rekurzív.

További képzés szükséges: a minimalizáció!

Definíció: Az $f : N^n \to N$ függvény a $g : N^{n+1} \to N$ függvényből minimalizációval keletkezik, ha $f(x_1, \dots, x_n) = k$, amennyiben
1. $g(x_1, \dots, x_n, k) = 0$
2. $y < k$-ra $g(x_1, \dots, x_n, y) \ne 0$
3. $g(x_1, \dots, x_n, y)$ értelmezve van $y \le k$-ra.

Nincs értelmezve f az $(x_1, \dots, x_n)$ helyen, ha vagy $g(x_1, \dots, x_n, k) \ne 0$, minden k-ra, vagy nem teljesül a 3. feltétel.

Ezzel parciális függvényeket kapunk.

Jelölés: $f(\vec{x}) = \mu y[g(\vec{x}, y) = 0]$)

Speciális minimalizáció: reguláris minimalizáció

Az $f: N^n \to N$ a $g: N^{n+1} \to N$ -ből reguláris minimalizációval keletkezik, ha minimalizációval keletkezik

a) – minimalizációval keletkezik

1., minden $x_1,\dots,x_n$-hez létezik y, hogy

$$g(x_1,\dots,x_n,y)=0$$

2. A g totális függvény.

Parciális rekurzív függvények halmaza:
primitív rekurzív függvények + minimalizáció.

Rekurzív (totális) függvények halmaza:
primitív rekurzív függvények + reguláris minimalizáció.

Zártság – generáló sorozat ugyanúgy definiálható, mint primitív rekurzióra nézve.

Lényeges különbség a két osztály között:
- a parciális rekurzív függvények generáló sorozata szintaktikusan jellemezhető,
- a rekurzív függvényeknél a reguláris minimalizáció szemantikus, nem ellenőrizhető.

Tétel Létezik $U: N^2 \to N$ univerzális parciális rekurzív függvény, amelyre minden $f: N \to N$ parc. rek. függvényhez létezik k, hogy

$$U(x,k) = f(x),$$ és mindkettő ugyanott van értelmezve.

Az átló-konstrukció itt nem ellenpélda:
$h(x) = U(~~x~~ x,x)+1$, és k a h kódja,
$h(k) = \begin{cases} U(k,k)+1 \\ U(k,k) \end{cases}$ csak azt jelenti, hogy h nincs értelmezve k-ban.

Az U konstrukciója ugyanaz, mint a primitív rekurzív függvények esetében.

U képzésében a felismerő rész primitív rekurzív, a sorszám keresése minimalizáció. Ezért U parciális rekurzív függvény.

Kleene normál alak: Létezik $U: N \to N$ és $V: N^3 \to N$ primitív rekurzív függvény, hogy bármely $f: N \to N$ parciális rekurzív függvényhez létezik k, hogy

$$f(x) = U(\mu t(V(k,x,t)=0)),$$ minden x-re.

Def. $A \subseteq N$ rekurzív: ha karakterisztikus függvénye rekurzív. (eldönthető $x \in A$)

Def. $A \subseteq N$ rekurzívan felsorolható, ha létezik $f: N \to N$ rekurzív függvény, hogy $A = f(N)$. (Ha $x \in A$, akkor azt véges időben megtaláljuk).

Rice-tétel:
Legyen ~~$A \subseteq N$ tetszőleges~~ $\mathcal{A}$ a parciális rekurzív függvények tetszőleges halmaza. Az $A = \{n \mid f_n \in \mathcal{A}\}$ halmaz a.cs.a. eldönthető, ha $\mathcal{A} = \emptyset$, vagy $\mathcal{A}$ az összes parc. rekurzív fv.

Turing gép:

I/O fej, állapot,

(Q, T, δ, q_0, E)

állapot jelek ABC

q_0 kezdő
E elfogadó állapot(ok)

$\delta: Q \times T \to Q \times T \times (\text{Jobb, bal, helyben})$

Mási jelentés: ha egy függvény kiszámítható, az azt jelenti, hogy minden véges halmazon belül az összes értelmezési helyen véges időben meghatározható értékeit. Viszont (általában) nem tudjuk, hogy ahol még nem kaptunk értéket, ott értelmezve van-e.

| Kiszámítható függvény tér-idő táblázata;
| Végét is: annak hasonló csupa Λ;
| Az ilyen táblázatok nem mind kiszámítható függvények táblázatai!

—— . ——

Kolmogorov bonyolultság
(Kolmogorov – Chaitin – Solomonoff)

Technikai előkészítés.

