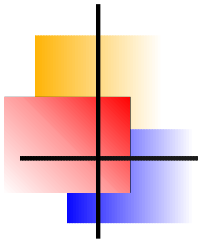
Zavedenie pojmu informácie

Stanislav Palúch

Fakulta riadenia a informatiky, Žilinská univerzita

25. februára 2013

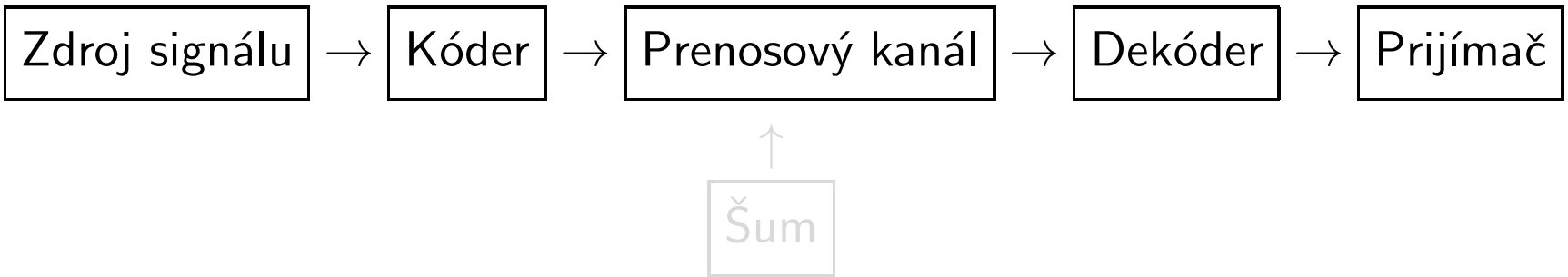


Informačné revolúcie

- Vznik reči
- Objavenie písma
- Kníhtlač 1452 Guttenberg
- Objavenie počítačov a komunikačnej techniky
  - vstup do informačného veku

edi.fmph.uniba.sk/ winczer/historia/GHWZ/index.html

Prenosový reťazec



Prenosový kanál

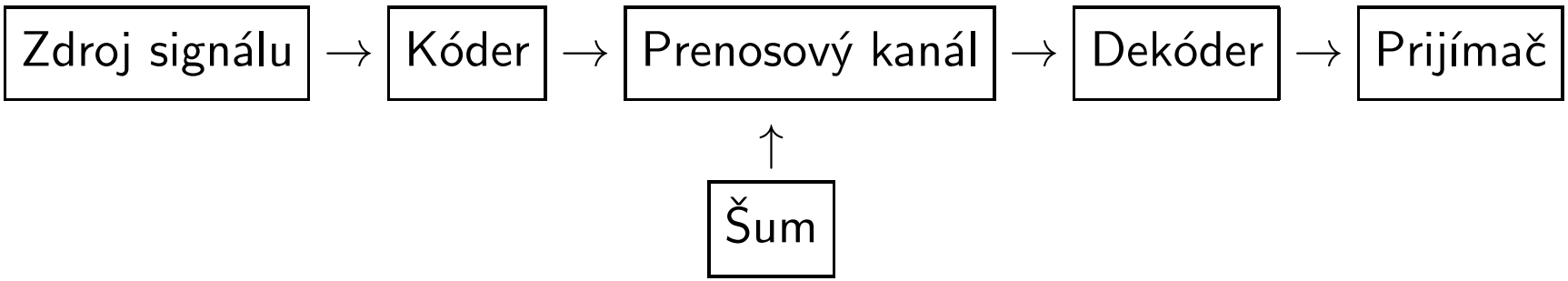
- Prenosový kanál – komunikačné zariadenie so vstupom a výstupom.
- Vstup dokáže spracovávať znaky vstupnej abecedy Y.
- Z výstupu kanála vystupujú znaky výstupnej abecedy Z.

$$y_1, y_2, y_3, \cdots \rightarrow \boxed{\text{Prenosový kanál}} \rightarrow z_1, z_2, z_3, \ldots$$

- Vo väčšine prípadov $Y = Z$.

