Informacija in kodi

UN2-1-AV 2024/2025

Diskretni viri informacije

Simon Dobrišek

oktober 2024

Teme predavanja

Uvod

Matematični model diskretnega vira informacije

Entropija stacionarnega vira

Ergodični stacionarni viri

Vir brez spomina

Entropija vira brez spomina

Vir s spominom (Markovov vir)

Entropija Markovovega vira

Odvečnost vira

Uvod

Vir informacije je podsistem komunikacijskega sistema, ki oddaja (pošilja) informacijo v kanal.

Teorija informacije obravnava le tiste vire, ki oddajajo informacijo zajeto v časovno in amplitudno (ali frekvenčno ali/in fazno) diskretnih signalih.

Če signali, ki nosijo informacijo vira, niso časovno in amplitudno (ali frekvenčno ali/in fazno) diskretni, jih v to obliko pretvorimo s postopkoma *vzorčenja* in *kvantizacije* signalov[1].

Pravimo jim ***diskretni viri informacije***.

Spoznali bomo, kako matematično modeliramo diskretni vir informacije.

[1] Glej literaturo, ki obravnava teorijo signalov, na primer: F. Mihelič, *Signali*, Založba FE in FRI, Ljubljana, 2006.

Uvod

Zamislimo si ***sistem, ki ga sestavljata človek in tipkovnica osebnega računalnika***.

Proces, ki ga opazujemo, je tipkanje človeka po tipkovnici med pisanjem besedila v slovenskem jeziku.

V določenem času natipka človek določen niz znakov (črk, števk, ločil,...), ki se shranijo v pomnilniku računalnika[2].

Takšen sistem lahko imamo za vir informacije, proces tipkanja po tipkovnici računalnika pa za proces ustvarjanja informacije.

[2] Pomnilnik računalnika je kanal, ki prevaja informacijo skozi čas.

Uvod

Vsak človek tipka po tipkovnici z določeno končno hitrostjo.

Če kot primer vzamemo človeka, ki vsako sekundo pritisne na nek znak tipkovnice, ustvari v n sekundah določen niz znakov dolžine[3] n.

Če označimo množico znakov, ki jih vsebuje tipkovnica računalnika, z A, njeno moč[4] pa z a, je niz znakov dolžine n element množice A^n z močjo a^n.

V n sekundah lahko človek ustvari vsak niz iz množice A^n tako, da prej natipkani znaki vplivajo na pozneje natipkane znake, oziroma tako, da ne vplivajo.

[3] Dolžino niza definiramo kot število znakov v nizu.

[4] Število elementov množice $A = \{x_1, x_2, ..., x_a\}$. Označimo ga z a ali $\mid A \mid$.

Uvod

Splošno velja, da če poznamo $n-1$ znakov v nizu, v katerem prej natipkani znaki vplivajo na pozneje natipkane znake, lahko bolj zanesljivo napovemo n-ti znak, kot v nizu, v katerem se znaki vrstijo neodvisno drug za drugim.

Na primer, če vemo, da so natipkani znaki *'Mera informacij'*, naslednji znak (črko *e*) lahko pravilno napovemo z veliko verjetnostjo.

Če pa poznamo enako dolg niz natipkanih znakov, ki so bili natipkani neodvisno en od drugega, na primer *'Na se rč nela e'*, je praktično nemogoče napovedati naslednji znak v nizu.

Uvod

Velja tudi, da so nekateri nizi znakov dolžine n bolj verjetni, drugi pa manj.

Na primer, gotovo je večja verjetnost, da človek, ki piše besedilo v slovenskem jeziku, natipka niz 16-ih znakov *'Mera informacije'* kot niz ravno tako 16-ih znakov *'Nera informacije'*.

Torej lahko strnemo: *matematično modeliranje diskretnih virov informacije temelji na opazovanju nizov znakov dolžine n.*

Matematični model diskretnega vira informacije

Diskretni vir informacije, ki oddaja znake iz končne neprazne množice znakov $A = \{x_1, x_2, ..., x_a\}$, $a \in \aleph$, ki ji pravimo tudi *abeceda vira*, matematično opišemo kot *naključni proces* z diskretnim parametrom (trenutek oddaje znaka) in diskretnimi vrednostmi (znaki, ki jih vir oddaja), to je z:

- nizom medsebojno odvisnih (ali neodvisnih) naključnih spremenljivk
$$\{X_t,\ t = 1, 2, ..., n\}$$
z zalogami vrednosti[5] $\mathcal{Z}(X_t) = A = \{x_1, x_2, ..., x_a\}$ in

[5] Za potrebe modeliranja virov z nizi diskretnih naključnih spremenljivk vzamemo, da so znaki iz abecede vira števila, za katere velja: $x_1 < x_2 < \cdots < x_a$.

Matematični model diskretnega vira informacije

- porazdelitvijo verjetnosti, da vir odda znak x_1 v trenutku $t = 1$, znak x_2 v trenutku $t = 2$, ..., znak x_n v trenutku $t = n$, to je

$$P(X_1 = x_1, ..., X_n = x_n) = P(x_1, ..., x_n) \geq 0, \quad (x_1, ..., x_n) \in A^n, \tag{1}$$

  kjer je A^n množica urejenih n-teric (nizov) znakov iz A, to je

$$A^n = \{(x_1, x_2, ..., x_n) : \ x_i \in A; \ i = 1, 2, ..., n\}.$$

Pravimo, da vir poznamo, če poznamo abecedo vira A in porazdelitev verjetnosti (1) za vsako naravno število $n \geq 1$.

Matematični model diskretnega vira informacije

DEFINICIJA 4.1 *Vir informacije je stacionaren, če za vsaki dve naravni števili n in k velja*

$$P(X_{k+1} = x_1, ..., X_{k+n} = x_n) = P(X_1 = x_1, ..., X_n = x_n). \quad (2)$$

Vidimo, da so verjetnostne lastnosti stacionarnega vira informacije neodvisne od časa, zato je verjetnost oddaje določenega niza znakov $x_1 x_2 ... x_n$ neodvisna od trenutka njegove oddaje.

Vsak stacionaren vir ima določeno nedoločenost H.

