Információs rendszerek elméleti alapjai

Információelmélet

Az információ nem növekedés törvénye

Adatbázis – *x* (*x* adatbázis tartalma)

Kérdés *: y*

Válasz*: a = f(y, x)*

Mennyi *az a* információtartalma*:*

Az információ nem növekedés törvénye

$C(a|x) \leq C(y)+h$

ui:

$C(a|x)= C_f(a|x)+k_f \leq C(y)+ k_f$

Ha egy *y* stringet transzformálunk egy parciális rekurzív fv. segítségével., akkor nem nyerhetünk több információt mint az *y* stringbeli információ plusz a transzformáció leírásához szükséges információ mennyisége.

Az információ nem növekedés törvénye

Motiváció: (Solomonoff)

Algoritmikus bonyolultságot vezessük be mint egy univerzális, *a priori valószínűséget* minden véges bináris stringre.

Az információ nem növekedés törvénye

Az <u>*invariancia tétel*</u> bizonyításában használt *U* univerzális függvényt választva, *U* egy *P (i)* valószínűségi eloszlást indukál az egész számok fölött (ugyanígy $\{0,1\}^*$ fölött).

Az információ nem növekedés törvénye

$$-\sum_{i=1}^{n} \underbrace{p_i}_{\text{valószínűség}} \underbrace{\left(log_2 p_i\right)}_{l_i \text{ kódhossz}}$$ mennyiségeket rendeljük

$$p_i = 2^{-l_i}$$

$$x \quad - \underbrace{C(x)}_{\text{kódhossz}}$$ kódhossz (Kolmogorov esetben ezt rendeljük)

Egész számokon adjunk meg egy val.szeg. eloszlást

$$P(i) = 2^{-l_i}$$

$$\sum_{i=0}^{\infty} P(i) = \infty,$$ (nem jó val.szeg. eloszlásnak)

Ha x-et rögzítjük, $$\sum_{\underbrace{f_0(p)=x}_{\text{az összes lehetséges kódra vesszük az összeget}}} 2^{-l(p)} = \infty,$$

Miért? Mert, amit az f_0 felhasznál, nem alkot prefix-mentes kódrendszert

Az információ nem növekedés törvénye

Ha x-et rögzítjük,

$$P(x) = \underbrace{\sum_{U(f,p)=x} 2^{-l(p)}}_{\substack{\text{az összes}\\ \text{lehetséges}\\ \text{kódra vesszük}\\ \text{az összeget}}} = \infty,$$

U az univerzális Turing Gép

bemenet (f,p), két kód

$(l(f), l(p))$, a kódok hossza

f_0, az optimális fv.,

amely ik a legrövidebb kód hosszú

p meghatározásához szükséges

Az információ nem növekedés törvénye

$$P(x) \ge \underbrace{2^{-l\overbrace{(f_0)}^{\substack{x\text{-et kiszámító}\\ \text{függvény}\\ \text{leírása}\\ \text{(bináris kódja/indexe)}}}}}_{\text{konstans}} \cdot \sum_{\underbrace{p\in\{0,1\}^*}_{\substack{\text{az összes,}\\ \text{x-et előállító}\\ \text{lehetséges kódra}\\ \text{vesszük}\\ \text{az összeget}}}} 2^{-l(p)} = \infty$$

Az információ nem növekedés törvénye

Ha x-et rögzítjük, most csak a legrövidebb programokat vegyük figyelembe, amely előállítja x-et.

$P(x) = 2^{-C(x)}$.

$\sum_x P(x) = \infty$

$C(x) \approx \log x$, majdnem minden x-re,

$[\log x \approx I(x)]$

$P(x) \approx \frac{1}{x}$, majdnem minden x-re,

$\sum_x \frac{1}{x} = \infty, x \in \mathcal{N}$ (harmonikus sor)

Az információ nem növekedés törvénye

f_0 – univerzális T-gép segítségével definiáljuk

egy bináris string univerzális valószínűségét ($M(x)$)

$U(i,p) = T_i(p)$.