Prenosový reťazec a prenosový kanál

Prenosový reťazec

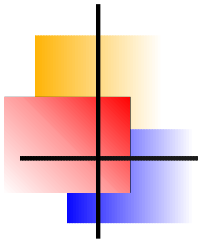


Prenosový kanál

- Prenosový kanál – komunikačné zariadenie so vstupom a výstupom.
- Vstup dokáže spracovávať znaky vstupnej abecedy Y.
- Z výstupu kanála vystupujú znaky výstupnej abecedy Z.

$$y_1, y_2, y_3, \cdots \rightarrow \boxed{\text{Prenosový kanál}} \rightarrow z_1, z_2, z_3, \ldots$$

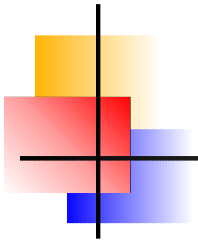
- Vo väčšine prípadov $Y = Z$.



Príklad problému pri prenose informácie

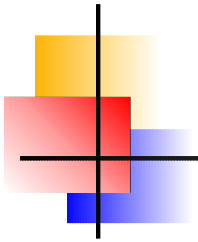
1. pokus – 1 000 000 hodov riadnej mince
   Na oznámenie výsledku potrebujeme 1 000 000 bitov

2. pokus – 1 000 000 hodov falošnej mince
   Padne 999 000 x líc a 1000 x rub
   Na oznámenie výsledku stačí udať poradové čísla hodov,
   v ktorých padol rub – 1000 20-bitových čísel
   – treba vyslať 20 000 bitov



Obsah predmetu Teória informácie

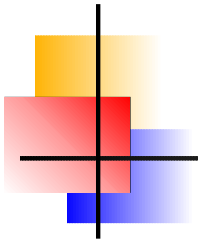
- Pojem množstva informácie a jej meranie
- Entropia pokusu
- Zdroje informácie
- Kódovanie
  - prispôsobenie abecedy zdroja abecede prenosového kanála
  - Kompresia dát
  - Detekcia vzniku chyby pri prenose
  - Schopnosť opraviť chybu pri prenose
- Prenosové kanály a ich kapacita



Príklady informácie

Informácia – skutočnosť (vec, fakt, udalosť, čosi) o ktorej keď sa dozvieme, vieme viac.
Vzniká problém merania množstva informácie
Príklady

- Odchod vlaku o 19:35 – Odchod vlaku poobede
- Skúšku som urobil za B – Skúšku som urobil
- Vonku je - 8°C – Vonku mrzne



Zavedenie informácie

- Where is wisdom?
- Lost in knowledge
- Where is knowledge?
- Lost in information
- Where is information
- Lost in data

T.S.Eliot

Čo je nositeľom informácie

- Veta – ale nie každá
- Výrok
- Výrok typu: Nastal jav A

Informáciu budeme priraďovajť javom - systému podmnožín nejakej univerzálnej množiny Ω.

Definícia

Nech Ω je neprázdna množina, ktorú budeme volať aj **základný priestor**. *σ-***algebrou** *podmnožín základného priestoru Ω nazývame taký systém $\mathcal{A}$ podmnožín množiny Ω, pre ktorý platí:*

1. $\Omega \in \mathcal{A}$
2. *Ak* $A \in \mathcal{A}$ *potom aj* $A^C = (\Omega - A) \in \mathcal{A}$
3. *Ak* $A_n \in \mathcal{A}$ *pre* $n = 1, 2, \ldots,$ *potom aj* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.$

Analógia elementárneho postupu zavedenia pravdepodobnosti

Predpokladáme, že základný priestor Ω je zjednotením konečného počtu n rovnako pravdepodobných disjunktných javov:

$$\Omega = A_1 \cup A_2 \cup \cdots \cup A_n \ .$$

Pretože $1 = P(\Omega) = P(A_1) + P(A_2) + \cdots + P(A_n)$, pravdepodobnosť každého z nich sa musí rovnať $\frac{1}{n}$
– t. j. $P(A_i) = \frac{1}{n}$ pre každé $i = 1, 2, \ldots, n$.