Entropija stacionarnega vira

Izhajajoč iz

$$P(X_1 = x_1, ..., X_n = x_n) = P(x_1, ..., x_n) \geq 0 \qquad x_1, ..., x_n \in A^n$$

in

$$P(X_{k+1} = x_1, ..., X_{k+n} = x_n) = P(X_1 = x_1, ..., X_n = x_n)$$

lahko za vsak $n \geq 1$ zapišemo *povprečno lastno informacijo* (entropijo) *znaka* v n-členem nizu kot

$$\boxed{H_n = \frac{1}{n} H(X_1, X_2, ..., X_n) = -\frac{1}{n} K \sum P(x_1, ..., x_n) \log_d P(x_1, ..., x_n) \quad (3)}$$

Entropija stacionarnega vira

in za vsak $n \geq 2$ *pogojno entropijo n-tega (zadnjega) znaka*, ko poznamo predhodnih $(n-1)$ znakov:

$$
\boxed{
\begin{aligned}
H'_n &= H(X_n|(X_1, \ldots, X_{n-1})) = \\
&= -K \sum P(x_1, \ldots, x_n) \log_d P(x_n|(x_1, \ldots, x_{n-1})), \qquad (4)
\end{aligned}
}
$$

kjer poteka seštevanje po vseh tistih $x_1, \ldots, x_n \in A^n$, za katere so vezane in pogojne verjetnosti večje od nič.

Entropija stacionarnega vira

Za diskreten stacionarni vir informacije s končno abecedo A bi lahko dokazali naslednji izrek:

IZREK 4.1

a) *Zaporedji* $H'_n (n = 2, 3, ...)$ *in* $H_n (n = 1, 2, ...)$ *sta konvergentni in*

b)
$$\lim_{n \to \infty} H'_n = \lim_{n \to \infty} H_n = H < \infty. \tag{5}$$

Dokaz: Glej učbenik *Informacija in kodi*, str. 45

Entropija stacionarnega vira

Količino $H \geq 0$ imenujemo *entropija diskretnega stacionarnega vira informacije*.

Lahko jo tolmačimo kot povprečno lastno informacijo znaka v nizu dolžine n (velja za velike n), ki ga odda stacionaren diskretni vir informacije, vendar le v primeru, ko je obravnavani stacionarni vir hkrati tudi *ergodičen*.

Ergodični stacionarni viri

Nizi znakov, ki jih oddajajo ergodični stacionarni viri, se statistično ne razlikujejo, zato lahko porazdelitev verjetnosti vira

$$P(X_1 = x_1, ..., X_n = x_n) = P(x_1, ..., x_n) \geq 0, \quad (x_1, ..., x_n) \in A^n,$$

za vsak $n \geq 1$ ocenimo iz samo enega dovolj dolgega niza oddanih znakov.

DEFINICIJA 4.2 *Stacionaren vir informacije je ergodičen, če velja za vsako naravno število $m < n$ in za vsak niz dolžine m znakov $\mathrm{y} = (b_1 ... b_m) \in A^m$, da relativna pogostost niza y v nizu dolžine n znakov x verjetnostno konvergira proti vrednosti $P(\mathrm{y})$, ko gre n proti neskončnosti.*

Pojasnitev pojma verjetnostne konvergence

Bernoullijev zakon velikih števil, ki pravi, da *z verjetnostjo, ki je poljubno blizu 1, lahko pričakujemo, da se bo pri zadosti velikem številu poskusov pogostost dogodka poljubno malo razlikovala od njegove verjetnosti*, lahko zapišemo kot

$$\lim_{n\to\infty} P\left(\left|\frac{k}{n} - p\right| < \delta\right) = 1, \tag{6}$$

kjer so: p verjetnost dogodka, k/n relativna pogostost dogodka v n ponovitvah poskusa in δ poljubno majhno pozitivno število.

Pojasnitev pojma verjetnostne konvergence

V izrazu (6) količina k/n ne konvergira proti p v smislu, da bi bil izraz $|\frac{k}{n} - p|$, od zadosti velikega n naprej, manjši od vnaprej predpisanega števila δ, temveč da je mogoče najti za poljubno majhni pozitivni števili δ in γ takšno naravno število n_0, da je pri $n \geq n_0$

$$P\left(\left|\frac{k}{n} - p\right| < \delta\right) > 1 - \gamma.$$

Tovrstni konvergenci pravimo *verjetnostna konvergenca*.

Primer stacionarnega vira, ki ni ergodičen.

Primer 4.1

Vzemimo stacinaren vir z abecedo $A = \{0, 1\}$, ki se z verjetnostjo $1/2$ nahaja v enem izmed dveh načinov delovanja:

- v prvem načinu lahko odda le niz dolžine n znakov $\mathrm{x}_0 = 00\cdots0$,
- v drugem pa lahko odda katerikoli drug niz iz A^n z enako verjetnostjo oddaje obeh znakov.

Če vzamemo $m = 1$ in $\mathrm{y} = 0$, je relativna pogostost niza $\mathbf{y}$ v oddanem nizu x_0 enaka 1 za vsak $n \in \aleph$.

Ker pa je verjetnost oddaje niza $\mathrm{y} = 0$ enaka 0,75, obravnavani stacionarni vir očitno ni tudi ergodičen. □

Ergodični stacionarni viri

Iz primera 4.1 sklepamo, da ergodični viri ne smejo imeti dva ali več načinov delovanja.

Dovolj dolge nize, ki jih oddajajo ergodični stacionarni viri, lahko razvrstimo v dve disjunktni množici:

- v množico zelo verjetnih – *značilnih* – nizov, ki imajo vsi približno enako verjetnost, Ξ_z, in
- v množico malo verjetnih – *neznačilnih* – nizov Ξ_n.

To lastnost ergodičnih virov imenujemo *asimptotična enakodelitvena lastnost (AEL)*[6].

Iz ergodičnih virov lahko pričakujemo le značilne nize, ker so neznačilni nizi malo verjetni. Značilni nizi so vsi približno enako verjetni in zato nosijo približno enako količino informacije.

[6] Angleški izraz: *Asymptotic Equipartition Property (AEP)*.

Ergodični stacionarni viri

Neergodični stacionarni viri nimajo asimptotične enakodelitvene lastnosti.

Primer 4.2

Vir iz primera 4.1 nima AEL, ker ne moremo sestaviti množice zelo verjetnih nizov, ki jih vir oddaja s približno enakimi verjetnostmi. Namreč, niz $x_0 = 00 \cdots 0$, ki je zelo verjeten, ima veliko večjo verjetnost kot katerikoli drugi niz. $\square$

Ergodični stacionarni viri

Za neergodične stacionarne vire informacije, količino H ne moremo tolmačiti kot povprečno informacijo, ki jo nosi znak v poljubnem dovolj dolgem nizu znakov.

