Zavod za telekomunikacije

Teorija informacije

Osnovni pojmovi teorije informacije

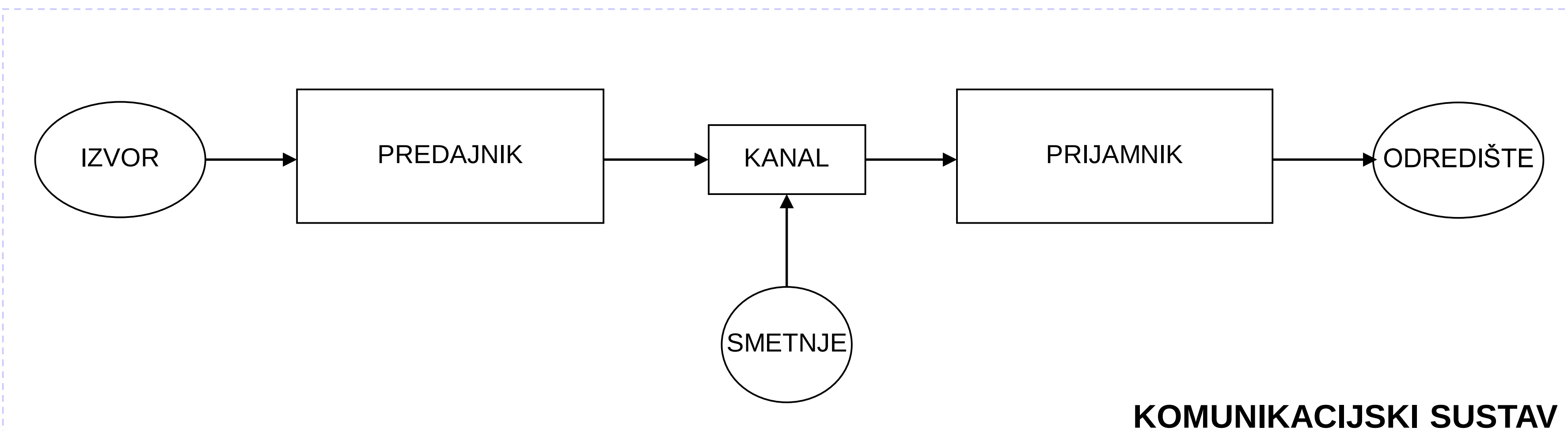
Osnovni pojmovi teorije informacije

- Opći model komunikacijskog sustava
  - Diskretni komunikacijski sustav
  - Poruka i prijenos poruke
- Sadržaj informacije, entropija
- Kodiranje
- Informacijski opis komunikacijskog sustava, informacijske mjere
- Kapacitet kanala
- Prijenos informacije komunikacijskim sustavom

DEFINICIJA

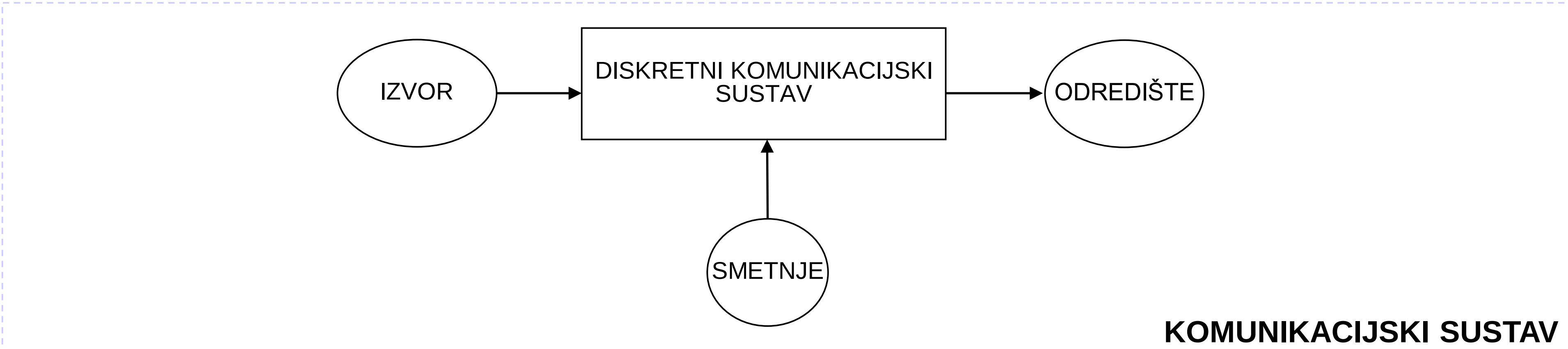
Opći model komunikacijskog sustava

Temeljni problem komunikacije je točno ili aproksimativno reproducirati u jednoj točki informacijskog prostora (odredište) poruku odabranu na nekoj drugoj točki (izvor) [Shannon 1948].