Za prvky σ-algebry $\mathcal{A}$ berieme $\emptyset$ a všetky konečné zjednotenia typu

$$A = \bigcup_{k=1}^{m} A_{i_k}, \tag{1}$$

kde $A_{i_k} \neq A_{i_l}$ pre $k \neq l$.
Potom každej množine $A \in \mathcal{A}$ tvaru (1) priradíme pravdepodobnosť

$$P(A) = \frac{m}{n}.$$

Analógia elementárneho postupu zavedenia pravdepodobnosti

Na to ale potrebujeme záhadnú binárnu operáciu $\oplus$, ktorá vypočíta informáciu disjunktného zjednotenia množín

$$I(A \cup B) = I(A) \oplus I(B) \tag{2}$$

Očakávame od informácie

1. $I(A) \geq 0$ pre všetky $A \in \mathcal{A}$ (3)
2. $I(\Omega) = 0$ (4)
3. Ak $A \in \mathcal{A},\ B \in \mathcal{A},\ A \cap B = \emptyset,$ potom $I(A \cup B) = I(A) \oplus I(B)$ (5)
4. Ak $A_n \nearrow A = \bigcup_{i=1}^{\infty} A_i,$ alebo $A_n \searrow A = \bigcap_{i=1}^{\infty} A_i,$ potom $I(A_n) \to I(A).$ (6)

Definícia

Hovoríme, že javy A, B sú **informačne nezávislé**, *ak platí*

$$I(A \cap B) = I(A) + I(B)\ . \tag{7}$$

Vlastnosti neznámej operácie $\oplus$

Teraz zosumarizujeme vlastnosti binárnej operácie $\oplus$

1. $x \oplus y = y \oplus x$ (8)
2. $(x \oplus y) \oplus z = x \oplus (y \oplus z)$ (9)
3. $I(A) \oplus I(A^C) = 0$ (10)
4. $\oplus : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$ je spojitá funkcia dvoch premenných (11)
5. $(x + z) \oplus (y + z) = (x \oplus y) + z$ (12)

Vlastnosti 1. a 2. vyplývajú z komutativity a asociativity množinového zjednotenia.
Vlastnosť 3. možno odvodiť z požiadavky $I(\Omega) = 0$ nasledujúcou postupnosťou rovností

$$0 = I(\Omega) = I(A \cup A^C) = I(A) \oplus I(A^C)$$

Vlastnosť 4. – spojitosť je prirodzená požiadavka vyplývajúca zo 4. požiadavky na spojitosť informácie I.

Vlastnosti neznámej operácie $\oplus$

Odôvodnenie požiadavky 5.
Majme dva disjunktné javy A, B také, že
A je nezávislé od C a tiež B je nezávislé od C.

Ak z toho, že nastal jav A, sa nič nedozvieme o jave C a ani z toho, že nastal jav B, sa nič nedozvieme o jave C, potom ani z toho, že nastal jav $A \cup B$, nedostaneme žiadnu informáciu o jave C, a teda javy $A \cup B$ a jav C sú nezávislé.
Označme $x = I(A)$, $y = I(B)$, $z = I(C)$ a počítajme informáciu
$I\left[(A \cup B) \cap C\right]$

$$I\left[(A \cup B) \cap C\right] = I(A \cup B) + I(C) = I(A) \oplus I(B) + I(C) = x \oplus y + z \quad (13)$$

$$\begin{aligned} I\left[(A \cup B) \cap C\right] = I\left[(A \cap C) \cup (B) \cap C\right] = I(A \cap C) \oplus I(B \cap C) = \\ = \left[I(A) + I(C)\right] \oplus \left[I(B) + I(C)\right] = (x + z) \oplus (y + z) \end{aligned} \quad (14)$$

Porovnaním vzťahov (13), (14) dostaneme žiadanú vlastnosť 5.