Primer 4.3

Če vzamemo, da oddaja vir iz primera 4.1 v drugem načinu delovanja Bernoullijeve nize znakov[7], je verjetnost oddaje niza x dolžine n enaka

$$P(\mathrm{x}) = \begin{cases} \frac{1}{2} + \frac{1}{2^{n+1}}, & \text{če } \mathrm{x} = \mathrm{x}_0 \\ \frac{1}{2^{n+1}}, & \text{če } \mathrm{x} \neq \mathrm{x}_0. \end{cases}$$

(Ker je nizov $\mathrm{x} \neq \mathrm{x}_0$ natanko $2^n - 1$, velja $\frac{1}{2} + \frac{1}{2^{n+1}} + \frac{2^n-1}{2^{n+1}} = 1$.)

[7] Vir oddaja znake iz abecede $A = \{0, 1\}$ neodvisno enega od drugega. Znak 0 oddaja z verjetnostjo p, znak 1 pa z verjetnostjo $1 - p$.

Primer 4.3

Za dan vir je

- lastna informacija niza $\mathrm{x} = \mathrm{x}_0$, to je

$$I(\mathrm{x} = \mathrm{x}_0) = -\ln P(\mathrm{x} = \mathrm{x}_0) = -\ln\left(\frac{1}{2} + \frac{1}{2^{n+1}}\right) \quad \text{natov} \qquad (7)$$

- ter lastna informacija niza $\mathrm{x} \neq \mathrm{x}_0$, to je

$$I(\mathrm{x} \neq \mathrm{x}_0) = -\ln P(\mathrm{x} \neq \mathrm{x}_0) = (n+1)\ln 2 \quad \text{natov.} \qquad (8)$$

Iz (7) izhaja, da je lastna informacija, ki jo nosi znak v nizu $\mathrm{x} = \mathrm{x}_0$ ko narašča n čez vse meje, v povprečju enaka

$$\lim_{n\to\infty} \frac{I(\mathrm{x} = \mathrm{x}_0)}{n} = 0 \quad \text{natov,}$$

iz (8) pa, da je lastna informacija, ki jo nosi znak v nizu $\mathrm{x} \neq \mathrm{x}_0$ ko narašča n čez vse meje, v povprečju enaka

$$\lim_{n\to\infty} \frac{I(\mathrm{x} \neq \mathrm{x}_0)}{n} = \ln 2 \quad \text{natov.}$$

Primer 4.3

Polovica oddaje danega vira je torej sestavljena iz nizov, ki nosijo v povprečju 0 natov informacije na znak, polovica pa iz nizov, ki nosijo v povprečju $\ln 2$ natov na znak.

Izračunajmo sedaj entropijo neergodičnega vira iz primera 4.1. Da bi jo izračunali, izračunajmo najprej vezano entropijo n-terice diskretnih naključnih spremenljivk

$$\begin{aligned}
H(X_1, X_2, \ldots, X_n) &= -\sum P(x_1, \ldots, x_n) \ln P(x_1, \ldots, x_n) \\
&= -\left(\frac{1}{2} + \frac{1}{2^{n+1}}\right) \ln \left(\frac{1}{2} + \frac{1}{2^{n+1}}\right) \\
&\quad -\frac{2^n - 1}{2^{n+1}} \ln \left(\frac{1}{2^{n+1}}\right) \\
&= \cdots \approx \frac{n+2}{2} \ln 2 \quad \text{natov}^8,
\end{aligned}$$

[8] Za velike n zelo blizu prave vrednosti.

Primer 4.3

Nato pa še njeno limitno vrednost, to je entropijo vira H. Imamo

$$H = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, ..., X_n) = \frac{1}{2} \ln 2 \quad \text{natov.}$$

Ker sta $\lim_{n\to\infty} I(\mathrm{x} = \mathrm{x}_0)/n = 0$ natov in $\lim_{n\to\infty} I(\mathrm{x} \neq \mathrm{x}_0)/n = \ln 2$ natov, sledi da lastna informacija znaka $I(\mathrm{x})/n$ v nizih x, ki jih odda neergodičen vir iz primera 4.1, ne konvergira k entropiji vira H, ko dolžina nizov n narašča čez vse meje. $\square$

Ergodični stacionarni viri

V nadaljevanju se bomo bolj natančno seznanili s stacionarnimi ergodičnimi viri informacije.

Ugotovili bomo, da imata vira, ki ju bomo obravnavali v nadaljevanju, asimptotično enakodelitveno lastnost ter da lastna informacija znaka $I(\mathrm{x})/n = -K \log_d P(\mathrm{x})/n$ v nizu x dolžine n *verjetnostno konvergira* proti entropiji vira H, ko narašča dolžina niza čez vse meje.

To pomeni, da entropijo teh virov lahko tolmačimo kot povprečno informacijo, ki jo nosi znak v poljubnem dovolj dolgem nizu znakov, oddanem iz teh virov.

Vir brez spomina

DEFINICIJA 4.3 *Vir, za katerega je pogojna verjetnost*

$$P(x_n|(x_1, ..., x_{n-1})),$$

da v n-tem trenutku oddamo znak x_n*, glede na pogoj, da so prehodno oddani znaki* $x_1, ..., x_{n-1}$*, enaka brezpogojni verjetnosti* $P(x_n)$ *znaka* $x_n \in A$*, to je*

$$P(x_n|(x_1, ..., x_{n-1})) = P(x_n) \qquad n = 2, 3, ... \qquad x_n \in A,$$

je vir brez spomina.

Vir brez spomina

Če poznamo verjetnosti posameznih znakov

$$p_i = P(x_i)(i = 1, 2, ..., a)$$

iz A, lahko izračunamo verjetnost poljubnega n-členega niza kot

$$\boxed{P(x_1, x_2, ..., x_n) = P(x_1)P(x_2)\cdots P(x_n). \qquad (9)}$$

Entropija vira brez spomina

Entropijo znaka v n-členem nizu, ki ga odda diskretni stacionarni vir brez spomina, lahko zapišemo kot

$$H_n = \frac{1}{n}[H(X_1) + H(X_2) + \cdots + H(X_n)]. \tag{10}$$

Če vzamemo, da so naključne spremenljivke $X_1, X_2, \ldots$ enako porazdeljene, lahko zapišemo (10) kot

$$\begin{aligned} H_n &= \frac{1}{n} n H(X_1) = H(X_1) \\ &= -K \sum_{i=1}^{a} p_i \log_d p_i , \end{aligned}$$

kjer je a moč abecede A.

