Teorija informacij in sistemov, predavanje 5

Uroš Lotrič

Univerza v Ljubljani,
Fakulteta za računalništvo in informatiko

4. Kanal 1

- Strukturo, ki opisuje medsebojno povezanost, imenujemo kanal

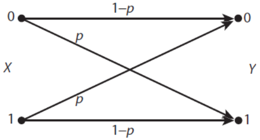
- Kanal prenaša informacijo o spremenljivki X do spremenljivke Y
- Kanal matematično opišemo s **pogojnimi verjetnostmi**, ki povezujejo izhodne verjetnosti z vhodom.
- Ta matematični model opisuje mnoge realne situacije
- Za enkrat se bomo omejili na diskretni kanal brez spomina.

4.1 Diskretni kanal brez spomina 1

- **Diskretni kanal brez spomina** povezuje diskretni naključni spremenljivki s končno množico stanj X in Y.
  - Naključni spremenljivki zavzameta končno mnogo stanj $X = \{x_1, \ldots, x_r\}$, $Y = \{y_1, \ldots, y_s\}$ z verjetnostmi $P_X = \{p(x_1), \ldots, p(x_r)\}$ in $P_Y = \{p(y_1), \ldots, p(y_s)\}$.
  - Kanal je definiran kot množica pogojnih verjetnosti $p(y_j|x_i)$. Pogojna verjetnost nam pove verjetnost za dogodek y_j na izhodu iz kanala, če je na vhodu v kanal dogodek x_i.
  - Brez spomina je zato, ker so pogojne verjetnosti $p(y_j|x_i)$ konstantne in torej neodvisne od predhodnih simbolov, $\sum_j p(y_j|x_i) = 1$
  - Kanal je popolnoma podan z $r \times s$ pogojnimi verjetnostmi (tudi **prehodnimi verjetnostmi**).
  - Slika kanala (nariši za $r = 4$, $s = 3$)

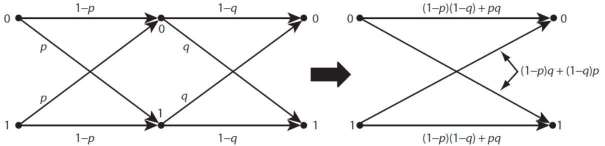
4.1 Diskretni kanal brez spomina 2

- Pomemben poseben primer je **binarni simetrični kanal**.
  - napaka kanala je p, saj se z verjetnostjo p znak prenese v napačnega



$$P_k = X \rightarrow \begin{array}{c} Y \downarrow \\ \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \end{array}$$

  - vhodni in izhodni znaki so 0, 1
- Kanal z notranjo strukturo z vmesnimi spremenljivkami (ojačevalnik v strojni opremi)
  - Verjetnost za napako do ojačevalnika je drugačna kot naprej



4.2 Pogojna entropija 1

- Imamo dve naključni spremenljivki X in Y.
- Pogojna entropija spremenljivke Y pri znanem X se zapiše kot $H(Y|X)$.
- Vzemimo, da se je zgodil dogodek $x_i \in X$. Entropija dogodka Y je potem

$$H(Y|x_i) = -\sum_{j=1}^{s} p(y_j|x_i) \log p(y_j|x_i)$$

- iz znanja o entropiji: $0 \leq H(Y|x_i)$
- Primer: mečemo kocko, dogodek X: veliko (5 ali 6), dogodek Y: točno število pik. H(Y|veliko) = 1 bit. H(Y|malo) = 2 bit.