Vlastnosti neznámej operácie $\oplus$

Veta

Nech binárna operácia $\oplus$ na množine $\mathbb{R}_0^+$ vyhovuje axiómam (8) až (12). Potom

$$\text{buď} \quad \forall x, y \in \mathbb{R}_0^+ \qquad x \oplus y = \min\{x, y\}, \tag{15}$$

$$\text{alebo} \quad \exists k > 0 \; \forall x, y \in \mathbb{R}_0^+ \qquad x \oplus y = -k \log_2 \left(2^{-\frac{x}{k}} + 2^{-\frac{y}{k}}\right). \tag{16}$$

Veta

Nech $x \oplus y = -k \log_2 \left(2^{-\frac{x}{k}} + 2^{-\frac{y}{k}}\right)$ pre všetky nezáporné reálne x, y. Nech $x_1, x_2, \ldots, x_n$ sú nezáporné reálne čísla. Potom

$$\bigoplus_{i=1}^{n} x_i = x_1 \oplus x_2 \oplus \cdots \oplus x_n = -k \log_2 \left(2^{-\frac{x_1}{k}} + 2^{-\frac{x_2}{k}} + \cdots + 2^{-\frac{x_n}{k}}\right) \tag{17}$$

Elementárna definícia informácie

Nech $\{A_1, A_2, \ldots, A_n\}$ je rozklad priestoru Ω na javy s rovnakou informáciou, t. j. nech

$$1. \quad \Omega = \bigcup_{i=1}^{n} A_i, \text{ kde } A_i \cap A_j = \emptyset \text{ pre } i \neq j \tag{18}$$

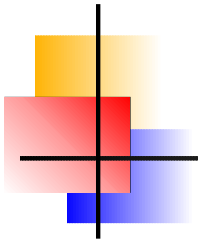
$$2. \quad I(A_1) = I(A_2) = \cdots = I(A_n) = a \tag{19}$$

Chceme určiť veličinu a. Z (18), (19) vyplýva

$$0 = I(\Omega) = I(A_1) \oplus I(A_2) \oplus \cdots \oplus I(A_n) = \underbrace{a \oplus a \oplus \cdots \oplus a}_{n-\text{krát}} = \bigoplus_{i=1}^{n} a \tag{20}$$

$$0 = \bigoplus_{i=1}^{n} a =$$

$$= \begin{cases} \min\{a, a, \ldots, a\} = a & \text{ak } x \oplus y = \min\{x, y\} \\ -k \log_2 \left(2^{-a/k} + \cdots + 2^{-a/k}\right) & \text{ak } x \oplus y = -k \log_2 \left(2^{-x/k} + 2^{-y/k}\right) \end{cases} \tag{21}$$



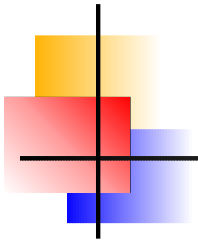
Elementárna definícia informácie

Pre prvý prípad $\bigoplus_{i=1}^{n} = a = 0$ a teda každý jav rozkladu $\{A_1, A_2, \ldots, A_n\}$ nesie so sebou nulovú informáciu. Toto je výsledok nezaujímavý a nemá význam sa ním ďalej zaoberať.
Pre druhý prípad

$$\bigoplus_{i=1}^{n} a = -k \log_2 \left(\underbrace{2^{-a/k} + \cdots + 2^{-a/k}}_{n\text{–krát}} \right) = -k \log_2 \left(n.2^{-a/k} \right) = a - k \log_2(n) = 0$$