Entropija vira brez spomina

Ker H_n ni odvisen od n, je v tem primeru entropija vira (povprečna informacija, ki jo vsebuje znak, oddan iz vira) enaka

$$\boxed{H = -K \sum_{i=1}^{a} p_i \log_d p_i . \qquad (11)}$$

Značilni nizi znakov, ki jih oddaja vir brez spomina

Pokazali bomo, da lahko dovolj dolge nize znakov, ki jih odda vir brez spomina, razvrstimo v dve množici: v množico *zelo verjetnih* nizov ter v množico nizov z *verjetnostjo, ki je približno enaka nič*.

Primer 4.4 Vzemimo, da ima vir abecedo $A = \{0, 1\}$ z

verjetnostmi oddaje znakov

$$P(0) = p_1 = 1/3$$

in

$$P(1) = 1 - p_1 = 2/3$$

ter da vir oddaja znake neodvisno enega od drugega.

Značilni nizi znakov, ki jih oddaja vir brez spomina

Entropija vira je

$$H = H(p_1\,,\,1 - p_1) = -\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3} = 0{,}918 \quad \text{bitov.}$$

Če je v nizu dolžine n znakov m znakov 0, kjer je $m \leq n$, je verjetnost takšnega niza[9]

$$p_1^m(1 - p_1)^{n-m}. \tag{12}$$

Nizov z verjetnostjo (12) je $\binom{n}{m} = \frac{n!}{m!(n-m)!}$, zato je celotna verjetnost nizov dolžine n z m znakov 0 enaka

$$\binom{n}{m} p_1^m(1 - p_1)^{n-m}.$$

[9] Verjetnost niza $p_1^m(1-p_1)^{n-m}$ zapišemo tudi kot $2^{-n\cdot x}$, kjer je $x = -[m\log_2 p_1 + (n-m)\log_2(1-p_1)]/n$.

Značilni nizi znakov, ki jih oddaja vir brez spomina

m	Število nizov $\binom{n}{m}$	Verjetnost niza $p_1^m(1-p_1)^{n-m}$	Celotna verjetnost $\binom{n}{m}p_1^m(1-p_1)^{n-m}$
0	1	$2^{-15\cdot 0,585}$	0,002284
1	15	$2^{-15\cdot 0,652}$	0,017127
2	105	$2^{-15\cdot 0,718}$	0,059946
3	455	$2^{-15\cdot 0,785}$	0,129883
4	1365	$2^{-15\cdot 0,852}$	0,194825
5	3003	$2^{-15\cdot 0,918}$	0,214307
6	5005	$2^{-15\cdot 0,985}$	0,178589
7	6435	$2^{-15\cdot 1,052}$	0,114807
8	6435	$2^{-15\cdot 1,118}$	0,057404
9	5005	$2^{-15\cdot 1,185}$	0,022324
10	3003	$2^{-15\cdot 1,252}$	0,006697
11	1365	$2^{-15\cdot 1,318}$	0,001522
12	455	$2^{-15\cdot 1,385}$	0,000254
13	105	$2^{-15\cdot 1,452}$	0,000029
14	15	$2^{-15\cdot 1,518}$	0,000002
15	1	$2^{-15\cdot 1,585}$	0,000000

Tabela: Verjetnosti nizov znakov iz abecede A dolžine $n = 15$.

Značilni nizi znakov, ki jih oddaja vir brez spomina

Opazimo, da so najbolj verjetni nizi, ki imajo m blizu vrednosti $np_1 = 15 \cdot (1/3) = 5$.

Verjetnost nizov z $2 \leq m \leq 8$ je 0,95, zato lahko rečemo, da je verjetnost niza z m znakov 0, kjer se m pomembno razlikuje od np_1, zelo majhna.

Lahko tudi opazimo, da so verjetnosti nizov s številom ničel v nizu blizu 5 med $2^{-15 \cdot 0{,}718}$ in $2^{-15 \cdot 1{,}18}$, kar je zelo blizu $2^{-nH(p_1, 1-p_1)} = 2^{-15 \cdot 0{,}918}$.

Zato lahko rečemo, da so vsi zelo verjetni nizi - značilni nizi - skoraj enakoverjetni z verjetnostjo blizu $2^{-nH(p_1, 1-p_1)}$.

Značilni nizi znakov, ki jih oddaja vir brez spomina

Končno lahko tudi opazimo, da je število nizov z m med 2 in 8 enako $22\,803 = 2^{15 \cdot 0{,}965}$, kar je blizu $2^{nH(p_1, 1-p_1)}$.

Lahko torej rečemo, da je celotno število zelo verjetnih nizov blizu $2^{nH(p_1, 1-p_1)}$.

Če rezultate primera 4.4 posplošimo, vidimo, da so značilni nizi tisti nizi dolžine n, ki vsebujejo $n_1 = np_1$ znakov x_1, $n_2 = np_2$ znakov x_2 in tako naprej, kjer je $n = \sum_i n_i$. Verjetnost množice značilnih nizov je blizu ena. Neznačilni nizi pa so tisti nizi dolžine n, ki ne vsebujejo $n_1 = np_1$ znakov x_1, $n_2 = np_2$ znakov x_2 in tako naprej, verjetnost množice neznačilnih nizov pa je blizu nič.

$\square$

Značilni nizi znakov, ki jih oddaja vir brez spomina

Oglejmo si sedaj, kaj se zgodi, ko teži dolžina nizov n proti neskončnosti.