Diskretni komunikacijski sustav

- Jednostavniji slučaj – diskretni signali
- Ključna pitanja:
  - Što je poruka?
  - Što znači prenijeti poruku?
  - Koja je mjera za količinu informacije u nekoj poruci, te informacije prenesene sustavom?

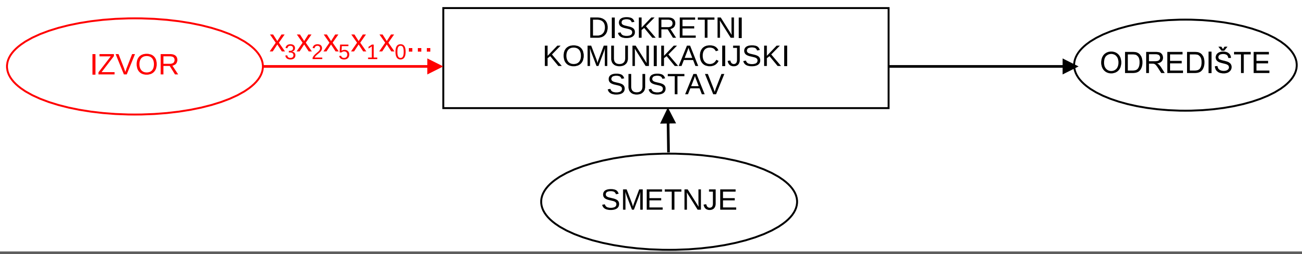


Poruka

- Niz simbola odabranih iz konačne abecede *X*
  - Abeceda je skup elementarnih simbola

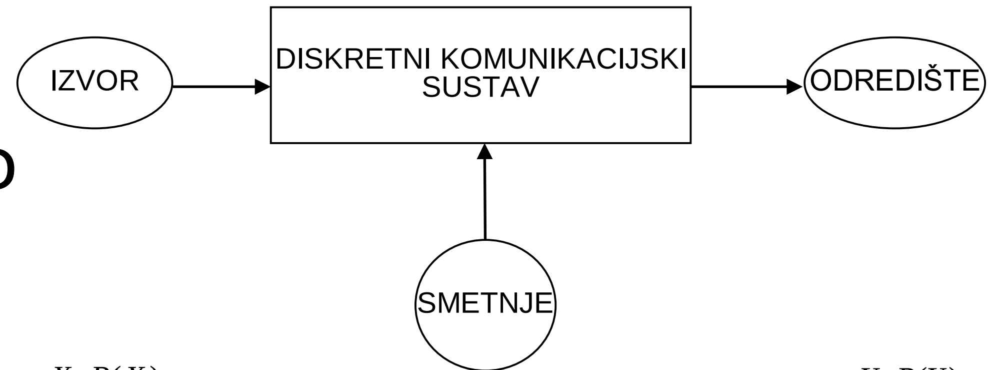
$$X = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$$

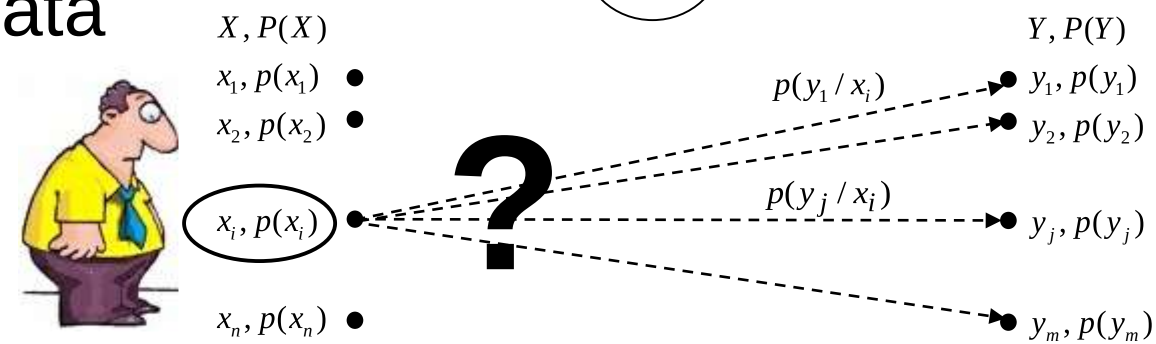
- Svaki simbol pri *N*-tom biranju ima vjerojatnost pojavljivanja: $x_i \longrightarrow p_N(x_i)$
- Pretpostavka (za sada): odabir simbola neovisan o prethodno odabranim simbolima: $x_i \longrightarrow p(x_i)$