Pénzfeldobással állítsuk elő a T-gép menetét (input)

és állítsa elő az x-el kezdődő stringet

Def: Annak a valószínűsége, hogy egy pénzfeldobási sorozatra az x-el kezdődő stringet kapjuk kimenetként legyen $M(x)$)

Az információ nem növekedés törvénye

A pénzfeldobási sorozat legyen „*az univerzális a priori valószínűség*”, ebből a *Bayes szabály* segítségével extrapolálható a valószínűség.

És ez legyen ez a valószínűségi eloszlás

Annak a valószínűsége, hogy *x*-et 1 követi:

M(*x1*)/[(*M*(*x0*) + *M*(*x1*))]

Nem jó, mert ….

Nem egyértelmű, hogy mire vezetett a pénzfeldobás.

Prefix Kolmogorov bonyolultság

Csak prefix-mentes függvényeket használunk.

f: $N \to N$,

$\Omega \to \Omega$

Ha az *f* értelmezve van *p*-n és *q*-n, akkor egyik sem prefixe a másiknak ($p,q \in N$)

(Kraft egyenlőtlenség: $\{p_1, \ldots,\}$ prefix-mentes rendszer (prefix-mentes kódfa), akkor :

$$\sum_{i=0}^{\infty} 2^{-l(p_i)} \leq 1$$

Prefix Kolmogorov bonyolultság

Ha f prefix-mentes, akkor a $C_f(x)$ értelmezhető.
Létezik optimális
$g_0:\Omega \to \Omega$, prefix-mentes fv., amire
$\forall$ f prefix-mentes fv. $\exists$ k_f konstans
$C_{g0}(x) \le C_f(x) + k_f$
ui. létezik univerzális felsoroló fv.
Tetszőleges ($\forall$) h: $:\Omega \to \Omega$ parciális rekurzív függvény ,
$\exists$ h_p: prefix-mentes fv.
Idő- tér konstansnak tekintve, futtatjuk a kiértékelést, (átlósan), és ha találtunk értelmezett helyen értelmezett értéket, akkor az tartozik a kódhoz, ha a következő prefixe az előzőnek, akkor kihagyjuk.

Prefix Kolmogorov bonyolultság

T_p – prefix-mentes T-gép – véletlen sorozatok

k bit hosszú helyen, 2^{-k} valószínűséggel áll meg.

Prefix Kolmogorov bonyolultság

$$\sum_{i=0}^{\infty} 2^{-k} \leq 1 \left(\text{Kraft}\right)$$

$K(x) = Cg_0(x)$ - prefix - mentes Kolmogorov bonyolults ág

$\pi(x) = 2^{-K(x)}$ valószínűs égek,

$$\sum_{i=0}^{\infty} 2^{-K(x)} \leq 1$$

$$\pi^*(x) := \sum_{\substack{\infty \\ \underbrace{g_0(p)=x}_{x-g_0 \text{ szerinti összes lehetséges kódja}}}} 2^{-l(p)} < 1$$

Prefix Kolmogorov bonyolultság

g_0-t képzeljük el T-gépnek, akkor annak a valószínűsége, hogy *x*-et ad éppen az a mennyiség

Azaz: a megállás valószínűsége *x* eredménnyel.

(Azok fordulnak elő nagyobb valószínűséggel, amit egyszerűbben tudunk jellemezni)

Prefix Kolmogorov bonyolultság

A $\pi(x)$, $\pi^*(x)$ az egész számok fölött értelmezett valószínűség eloszlás.

A legnagyobb bizonytalanság: az egyenletes eloszlás

Tudunk-e az egész számok halmazán egyenletes eloszlást megadni? Nem !

Mi a lehető legegyenletesebb eloszlás (Ami a lehető leglassabban csökken).