IZREK 4.2 AEL za vir brez spomina. *Če so naključne spremenljivke $X_1, \ldots, X_n$ medsebojno neodvisne in enako porazdeljene, lastna informacija po znaku niza $(1/n)\, I_n(\mathrm{x})$ verjetnostno konvergira, ko gre n proti neskončnosti, k entropiji vira brez spomina H :*

$$\frac{1}{n} I_n(\mathrm{x}) \xrightarrow{P} H. \tag{13}$$

Dokaz: Glej učbenik *Informacija in kodi*, str. 52

Značilni nizi znakov, ki jih oddaja vir brez spomina

TRDITEV 4.1 Verjetnost značilnih nizov. *Nizi* $\mathrm{x} \in A^n$ *z verjetnostjo oddaje*

$$d^{\frac{-n(H+\delta)}{K}} \leq P(\mathrm{x}) \leq d^{\frac{-n(H-\delta)}{K}} \tag{14}$$

tvorijo podmnožico značilnih nizov Ξ_z *vira brez spomina. Pri tem so:* H *entropija vira brez spomina,* $K > 0$ *poljubna konstanta,* $d > 1$ *osnova logaritma in* $\delta > 0$ *poljubno majhno število.*

Dokaz: Glej učbenik *Informacija in kodi*, str. 53

Značilni nizi znakov, ki jih oddaja vir brez spomina

Vsak element množice Ξ_z je *značilen* niz znakov, ki ga odda vir z verjetnostjo približno $d^{\frac{-nH}{K}}$.

Niz znakov iz množice Ξ_n pa je *neznačilen* niz znakov, ki ga odda vir z verjetnostjo približno nič.

TRDITEV 4.2 Moč množice značilnih nizov $|\Xi_z|$. *V množici značilnih nizov je največ $d^{\frac{n(H+\delta)}{K}}$ nizov oziroma*

$$| \Xi_z | \leq d^{\frac{n(H+\delta)}{K}}. \tag{15}$$

Pri tem so: H entropija vira brez spomina, $K > 0$ poljubna konstanta, $d > 1$ osnova logaritma in $\delta > 0$ poljubno majhno število.

Dokaz: Glej učbenik *Informacija in kodi*, str. 54

Značilni nizi znakov, ki jih oddaja vir brez spomina

Število značilnih nizov vira brez spomina je navzgor omejeno z $d^{\frac{n(H+\delta)}{K}}$, kar je manj ali kvečjemu enako a^n.

Navadno je znatno manj kot a^n, enako pa je le v primeru, da velja $p_i = P(x_i) = 1/a$ $(i = 1, 2, ..., a)$.

Vir s spominom (Markovov vir)

Če je oddaja znaka v sedanjem trenutku statistično odvisna od določenega števila ali vseh predhodno oddanih znakov, imenujemo takšen vir *diskretni vir s spominom*.

DEFINICIJA 4.4 Markovov vir. *Za Markovov vir je verjetnost oddaje znaka v n-tem trenutku odvisna le od znaka, ki je oddan v trenutku* $(n-1)$*, to je*

$$P(X_n = x_j | (X_1 = x_k, ..., X_{n-1} = x_i)) = P(X_n = x_j | X_{n-1} = x_i), \tag{16}$$

kar velja za poljubne $x_i, x_j, x_k \in A$ *ter* $n = 2, 3, ...$

Vir s spominom (Markovov vir)

Pogojni verjetnosti, da odda vir v trenutku n znak $x_j \in A$ pri pogoju, da je v trenutku $(n-1)$ oddan znak $x_i \in A$,

$$q_{ij} = P(X_n = x_j | X_{n-1} = x_i) \geq 0 : \quad n \geq 2 \; i, j = 1, 2, ..., a \; \sum_{j=1}^{a} q_{ij} = 1,$$

pravimo *prehodna verjetnost*.

Vir s spominom (Markovov vir)

Verjetnost, da v trenutku n odda vir znak $x_j \in A$ ne glede na predhodno oddani znak, je

$$
\begin{aligned}
P(X_n = x_j) &= \sum_{i=1}^{a} P(X_n = x_j \, , \, X_{n-1} = x_i) \\
&= \sum_{i=1}^{a} P(X_{n-1} = x_i) P(X_n = x_j | X_{n-1} = x_i)
\end{aligned}
$$

oziroma

$$
\boxed{P(X_n = x_j) = \sum_{i=1}^{a} P(X_{n-1} = x_i) q_{ij} \qquad \text{za} \;\; j = 1, \ldots, a. \qquad (17)}
$$

Vir s spominom (Markovov vir)

Enačbe (17) lahko zapišemo v matrični obliki kot

$$\boxed{\mathrm{p}_n = \mathrm{p}_{n-1}\mathrm{P}_Q, \qquad (18)}$$

kjer so:

$$\mathrm{p}_n = \big(P(X_n = x_1), P(X_n = x_2), ..., P(X_n = x_a)\big)$$

porazdelitev verjetnosti n-tega znaka v nizu,

$$\mathrm{P}_Q = \begin{bmatrix} q_{11} & \cdots & q_{1a} \\ \vdots & & \vdots \\ q_{a1} & \cdots & q_{aa} \end{bmatrix}$$

matrika prehodnih verjetnosti homogenega Markovovega vira in

$$\mathrm{p}_{n-1} = \big(P(X_{n-1} = x_1), P(X_{n-1} = x_2), ..., P(X_{n-1} = x_a)\big)$$

porazdelitev verjetnosti $(n-1)$-ga znaka v nizu, ki ga odda Markovov vir.

Vir s spominom (Markovov vir)

Primer 4.5

Dan je homogen Markovov vir z abecedo $A = \{0, 1\}$ in matriko prehodnih verjetnosti

$$\mathrm{P}_Q = \left[\begin{array}{cc} 1/4 & 3/4 \\ 1/2 & 1/2 \end{array}\right].$$

(Znak 0 ima verjetnost $1/4$, da bo ponovljen, in verjetnost $3/4$, da mu bo sledil znak 1. Znak 1 ima verjetnost $1/2$, da bo ponovljen in da mu bo sledil znak 0.)

Vzemimo, da je vir v trenutku $n = 0$ oddal znak 0. Določi verjetnosti, da bo vir oddal znak 0 tudi v trenutkih $n = 1, 2$ in 3.

Vir s spominom (Markovov vir)

Vektor začetnih verjetnosti znakov 0 in 1 je $\mathrm{p}_0 = (1, 0)$, zato je

$$\mathrm{p}_1 = \mathrm{p}_0 \, \mathrm{P}_Q = (1, 0) \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix} = (1/4,\ 3/4)\,.$$

Verjetnost, da bo vir oddal znak 0 v trenutku $n = 1$, je $1/4 = 0{,}25$.
V trenutku $n = 2$ je

$$\begin{aligned} \mathrm{p}_2 &= \mathrm{p}_1 \, \mathrm{P}_Q \\ &= (1/4, 3/4) \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix} \\ &= (7/\,16,\ 9/\,16)\,. \end{aligned}$$

Verjetnost, da bo vir oddal znak 0 v trenutku $n = 2$, je
$7/16 \approx 0{,}44$.