Z posledného vzťahu vyplýva, že

$$a = k. \log_2(n) = -k. \log_2 \left(\frac{1}{n} \right) \tag{22}$$



Elementárna definícia informácie

Nech jav A je zjednotením m rôznych základných javov $A_{i_1}, A_{i_2}, \ldots, A_{i_m}$. Potom

$$
\begin{aligned}
I(A) &= I(A_{i_1}) \oplus I(A_{i_2}) \oplus \cdots \oplus I(A_{i_m}) = \underbrace{a \oplus a \oplus \cdots \oplus a}_{m\text{–krát}} = \\
&= -k.\log_2 \left(\underbrace{2^{-a/k} + 2^{-a/k} + \cdots + 2^{-a/k}}_{m\text{–krát}} \right) = -k \log_2 \left(m.2^{-a/k} \right) = \\
&= -k.\log_2(m) - k.\log_2 \left(2^{-a/k} \right) = -k.\log_2(m) - k.(-a/k) = \\
&= -k.\log_2(m) + a \ = -k.\log_2(m) + k.\log_2(n) = \\
&= k.\log_2 \left(\frac{n}{m} \right) = -k.\log_2 \left(\frac{m}{n} \right) \qquad (23)
\end{aligned}
$$

Veta

Nech $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$ *je rozklad priestoru* Ω *na javy s rovnakou informáciou. Potom pre informáciu* $I(A_i)$ *každého javu* A_i $i = 1, 2, \ldots, n$ *platí*

$$I(A_i) = -k \log_2 \frac{1}{n}. \tag{24}$$

Nech $A = A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_m}$ *je zjednotenie* m *rôznych množín rozkladu* $\mathbf{A}$, *t. j.* $A_{i_k} \in \mathbf{A}$, $A_{i_k} \neq A_{i_l}$ *pre* $k \neq l$. *Potom pre informáciu* $I(A)$ *javu* A *platí*

$$I(A) = -k \log_2 \frac{m}{n}. \tag{25}$$

Nech $\Omega = \{0, 1\}$ je množina hodnôt, ktoré môže nadobúdať jeden binárny znak, $A_1 = \{0\}$, $A_2 = \{1\}$, nech obe tieto jednoprvkové množiny nesú rovnakú informáciu a, ktorú by sme radi prehlásili za jednotkovú. Chceme, aby $I(A_1) = I(A_2) = a = 1$.
Podľa (22) $1 = a = k.\log_2(2) = k$.

Informácia ako funkcia pravdepodobnosti

Predpokladajme teda, že informácia $I(A)$ javu A závisí iba od pravdepodobnosti $P(A)$ javu A, t. j.

$$I(A) = f\left(P(A)\right),$$

pričom funkcia f nezávisí od toho, aký je pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$.
Ukážeme, že jedinou funkciou pripadajúcou do úvahy je funkcia $f(x) = -k.\log_2(x)$.

Definícia

Konečná alebo nekonečná postupnosť javov $\{A_n\}_n$ *sa nazýva* **postupnosťou (informačne) nezávislých javov**, *ak pre každú konečnú vybranú postupnosť* $A_{i_1}, A_{i_2}, \ldots, A_{i_m}$ *platí*

$$I\left(\bigcap_{k=1}^{m} A_{i_k}\right) = \sum_{k=1}^{m} I\left(A_{i_k}\right). \tag{26}$$

Informácia ako funkcia pravdepodobnosti

Aby informácia mala „rozumné“ vlastnosti, treba požadovať, aby funkcia f bola spojitá a aby javy nezávislé v pravdepodobnostnom zmysle ostali nezávislými v zmysle teórie informácie.