4.2 Pogojna entropija 2

- Kaj če o dogodku X vemo le, da se je zgodil?
- Potem lahko sklepamo na povprečen izid - utežena vsota

$$\begin{aligned} H(Y|X) &= \sum_i p(x_i) H(Y|x_i) \quad , \quad 0 \leq H(Y|X) \\ &= -\sum_{i=1}^{r} \sum_{j=1}^{s} p(x_i, y_j) \log p(y_j|x_i) \end{aligned}$$

- uporabili smo **vezano verjetnost** dogodkov X in Y, $p(x_i, y_j) = p(y_j|x_i) p(x_i)$. (vezana = pogojna + lastna)
- Primer: kocka od prej: $H(Y|X) = \frac{2}{6} \cdot 1 + \frac{4}{6} \cdot 2 = \frac{5}{3} = 1.66$ Opazimo: $H(Y|X) < H(Y) = 2.58$ bit.
- Splošno: $0 \leq H(Y|X) \leq H(Y)$: če poznamo spremenljivko X, se nedoločenost Y ne more povečati (lahko pa se zmanjša). Spodnja omejitev sledi iz definicije (enačbe za $H(Y|x_i)$ in utežena vsota)
- Primer: diskretni binarni simetrični kanal:

4.3 Vezana entropija spremenljivk 1

- Vezana entropija naključnih spremenljivk X in Y je entropija para (X, Y).
  - $r \times s$ možnih parov, verjetnost vezanega dogodka $p(x_i, y_j)$
  - pomemebne zveze: $p(x_i, y_j) = p(y_j|x_i)p(x_i)$, $\sum_j p(x_i, y_j) = p(x_i)$, $\sum_i p(x_i, y_j) = p(y_j)$, $\sum_{i,j} p(x_i, y_j) = 1$
  - na enak način lahko povežemo tudi več spremenljivk
- Velja $H(X, Y) = H(Y|X) + H(X)$
  - razlaga: če najprej izvemo, kaj se je zgodilo v dogodku X in potem dobimo še dodatne informacije o dogodku Y, vemo vse.

4.3 Vezana entropija spremenljivk 2

- Primer: kocka: X = {visoko (5, 6), nizko (1,2,3,4)}, Y = {1, 2, 3, 4, 5, 6}
  - $H(Y|X) = 5/3 = 1.67$
  - $H(X) = -1/3\log(1/3) - 2/3\log(2/3) = \log 3 - 2/3 = 0.92$
  - $H(X,Y) = \log 3 + 1 = 2.59$
- Ali velja tudi obratno, $H(X,Y) = H(X|Y) + H(Y)$?
- Ker sta spremenljivki enakovredni, ni razloga, da ne.
- Primer s kocko še enkrat:
  - $H(X|Y) = 0$
  - $H(Y) = \log 6$
  - $H(X,Y) = \log 6 + 0 = 1 + \log 3$

4.4 Obrat kanala 1

- Glede na zgornjo ugotovitev lahko kanal tudi obrnemo
- Ne obračamo fizičnega procesa (če obstaja) ampak samo verjetnostno strukturo informacije, ki definira kanal
- Primer:
  - Oddajanja radijskih signalov ne moremo obrniti v vsrkavanje
  - Kanala, ki povezuje težo z višine, pa lahko obrnemo v kanal, ki povezuje višino s težo.
- Pogoj za obrat je, da poznamo vhodne verjetnosti. Iz njih lahko določimo izhodne verjetnosti, ki jih lahko uporabimo kot vhodne verjetnosti v obrnjeni kanal.

4.4 Obrat kanala 2

- Izračun izhodnih verjetnosti $p(y_j) = \sum_i p(y_j|x_i)p(x_i)$
- Potrebujemo še **obratne pogojne verjetnosti**
  - velja $p(x_i, y_j) = p(y_j|x_i)p(x_i) = p(x_i|y_j)p(y_j)$
  - Bayesovo pravilo:

$$p(x_i|y_j) = \frac{p(y_j|x_i)p(x_i)}{p(y_j)}$$

- Za sprejemnika sporočila so obratne pogojne verjetnosti zelo pomembne, saj z njimi lahko iz prejetih znakov določi verjetnosti za vhodne znake

4.4 Obrat kanala 3

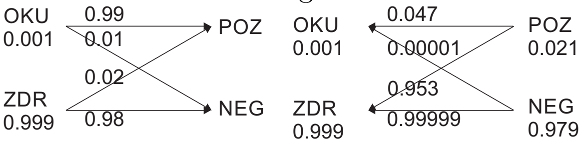
Primer: arena življenja in smrti

- redka bolezen napade vsakega tisočega v populaciji
- na srečo obstaja tester: okužbo zazna v 99% primerov, lažno okuženost pa ugotovi v 2% primerov
- Kakšna je verjetnost, da je oseba okužena, če je test pozitiven? $p(S|P)$?