Vir s spominom (Markovov vir)

V trenutku $n = 3$ je

$$\begin{aligned} \mathrm{p}_3 &= \mathrm{p}_2\,\mathrm{P}_Q \\ &= (7/16, 9/16) \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix} \\ &= (25/\,64,\ 39/\,64)\,. \end{aligned}$$

Verjetnost, da bo vir oddal znak 0 v trenutku $n = 3$, je $25/64 \approx 0{,}39$. (V tem trenutku je verjetnost, da bo vir oddal znak 1, enaka $39/64 \approx 0{,}61$.) $\square$

Vir s spominom (Markovov vir)

Za *stacionarne Markovove vire* porazdelitev znakov v nizu ni odvisna od trenutka oddaje, to je od indeksa n, zato

$$\boxed{\mathrm{p}_n = \mathrm{p}_{n-1} = \mathrm{p}}$$

ter

$$\boxed{\mathrm{p} = \mathrm{p}\,\mathrm{P}_Q. \qquad (19)}$$

Porazdelitvi $\mathrm{p} = (p_1, \ldots, p_a)$, ki naredi homogen Markovov vir stacionaren, pravimo *stacionarna porazdelitev vira*. Določimo jo iz matrične enačbe (19).

Vir s spominom (Markovov vir)

Primer 4.6

Določi stacionarno porazdelitev homogenega Markovovega vira, ki smo ga obravnavali v primeru 4.5.

Stacionarno porazdelitev vira p dobimo iz enačbe (19), ki se v našem primeru glasi

$$(p_1, p_2) = (p_1, p_2) \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix},$$

Vir s spominom (Markovov vir)

to je, iz sistema enačb[10]:

$$p_1 = \frac{1}{4}p_1 + \frac{1}{2}p_2,$$

$$p_2 = \frac{3}{4}p_1 + \frac{1}{2}p_2,$$

ki da rešitev $p_1 = 2/5$ in $p_2 = 3/5$. Stacionarna začetna

porazdelitev verjetnosti je torej

$$\mathrm{p} = (2/5, 3/5)\,,$$

kar pomeni, da je v vsakem trenutku verjetnost oddaje znaka 0 enaka $2/5 = 0{,}4$, verjetnost oddaje znaka 1 pa $3/5 = 0{,}6$. $\square$

[10] Lahko si pomagamo tudi z enačbo $\sum_i p_i = 1$.

Entropija Markovovega vira

Za vsak niz $x_1 x_2 ... x_n \in A^n$, ki ga odda Markovov vir, lahko zvezo

$$P(x_1, ..., x_n) = P(x_1)\, P(x_2 \mid x_1) \cdots P\big(x_n \mid (x_{n-1}, ..., x_1)\big)$$

zapišemo kot

$$P(x_1, ..., x_n) = P(x_1)\, P(x_2 \mid x_1) \cdots P(x_n \mid x_{n-1}),$$

zato iz (4) sledi

$$\begin{aligned} H'_n &= -K \sum_{x_1,...,x_n \in A^n} P(x_1) P(x_2 \mid x_1) \cdots P(x_n \mid x_{n-1}) \log_d P(x_n \mid x_{n-1}) \\ &= -\sum_{x_1,...,x_{n-1} \in A^{n-1}} P(x_1) P(x_2 \mid x_1) \cdots P(x_{n-1} \mid x_{n-2}) \cdot \\ &\quad \cdot K \sum_{x_n \in A} P(x_n \mid x_{n-1}) \log_d P(x_n \mid x_{n-1}). \end{aligned} \tag{20}$$

Entropija Markovovega vira

Ker je

$$\sum_{x_1,\ldots,x_{n-2}\in A^{n-2}} P(x_1)P(x_2 \mid x_1)\cdots P(x_{n-1} \mid x_{n-2}) = P(x_{n-1}),$$

izhaja iz (20)

$$H'_n = -\sum_{x_{n-1}\in A} P(x_{n-1})\, K \sum_{x_n\in A} P(x_n \mid x_{n-1}) \log_d P(x_n \mid x_{n-1}),$$

zaradi predpostavke o stacionarnosti vira pa

$$H'_n = -\sum_{x_1\in A} P(x_1)\, K \sum_{x_2\in A} P(x_2 \mid x_1) \log_d P(x_2 \mid x_1). \tag{21}$$

Entropija Markovovega vira

Iz (21) lahko razberemo, da H'_n ni odvisna od n ($n > 1$), zato je entropija Markovovega vira (povprečna informacija, ki jo vsebuje znak, oddan iz vira), enaka

$$\begin{aligned} H = \lim_{n\to\infty} H'_n &= -\sum_{i=1}^{a} p_i \, K \sum_{j=1}^{a} q_{ij} \log_d q_{ij} \\ &= \sum_{i=1}^{a} p_i H_i, \end{aligned} \tag{22}$$

kjer je $H_i = -K \sum_{j=1}^{a} q_{ij} \log_d q_{ij}$.

Entropijo stacionarnega Markovovega vira tako določata matrika prehodnih verjetnosti P_Q in stacionarna porazdelitev vira $\mathrm{p} = (p_1, ..., p_a)$.

Entropija Markovovega vira

Primer 4.7

Vzemimo, da ima stacionaren Markovov vir abecedo $A = \{0, 1\}$ in matriko prehodnih verjetnosti

$$\mathrm{P}_Q = \begin{bmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{bmatrix}.$$

(Vsak znak abecede ima verjetnost 1/4, da bo ponovljen, in verjetnost 3/4, da mu bo sledil drug znak.)

Iz izraza $H_i = -\sum_{j=1}^{2} q_{ij} \log_2 q_{ij}$ izračunamo entropijo posameznih znakov abecede:

- $\text{znak0} : H_1 = -\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4} = 0{,}811 \quad \textit{bitov}$
- $\text{znak1} : H_2 = H_1.$

Entropija Markovovega vira

Stacionarno porazdelitev vira $\mathrm{p} = (p_1, p_2)$ dobimo iz enačb (19), ki se v našem primeru glasijo

$$p_1 = \frac{1}{4} p_1 + \frac{3}{4} p_2,$$

$$p_2 = \frac{3}{4} p_1 + \frac{1}{4} p_2,$$

ter zveze $(p_1 + p_2 = 1)$. Rešitev je $p_1 = p_2 = 1/2$.
Entropija Markovovega vira informacije (22) je tako

$$H = \sum_{i=1}^{2} p_i H_i = 0{,}811 \quad \text{bitov.}$$

$\square$

Entropija Markovovega vira

Ker so besedila, napisana v naravnih jezikih, sestavljena iz nizov statistično odvisnih znakov, navadno vzamemo, da jih ustvarjajo stacionarni ergodični viri s spominom.