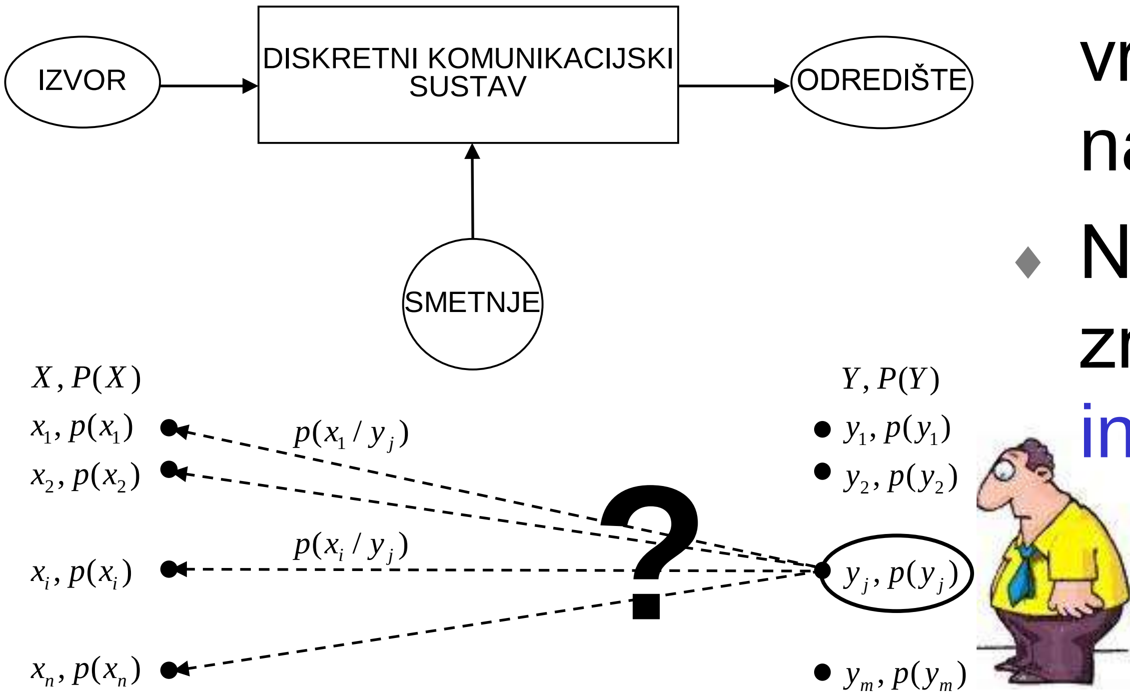
Prijenos poruke: pogled sa izvora

- Prijenos poruke = prijenos simbola
- Na izvoru odabran x_i: što se pojavi na odredištu?
- Pretpostavka: poznata statistička svojstva prijenosa





Prijenos poruke: pogled sa odredišta

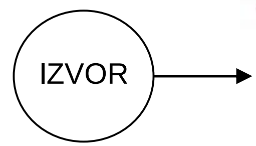


- Prije pojave y_i, znamo vrlo malo o događajima na izvoru
- Nakon opažanja y_i znamo više: primili smo informaciju!

Sadržaj informacije poruke - primjer

Zavod za telekomunikacije

- Koliko informacije možemo maksimalno prenijeti nekom porukom?
- Primjer: pismo ili glava



- Koliko informacije je primio promatrač?
- Što ako uvijek pada pismo?
- Što ako pismo pada 70% puta?