A félig kiszámítható („semi”) eloszlások között a lehető legegyenletesebb eloszlás a $\pi(x)$, $\pi^*(x)$

Félig kiszámítható: nem tudom megmondani, de tetszőlegesen jól közelítem a pontos értéket.

Prefix Kolmogorov bonyolultság

Ha $p(x)$ félig kiszámítható eloszlás, akkor $\exists$ c konstans, hogy $p(x) < c \cdot \pi(x)$ (nem tud $\pi(x)$-nél gyorsabban nőni).

Majoráló eloszlásnak nevezik ezt a tulajdonságot. (algoritmikus tanulás hátterében ez van.)

Probléma: nem tudjuk kiszámolni a $\pi(x)$, csak közelítő értékkel dolgozhatunk

Feltételes prefix (Kolmogorov) bonyolultság

$C(x) = k \Rightarrow \exists\, p : f_0(p) = x$ és $l(p) = k$

$K(x)$ $\quad g(p') = f_0(p) = x$ $\quad p'$ prefix - mentes kód

$l(p') = 2 \cdot log_2 k + k$

$\left[\begin{array}{l}\text{Emlékeztető} : x' = \bar{n}x,\ l(x) = n,\ l(x') = 2\,log_2\ n + l(x) + 1 = \\ = 2\,log 2(l(x)) + l(x) + 1\end{array}\right]$

$\Rightarrow K(x) \leq C(x) + 2 \cdot log_2 C(x) + k$ (Kolmogorov prefix bonyolultság)( )

Feltételes prefix - bonyolultság.

$f(x, p)$ p - ben prefix mentes

$C_f(y|x)$ (feltételes K. bonyolultság)

$g_0^{(2)} : N^2 \to N$ optimális

$C_{g_0}(y|x) \leq C_f(y|x) + k_f \Rightarrow K(y|x)$

Van értelme : $K(x|\mathcal{A})$, $x \in \mathcal{A}$ - nak

Feltételes prefix (Kolmogorov) bonyolultság

Shannon - entrópia esetében : $H(\xi,\eta)=H(\xi|\eta)+H(\eta)$

ugyanez Kolmogorov bonyolults ágra :

(x,y) rendezett párra : $(x,y)=x'y$

$$C((x,y))=C(x'y)$$

$$C(x,y)\le \underbrace{C(y|x)+C(x)}_{\text{túl nagy lehet}}$$

Ez az addítiv tulajdonsá g nem teljesül, de K - ra

$$K(x,y)\le K(x)+K(y|x)+2\log_2 K(x)$$

ha tudnám az x bonyolults ágát, akkor konstans erejéig :

$$K(x,y)\underbrace{=}_{\text{addítiv}} K(x)+\underbrace{K(y|x,K(x))}_{\text{van } x\text{-nek tanúja}}$$

$$x^* : g_0(x^*)=x$$

$$l(x^*)=K(x)$$

Összenyomhatatlanság – „véletlenszerű” string

$K(x) \geq l(x)$

$$C_M(x) = \min\{|\sigma| : M(\sigma) = x\}. \tag{2.1}$$

Here $\min \emptyset = \infty$.

For an example of a drastic compression, let M be the machine that takes an input σ, views it as a natural number and outputs

$$x = 2^{2^{2^{\sigma}}}.$$

Then $C_M(x) = |\sigma| =^{+} \log^{(4)} x$. For instance, the string $\sigma = 10$ corresponds to the number 5, so $M(\sigma) = 2^{2^{32}} = 2^{4294967296} = x$. The string σ is an M-description of x, and $C_M(x) = 2$, while x has length 2^{32}.