Oglejmo si poskus ustvarjanja slovenskih besedil z Markovovimi viri.

Primer 4.8

Vzemimo, da ima slovenska abeceda 26 znakov (25 črk in presledek).

Če najprej predpostavimo, da so vsi znaki enako verjetni, $p_i = 1/26, 1 \leq i \leq 26$, je značilni niz znakov, ki ga odda vir brez spomina z enako verjetnimi znaki, na primer

$$\text{DM ČAVKUDGFOZŠNŽ ZSDHLŽISCJ} \ldots .$$

Entropija Markovovega vira

Vidimo, da je takšen model slovenskega jezika neuporaben.

Če uporabimo vir brez spomina z ocenami verjetnosti posameznih znakov, ki jih vsebuje spodnja razpredelnica

znak	E	A	O	I	N	L	S	R	J
p_i	0,086	0,084	0,073	0,073	0,051	0,042	0,041	0,040	0,038
znak	T	V	K	D	P	M	Z	B	U
p_i	0,035	0,030	0,030	0,027	0,027	0,026	0,017	0,016	0,015
znak	G	Č	H	Š	C	Ž	F	presl.	
p_i	0,013	0,012	0,009	0,008	0,005	0,005	0,001	0,195	

dobimo

$$\text{NA SE RČ NELA E NLSBOTEI} \ldots.$$

Entropija Markovovega vira

Kljub temu, da je takšen model boljši, še vedno ne upošteva vzajemne odvisnosti med znaki v nizu.

To lahko najbolj preprosto upoštevamo s pogojnimi verjetnostmi dveh zaporednih znakov (bigramov).

Ocenjevanje pogojnih verjetnosti je časovno zahtevna naloga. Shannon je predlagal Monte Carlo metodo[11], ki poteka takole:

> *Naključno izberemo besedilo. Naključno izberemo prvo stran in na tej strani eno črko. Vzamemo, da je ta črka prvi znak, ki ga odda vir. Naključno izberemo naslednjo stran besedila in na tej strani poiščemo prvi oddani znak (črko). Drugi znak, ki ga odda vir, je črka, ki sledi prvi črki, itn.*

[11] V statističnih raziskavah se v primeru, ko so enote populacije težko dostopne, zatekamo k Monte Carlo metodam. Tedaj enote nadomeščamo z naključno generiranimi podatki.

Entropija Markovovega vira

Z uporabo tega pristopa dobimo na primer takšen približek slovenščine

PEŠESOLISPODO TEŽE J ZANA SEVZOSTI . . . .

Isti postopek lahko uporabimo za še boljšo aproksimacijo z izbiro črk iz knjige glede na *dve* predhodni črki.

Dobimo

POT RITEME JE ŠTERIH SEGAR SAMOŽO . . . .

Shannon je predlagal modeliranje angleškega jezika kot vira besed in ne črk. Namesto ocenjevanja frekvenc besed v angleških besedilih je Shannon predlagal *'naključno odpiranje knjige'*, ki smo ga opisali za črke.

Entropija Markovovega vira

Z upoštevanjem samo predhodne besede dobimo za primer slovenščine

JEZIK BREZ BITI VE VRSTI SE PESMIH POMENA KOT . . . ,

z upoštevanjem dveh predhodnih besed pa

JE KAKOR OTROK PRED KNJIGO TUDI NEKAJ MOŽ ZA UŠESI IN UDARIL FANTA PO RAMI . . . . □

Značilni nizi znakov, ki jih oddaja Markovov vir

Tudi za Markovov vir velja AEL, kar ima za posledico razvrstitev nizov $\mathrm{x} \in A^n$ v dve podmnožici: v podmnožico značilnih nizov z verjetnostjo oddaje približno $P(x_1)d^{-(n-1)H}$ ter v podmnožico neznačilnih nizov z verjetnostjo oddaje približno enako nič.

IZREK 4.3 (AEL za Markovov vir) *Če tvorijo naključne spremenljivke $X_1, \ldots, X_n$ Markovovo verigo*[12]*, lastna informacija po znaku niza $(1/n)\, I(\mathrm{x})$ verjetnostno konvergira k entropiji Markovovega vira H, ko narašča n čez vse meje:*

$$\frac{1}{n} I(\mathrm{x}) \xrightarrow{P} H.$$

Dokaz: Glej učbenik *Informacija in kodi*, Dodatek.

[12] Markovova veriga je niz soodvisnih naključnih spremenljivk, kjer sega odvisnost samo do sosednjih spremenljivk v nizu.

Odvečnost vira

Za vsak vir informacije lahko definiramo tudi mero O, ki ji pravimo *odvečnost* vira informacije.

DEFINICIJA 4.5 *Odvečnost vira informacije je*

$$O = 1 - \frac{\textit{dejanska entropija vira}}{\textit{največja možna entropija vira}}.$$

Za splošen, stacionaren vir informacije je odvečnost enaka

$$\begin{aligned} O = 1 - \frac{H}{K \log_d a} &= 1 - \frac{\lim\limits_{n\to\infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n)}{K \log_d a} \\ &= 1 - \frac{\lim\limits_{n\to\infty} H(X_n \mid (X_1, \ldots, X_{n-1}))}{K \log_d a}. \end{aligned}$$

Odvečnost vira

Za stacionaren vir brez spomina je odvečnost enaka

$$O = 1 - \frac{-\sum_{i=1}^{a} p_i \log_d p_i}{\log_d a},$$

za stacionaren Markovov vir pa

$$O = 1 - \frac{-\sum_{i=1}^{a} p_i \sum_{j=1}^{a} q_{ij} \log_d q_{ij}}{\log_d a}.$$

Zaloga vrednosti mere O je interval $[0, 1]$. Mera O nima enote in je zelo pogosto podana v odstotkih.

Odvečnost vira

Primer 4.9

Oglejmo si, kako je ocenjena odvečnost slovenskih leposlovnih besedil.

Navadno vzamemo, da so besedila, ki so napisana v nekem naravnem jeziku, ustvarjena v stacionarnem ergodičnem viru informacije.