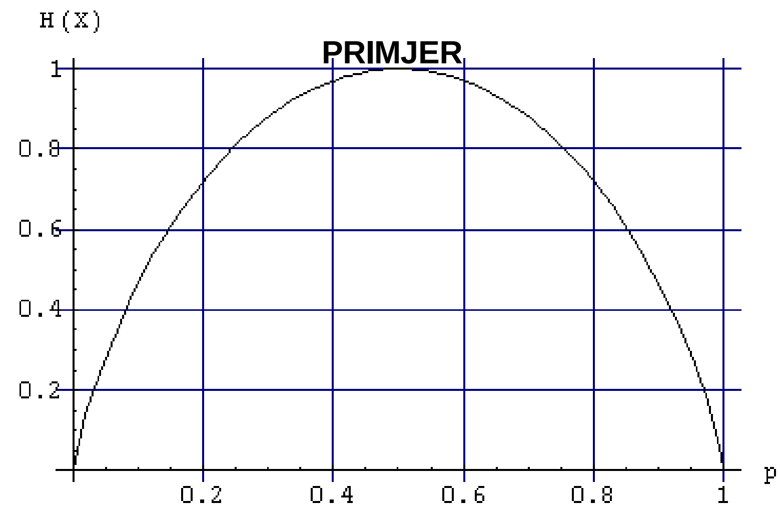
DEFINICIJA

Entropija

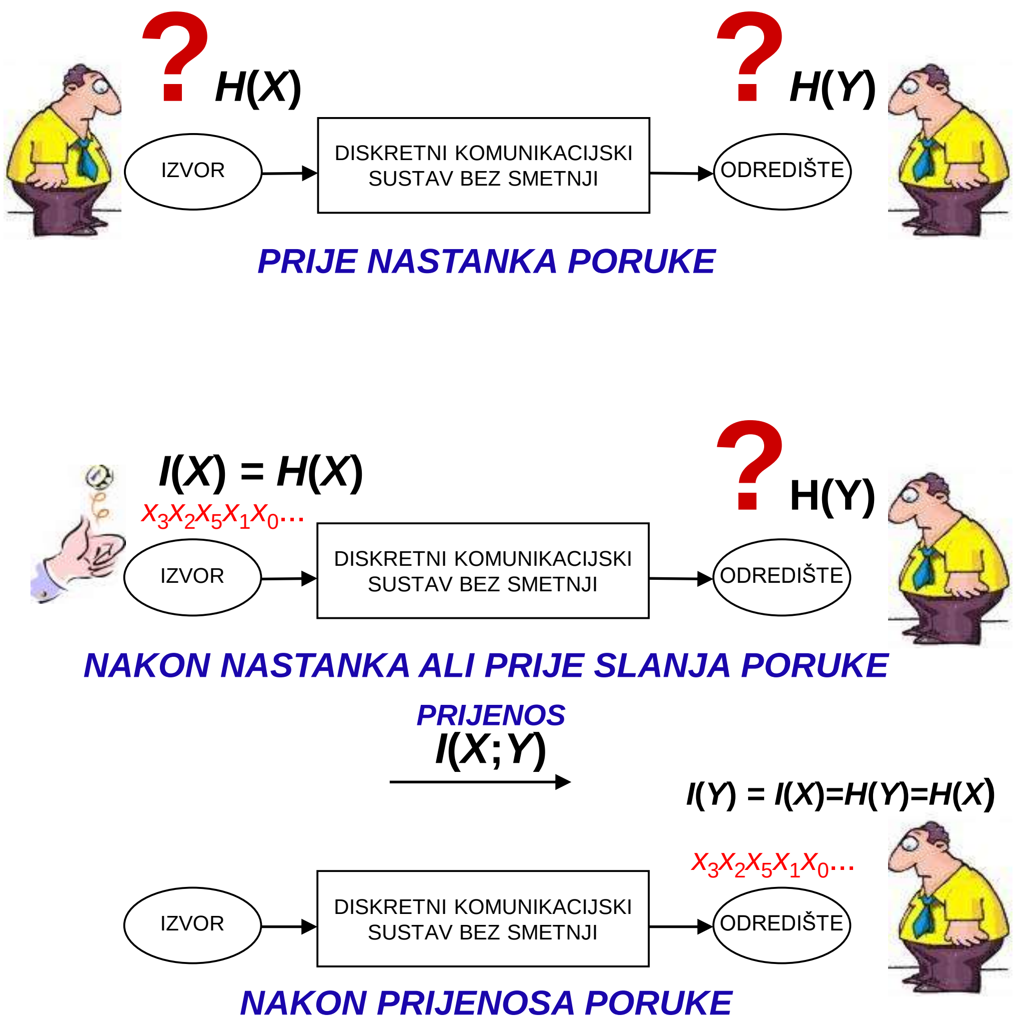
- Entropija diskretne slučajne varijable