$N = \{0,1,\dots\}$ nem-negatív egészek

$\Omega = \{0,1\}^*$, véges bináris szavak

Kölcsönösen egyértelmű megfeleltetés:

N	Ω
0	Λ
1	0
2	1
3	00
4	01
5	10
6	11
7	000
8	001
$\vdots$	

x egyaránt $\in N$, illetve $\in \Omega$.

x hossza: $l(x)$

xy : konkatenálás

$x * y$: szorzat.

Rendezett párok kódolása:

	0	1	2	3
0	0	2	5	9
1	1	4	8	
2	3	7		
3	6			

mellékátlók szerint ki,

$P(x,y) = z$ $\qquad l(z)$ – nagy mindkét változóban.

önkorlátozó kód: $x = x_1 \dots x_n$

$d(x) = x_1 x_1 \dots x_n x_n 01$

$\bar{x}^n = 1^n 0 x$ $\qquad$ | mindkettőt használják

Standard önkorlátozó kód:

$x' = \overline{l(x)}\, x$ $\quad ; \quad l(x') = 2\, l(l(x)) + l(x) + 1$

Rendezett pár kódolása:

$\gamma_0(x,y) = \bar{x} y$, $\qquad \pi_{01}(\bar{x} y) = x$, $\quad \pi_{02}(\bar{x} y) = y$

$\gamma_1(x,y) = x' y$, $\qquad \pi_{11}(x' y) = x$, $\quad \pi_{12}(x' y) = y$

$l(\gamma_1(x,y)) = 2\, l(l(x)) + 1 + l(x) + l(y)$

+: Ide a prefix-Turing gépek, és univerzális TM!

Bonyolultság adott függvény szerint:

Df. Az $x \in \Omega$ bonyolultsága az $f: N \to N$ parciális rekurzív függvény szerint:

$$C_f(x) = \min_p \{ l(p) : f(p) = x \}, \qquad \infty, \text{ ha}$$

nincs p, amin $f(p) = x$.

(Megjegyzés: $C_f(x)$ általában nem kiszámítható!)

Relatív bonyolultság. Kolmogorov – Chaitin – Solomonoff alaptétele, vagy újabban invariancia-tétel:

Tétel Létezik $f_0 : N \to N$ optimális parciális rekurzív függvény, amelyhez bármely $f : N \to N$ parciális rekurzív függvényhez létezik k_f konstans, hogy

$$C_{f_0}(x) \le C_f(x) + k_f, \quad \text{minden } x\text{-re.}$$

Következmény: (invariancia) f_0, g_0 optimális függvények,

$$|C_{f_0}(x) - C_{g_0}(x)| < k_{g_0,f_0}, \quad \text{minden } x\text{-re.}$$

Tehát $C_{f_0}(x)$ konstans erejéig egyértelmű.

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

Definíció Az $x \in N$ (Ω) Kolmogorov-bonyolultságán a $C(x) = C_{f_0}(x)$ számot értjük.

Invariancia tétel bizonyítása:

Legyen f, x tetszőleges. Ha $C_f(x) = k$, akkor legyen $f(p) = x$, és $l(p) = k$.

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

Legyen az f sorszáma n, azaz

$$U(y, n) = f(y).$$

A $z = n'p$ kódból

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

Ezért $C_{f_0}(x) \le l(z) = l(p) + 2\,l(l(n)) + l(n) + 1 = C_f(x) + k_f$.

Tehát f_0 optimális. □

Tulajdonságok:

Két konkrét tulajdonsággal lehet a bizonyításokat 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 $f(p) = k$.

1. Tulajdonság:

$$C(x) \le l(x) + k_1.$$

Bizonyítás: $f(p) = p$ -vel az (a) tulajdonságból.

2. $\lim_{x\to\infty} C(x) = \infty$.

Bizonyítás: ha ∞ sokszor N korlát alatt maradna, az a (b) tulajdonság miatt csak 2^N kód használatát jelentené, ami lehetetlen.

3. $C(x)$ totális függvény, de nem kiszámítható.

Bizonyítás: Indirekt: ha az lenne, akkor legyen

x_1 : első $C(x)$ először nagyobb 2-nél.

x_2 : $x_2 > x_1$, és $C(x)$ először nagyobb 2^2-nél.

⋮

⋮

x_k : $x_k > x_{k-1}$, és $C(x)$ először nagyobb 2^k-nál.

⋮

Legyen $f(p) = x_p$.

$$C_f(x_k) \le \log_2 k \le \log_2(\log_2(C(x_k))).$$

Nyilván

Így

$$C(x_k) \le C_f(x_k) + k_f \le \log_2(\log_2(C(x_k))) + k_f$$

csak véges sok x_k-ra teljesülhet, ellentmondás. □

Az (a) és (b) tulajdonságok alapján $C(x)$-nek mégis lehet valamit mondani!