Za slovenska leposlovna besedila so bile entropije znakov v n-členih nizih (v bitih na znak) ocenjene takole[13]:

n	1	2	3	4	5	6	7	8	9	10
H_n	4,46	4,00	3,67	3,39	3,15	2,94	2,75	2,57	2,40	2,23

[13] Entropije $H_n = (1/n)\,H(X_1, X_2, ..., X_n)$, $n = 1, ..., 10$, so izračunane na vzorcu slovenskih leposlovnih besedil, ki obsega 2 721 416 besed oziroma 16 784 110 znakov iz nabora 133-ih znakov.

Odvečnost vira

Vidimo, da povprečna lastna informacija znaka v n-členem nizu H_n z naraščanjem števila n monotono pada in za $n = 10$ velja ocena 2,23 bita po znaku.

Limitna vrednost, h kateri teži H_n (kakor tudi H'_n), je gotovo nižja kot 2,23, zato lahko rečemo, da 'nosi' znak v slovenskih leposlovnih besedilih v povprečju največ 2,23 bita informacije[14].

Če vzammo, da je dejanska entropija slovenskih leposlovnih besedil 2 bita po znaku, je odvečnost v slovenskih leposlovnih besedilih

$$
\begin{aligned}
O &= 1 - \frac{H_{dej}}{H_{maks}} \\
&\approx 1 - \frac{2}{\log_2 133} \approx \frac{5}{7} \approx 70\ \%.
\end{aligned}
$$

[14] Shannon je ocenil limitno vrednost entropije angleškega jezika na 2,14 bita/znak.

Odvečnost vira

Vendar iz tega podatka ne smemo sklepati, da bi človek znal obnoviti prvotno besedilo iz nekaj več kot ene same četrtine naključno skrajšanega besedila. Na primer, če naključno črtamo znake z verjetnostjo $0,5$, iz niza

$$\mathrm{INFORMACIJA\quad IN\quad KODI},$$

dobimo

$$\mathrm{NOMCIINOD},$$

kar bi bilo težko obnoviti, čeprav smo stavek skrajšali le na eno polovico. □

> Odvečnost virov lahko odpravimo, delno ali v celoti, le z gospodarnim kodiranjem vira, ki ga obravnavamo v naslednjem poglavju.

Odvečnost vira

Primer 4.10

Na plavalnih tekmovanjih v posameznih plavalnih disciplinah tekmuje navadno po osem plavalcev.

Vzemimo, da je verjetnost zmage plavalca v četrti progi 1/2, peti 1/4, tretji 1/8, šesti 1/16, prvi, drugi, sedmi in osmi pa 1/64.

Vzemimo naprej, da je naša naloga sporočati izide tekem.

Ker so plavalne tekme neodvisne ena od druge, lahko sporočanje izidov tekem obravnavamo kot vir informacije brez spomina.

Vir oddaja znake iz abecede vira $\{x_1, x_2, ..., x_8\}$, kjer je x_i $(i = 1, 2, ..., 8)$ sporočilo, da je zmagovalec tekme plavalec v i-ti progi.

Odvečnost vira

Povprečna informacija, ki jo nosi sporočilo o izidu tekme, to je dejanska entropija vira, je

$$
\begin{aligned}
H_{dej} &= -\sum_{i=1}^{8} p_i \log_2 p_i \\
&= -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{8}\log_2\frac{1}{8} - \frac{1}{16}\log_2\frac{1}{16} - 4\frac{1}{64}\log_2\frac{1}{64} \\
&= 2 \quad \text{bita},
\end{aligned}
$$

največja možna entropija vira pa

$$H_{maks} = \log_2 8 = 3 \quad \text{bitov}.$$

Sledi, da je odvečnost vira

$$O = 1 - \frac{H_{dej}}{H_{maks}} = 1 - \frac{2}{3} = \frac{1}{3} \approx 33\ \%.$$

Odvečnost vira

Ker imamo v abecedi vira osem različnih sporočil, jih navadno zapišemo (kodiramo) s tremi dvojiškimi znaki, na primer takole:

proga zmagovalca i	1	2	3	4	5	6	7	8
zapis sporočila	000	001	010	011	100	101	110	111
dolžina zapisa n_i	3	3	3	3	3	3	3	3

Tako bomo izide tekem sporočali v povprečju z

$$
\begin{aligned}
E\{n_i\} &= \sum_{i=1}^{8} p_i n_i \\
&= \left(\frac{1}{64}\cdot 3 + \frac{1}{64}\cdot 3 + \frac{1}{8}\cdot 3 + \frac{1}{2}\cdot 3 + \frac{1}{4}\cdot 3 + \frac{1}{16}\cdot 3 + \frac{1}{64}\cdot 3 + \frac{1}{64}\cdot 3\right) \\
&= 3 \quad \text{dvojiškimi znaki,}
\end{aligned}
$$

kar je preveč, saj sporočila v povprečju nosijo le 2 bita informacije.

Odvečnost vira

Če pa zapišemo sporočila o izidih tekem takole (gospodarno kodiranje):

proga zmagovalca i	1	2	3	4	5	6	7	8
zapis sporočila	111100	111101	110	0	10	1110	111110	111111
dolžina zapisa n_i	6	6	3	1	2	4	6	6

bodo izidi tekem sporočeni v povprečju z

$$
\begin{aligned}
E\{n_i\} &= \sum_{i=1}^{8} p_i n_i \\
&= \left(\frac{1}{64}\cdot 6 + \frac{1}{64}\cdot 6 + \frac{1}{8}\cdot 3 + \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{16}\cdot 4 + \frac{1}{64}\cdot 6 + \frac{1}{64}\cdot 6\right) \\
&= 2 \quad \text{dvojiškima znakoma.}
\end{aligned}
$$

Opazimo, da smo s 'primernim' zapisom sporočil, to je znakov iz abecede vira, uspeli odpraviti odvečnost vira. □

Vprašanja

- Kaj je vir informacije?
- Kako matematično opišemo diskretne vire informacije?
- Kateri diskretni viri informacije so stacionarni?
- Kakšno entropijo imajo diskretni stacionarni viri informacije?
- Kateri stacionarni diskretni viri informacije so ergodični?
- Kako tolmačimo entropijo stacionarnega ergodičnega diskretnega vira informacije?
- Kako izračunamo entropijo vira brez spomina?
- Kako izračunamo entropijo Markovovega vira informacije?