To znamená, že pre postupnosť nezávislých javov $A_1, A_2, \ldots, A_n$ platí

$$I(A_1 \cap A_2 \cap \cdots \cap A_n) = f(P(A_1 \cap A_2 \cap \cdots \cap A_n)) = f\left(\prod_{i-1}^{n} P(A_i)\right) \quad (27)$$

a súčasne

$$I(A_1 \cap A_2 \cap \cdots \cap A_n) = \sum_{i=1}^{n} I(A_i) = \sum_{i-1}^{n} f\left(P(A_i)\right) \quad (28)$$

Ľavé strany posledných dvoch vzťahov musia byť rovnaké, a preto

$$f\left(\prod_{i-1}^{n} P(A_i)\right) = \sum_{i-1}^{n} f\left(P(A_i)\right) \quad (29)$$

Informácia ako funkcia pravdepodobnosti

Nech sú všetky javy $A_1, A_2, \ldots, A_n$ rovnako pravdepodobné, nech $P(A_i) = x$. Potom $f(x^n) = n.f(x)$ pre všetky $x \in \langle 0, 1 \rangle$. Pre $x = 1/2$ máme

$$f(x^m) = f\left(\frac{1}{2^m}\right) = m.f\left(\frac{1}{2}\right). \tag{30}$$

Pre $x = \frac{1}{2^{1/n}}$ je $f(x^n) = f\left(\left(\frac{1}{2^{1/n}}\right)^n\right) = f\left(\frac{1}{2}\right) = n.f(x) = n.f\left(\frac{1}{2^{1/n}}\right)$,
z čoho máme

$$f\left(\frac{1}{2^{1/n}}\right) = \frac{1}{n}.f\left(\frac{1}{2}\right) \tag{31}$$

Konečne pre $x = \frac{1}{2^{1/n}}$ je

$$f(x^m) = f\left(\frac{1}{2^{m/n}}\right) = m.f(x) = m.f\left(\frac{1}{2^{1/n}}\right) = \frac{m}{n}.f\left(\frac{1}{2}\right),$$

takže

$$f\left(\frac{1}{2^{m/n}}\right) = \frac{m}{n}.f\left(\frac{1}{2}\right) \tag{32}$$

Informácia ako funkcia pravdepodobnosti

Pretože (32) platí pre všetky kladné čísla m, n a pretože predpokladáme, že funkcia f je spojitá, musí platiť

$$f\left(\frac{1}{2^x}\right) = x.f\left(\frac{1}{2}\right) \text{ pre všetky reálne čísla } x \in \langle 0, \infty).$$

Vytvorme pomocnú funkciu g predpisom $g(x) = f(x) + f\left(\frac{1}{2}\right).\log_2(x)$.

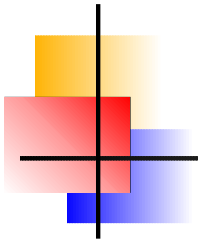
Potom platí

$$f(x) = f\left(2^{\log_2(x)}\right) = f\left(\frac{1}{2^{-\log_2(x)}}\right) = -\log_2(x).f\left(\frac{1}{2}\right)$$

$$f(x) = -f\left(\frac{1}{2}\right).\log_2(x) = -k.\log_2(x) \tag{33}$$

Pre množstvo informácie z posledného vzťahu vyplýva slávna
Shannonova – Hartleyova formula:

$$I(A) = -k.\log_2(P(A)) \tag{34}$$



Informácia ako funkcia pravdepodobnosti

Nech $\Omega = \{0, 1\}$ je množina hodnôt, ktoré môže nadobúdať jeden binárny znak, $A_1 = \{0\}$, $A_2 = \{1\}$, nech obe tieto jednoprvkové množiny majú rovnakú pravdepodobnosť $P(A_1) = P(A_2) = 1/2$.

Keďže veľkosť informácie je funkciou pravdepodobnosti, nesú obe množiny A_1, A_2 rovnakú informáciu a, ktorú by sme radi prehlásili za jednotkovú. Preto musí byť $1 = f\left(\frac{1}{2}\right) = -k.\log_2\left(\frac{1}{2}\right) = k$, čiže $k = 1$.

Shannonova – Hartleyova formula:

$$\boxed{I(A) = -\log_2(P(A))} \tag{35}$$