4. Tulajdonság

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

lehetséges kódok száma (b) miatt:

$$1 + 2 + 4 + \dots + 2^{n-\Delta-1} = 2^{n-\Delta} - 1$$

jelentés: a szavak l.a. fele nem tömöríthető!

5. Tulajdonság:

$m(x) = \min_{y > x} C(y)$, az alsó burkoló.

$\lim_{x\to\infty} m(x) = \infty$ (a (b) miatt)

Ezen túl: $m(x)$ minden kiszámítható, monoton ∞-hez tartó függvénynél lassabban nő, nem kiszámítható!

(Biz: indirekt úton: $m(x)$ növekedési helyén lévő x-nek legyen $m(x)$ a kódja. (a) miatt ez ellentmondásra vezet!)

Lassabb h-ra u.ez az ellentmondás.

6. Tulajdonság: kis változása mennyit változik?

$$C(x+h) \le C(x) + l(h) + 2\log l(h) + c.$$

$f_0(p)=x$, $l(p)=C(x)$ választással

$z = h'p$ -ből

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

$$C_f(x+h) \le l(p) + l(h) + 2\,l(l(h)) + c. \quad \square$$

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

[Adatbázisra értelmezve: nagy adatbázis kis módosítása nem nő a bonyolultság!]

Feltételes Kolmogorov bonyolultság

Kérdés: mennyit segít y ismerete x kiszámításában?
Mekkora a legrövidebb program kódjának hossza, ami y-ból x-et kiszámítja? p kód, $f_p: N\to N$ pr. függv. $f_p(y)=x$.
A programok (kódjai) $p_1, p_2, \dots, p_n, \dots$ legföljebb rekurzívan felsorolható halmaz, egy $h(i)=p_i$ a felsorolás, akkor

~~$f(i,y) = f_{p_i}(y)$ vehető~~

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

Def. Az $f: N^2\to N$ szerinti feltételes bonyolultsága x-nek feltéve y ismeretét:

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

Tétel Létezik optimális $f_0^{(2)}: N^2\to N$ parc. rekurzív függvény, amire bármely $f: N^2\to N$ parciális rek. függvényhez létezik k_f konstans, hogy:

$$C_{f_0}(x|y) \le C_f(x|y) + k_f, \quad \text{minden } x,y\in N\text{-re.}$$

Az f függvény a-ra $f(t,a)$ a értelmezve van $t=0,1,\dots,m_a-1$ helyeken, és máshol sehol. $f(t,a)\ne f(t',a)$, ha $t\ne t'$.
Legyen $x\in B_a$, akkor nyilván létezik t, hogy

$$f(t,a)=x, \text{ és } t \le \{m_a\}.$$

Ebből

$$C(x|a) \le C_f(x|a) + k_f \le l(t) + k_f \le \log_2 m_a + k_f. \quad \square$$

Jelentés: séma + paraméter jó megfogalmazás. Ezen belül már nem érdemes más kódolást használni, csak az ~~egyenlete~~ homogén hosszúságút, azaz $\log_2 m_a$ hosszút.

Összehasonlítás a Shannon entrópiával:

Hosszú sorozat N hosszú; H entrópia.
Tipikus halmaz: $2^{N(H-\delta)} < M(N) < 2^{N(H+\delta)}$.

$M(N)$ - tipikus halmaz elemszáma.
Mi jellemzi a tipikus halmazt?
Egy kiszámítható predikátum, $P_{TIP}(x) = \begin{cases} 0, & \text{ha tip,} \\ 1, & \text{ha nem} \end{cases}$

A paraméter halmazrendszer:

$$\mathcal{A} = \{ N, \{x \mid l(x)=N,\ P_{TIP}(x)=0\} \}$$

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

$$C(X|\mathcal{A}_N) \le \log_2|\mathcal{A}_N| + k_{\mathcal{A}} = \log_2 M(N) + k_{\mathcal{A}}$$

Bebizonyítható, tetszőleges Δ-ra az $\mathcal{A}_N$ Δ-tipikus elemeire az alsó becslés, a ~~tételben~~ teljesül a felső teljesül:

$$\log_2|\mathcal{A}_N| - \Delta = N(H-\delta) - \Delta \le C(x|\mathcal{A}_N) \le N(H+\delta) + k_{\mathcal{A}}.$$

Egy normálva:

$$(H-\delta) - \frac{\Delta}{N} \le \frac{C(x|\mathcal{A}_N)}{N} \le H+\delta + k_{\mathcal{A}}/N. \quad \square$$