kitérő –

$$C(x) = \min\{l(T) + l(p) : T(p) = x\} \pm 1,$$

$$C(x) = \min\{l(T) + C(x|T) : T \in \{T_0, T_1, \ldots\}\} + O(1),$$

$$C(x) \leq l(x) + c \textit{ and } C(x|y) \leq C(x) + c.$$

$$C(xx) \leq C(x) + O(1).$$

$$K(\sigma) \leqslant |\sigma| + K(|\sigma|) + O(1).$$

Matematikai kitérő –

$C(x) \leq l(x) + 8,\ C(\epsilon) = 2,$

$$C(x|y) \leq l(x) + 2,$$
$$K(x) \leq 2l(x) + 231,\ \ K(\epsilon) = 11,$$
$$K(x|l(x)) \leq l(x) + 689,$$
$$K(x) \leq l(x) + 2l(l(x)) + 1066.$$

Matematikai kitérő –

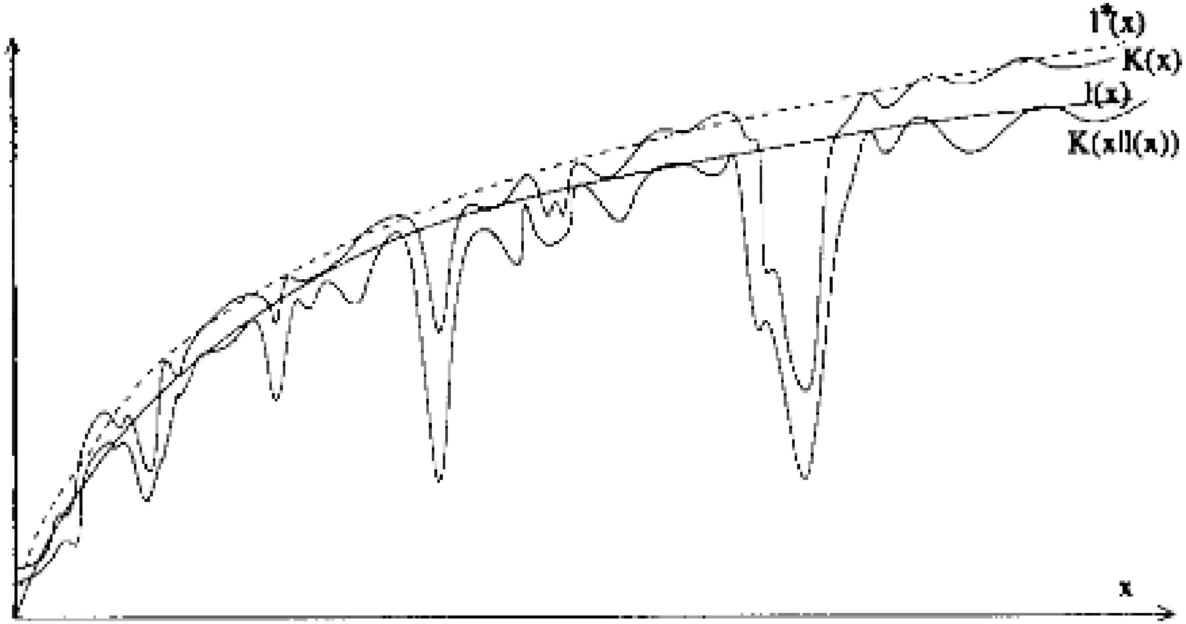

FIGURE 3.2. The graphs of $K(x)$ and $K(x|l(x))$

Matematikai kitérő –

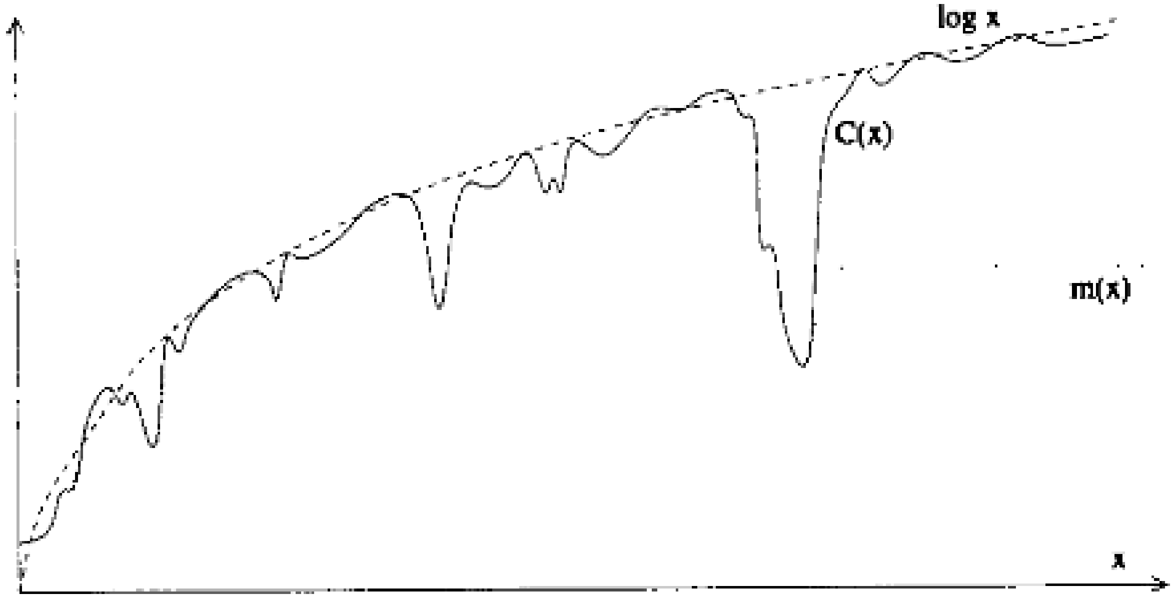

FIGURE 2.1. The graph of the integer function $C(x)$

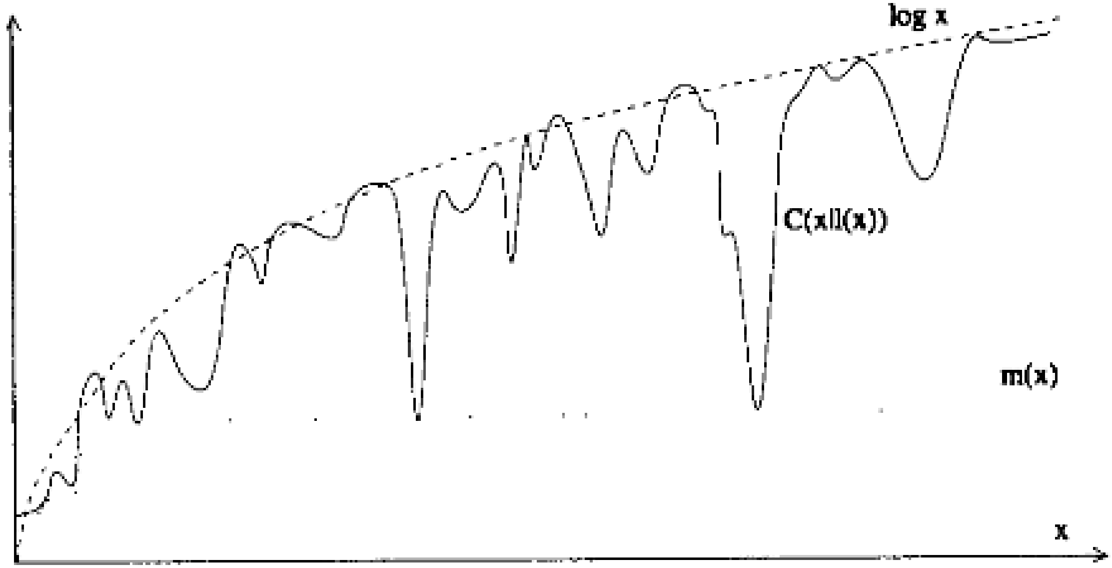

FIGURE 2.2. The graph of the integer function $C(x|l(x))$