4.4 Obrat kanala 4

Primer: okužba: arena življenja in smrti

- Slika kanala in obratnega kanala



- Račun: $p(O|P) = p(O,P)/p(P) = 0.047$
- Paradoks lažno pozitivnih testov: kljub temu, da je test zelo zanesljiv je manj kot 5% pozitivnih dejansko okuženih. Ob pozitivnem testu se je verjetnost za okužbo povečala iz 1/1000 na 1/21. Zdravnik naredi dodatne teste.

4.5 Medsebojna informacija 1

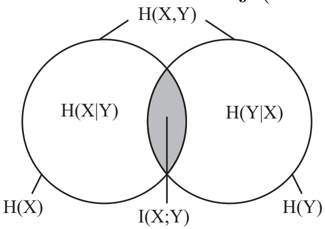
- Medsebojna informacija nam pove, koliko o eni spremenljivki izvemo iz druge spremenljivke
- $I(X;Y) = H(X) - H(X|Y)$
- Interpretacija
- Lastnosti z dokazi - 1. del
  1. $I(X;Y) = H(X,Y) - H(X|Y) - H(Y|X)$
  2. $I(X;Y) = H(X) - H(X|Y)$
  3. $I(X;Y) = H(Y) - H(Y|X)$

4.5 Medsebojna informacija 2

- Lastnosti z dokazi - 2. del
  4. $I(X;Y) = H(X) + H(Y) - H(X,Y)$
  5. $I(X;Y)$ je simetrična glede na X in Y [se vidi iz 4]
  6. $I(X;Y) = -\sum_i \sum_j p(x_i, y_j) \log \frac{p(x_i)p(y_j)}{p(x_i,y_j)}$
  7. $I(X;Y) \geq 0$
  8. $I(X;X) = H(X)$

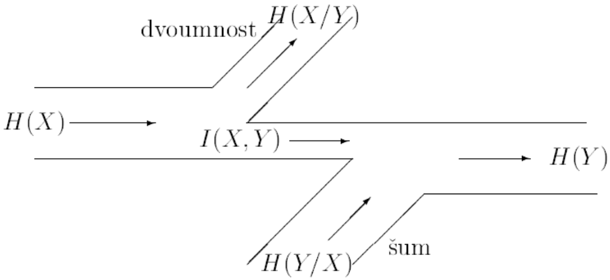
4.5 Medsebojna informacija 3

- Primer: kocka, $X = veliko(5,6), malo(1,...4)$, $Y = 1,...6$
  - $H(Y) = \log 6 = 2.59$
  - $H(Y|X) = 5/3 = 1.67$
  - $I(X;Y) = H(Y) - H(Y|X) = 0.92$
  - $I(X;Y) = H(X) - H(X|Y) = H(X) - 0 = 0.92$
- Iz medsebojne informacije lahko dokažemo še relacijo za zgornjo mejo v enačbi $0 \leq H(Y|X) \leq H(Y)$.
- Predstavitev relacij (Vennov diagram)



4.5 Medsebojna informacija 4

- Predstavitev relacij (kanal)



4.5 Medsebojna informacija 5

- Primer: Na nekem območju živi 25% blondink. Med blondinkami je 75% modrookih. Vemo še, da na tem območju živi 50% modrookih deklet.
  - B = {b, !b}, M = {m, !m}tabela

		$\rightarrow j$		
	B\M	m	!m	
$\downarrow i$	b	**3/4.1/4=3/16**	1/16	**1/4**
	!b	5/16	7/16	3/4
		1/2	1/2	

  - $I(B;M) = H(M) - H(M|B) = H(B) - H(B|M) = 0.06$