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) [bit / simbol]$$

- Entropija daje mjeru za sadržaj informacije



Entropija, neodređenost, sadržaj informacije u sustavu bez smetnji



- Neodređenost = entropija
- Informacija na izvoru, neodređenost na odredištu
- Prijenosom poruke neodređenost je nestala

Svojstva entropije $H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i)$

- Sadržaj informacije ne može biti negativan $H(X) \geq 0$
- Sadržaj informacije je 0 ako se uvijek pojavljuje samo jedan simbol $H(X) = 0 \Leftrightarrow \exists i \mid p(x_i) = 1$
- Neodređenost i sadržaj informacije su maksimalni ako su vjerojatnosti simbola jednako raspoređene

$$H(X) \leq \log n$$

$$p(x_i) = \frac{1}{n} \Rightarrow H(X) = \log n$$

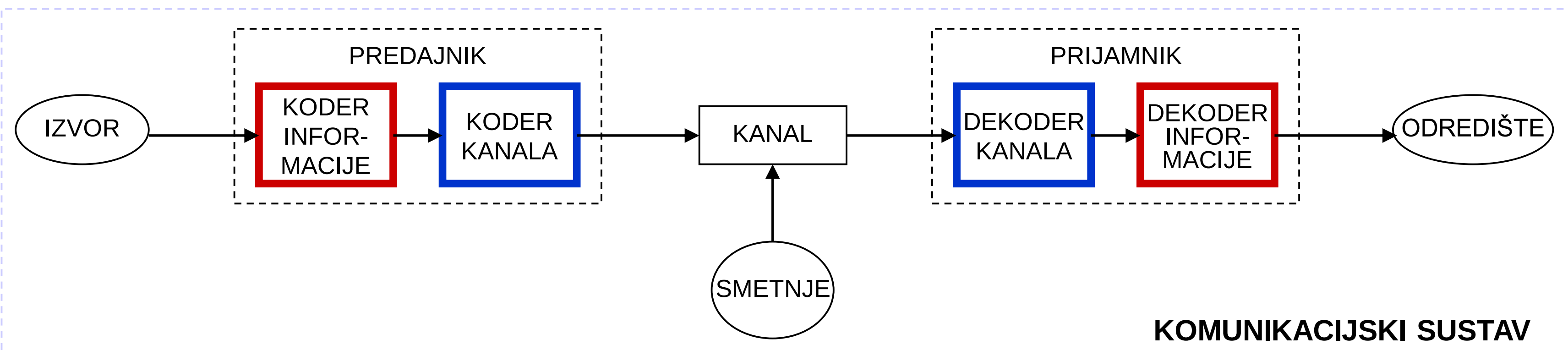
- Zašto baš logaritam? $H(XY) = H(X) + H(Y)$

Bit i binarna znamenka

- Teorija informacije: bit je osnovna jedinica informacije
- Ostatak svijeta: bit je binarna znamenka
- Bacamo “nepošteni” novčić, pismo=1, glava=0; koliko je ovo bitova: 1111111111 ?
- Kada znamo razliku, iz konteksta je jasno što se misli

Kodiranje

- Dodjela kodnih riječi simbolima poruke
- Poruka se “samo” pretvara u novi oblik (niz simbola)
- Zašto onda kodirati?
- U praksi, kodovi su binarni



Kodiranje i entropija

PRIMJER

SIMBOL (x_i)	VJEROJATNOST POJAVLJIVANJA $p(x_i) = p_i$	KODNA RIJEČ (C_i)	DULJINA KODNE RIJEČI (l_i)
1	1/2	0	1
2	1/4	10	2
3	1/8	110	3
4	1/8	111	3

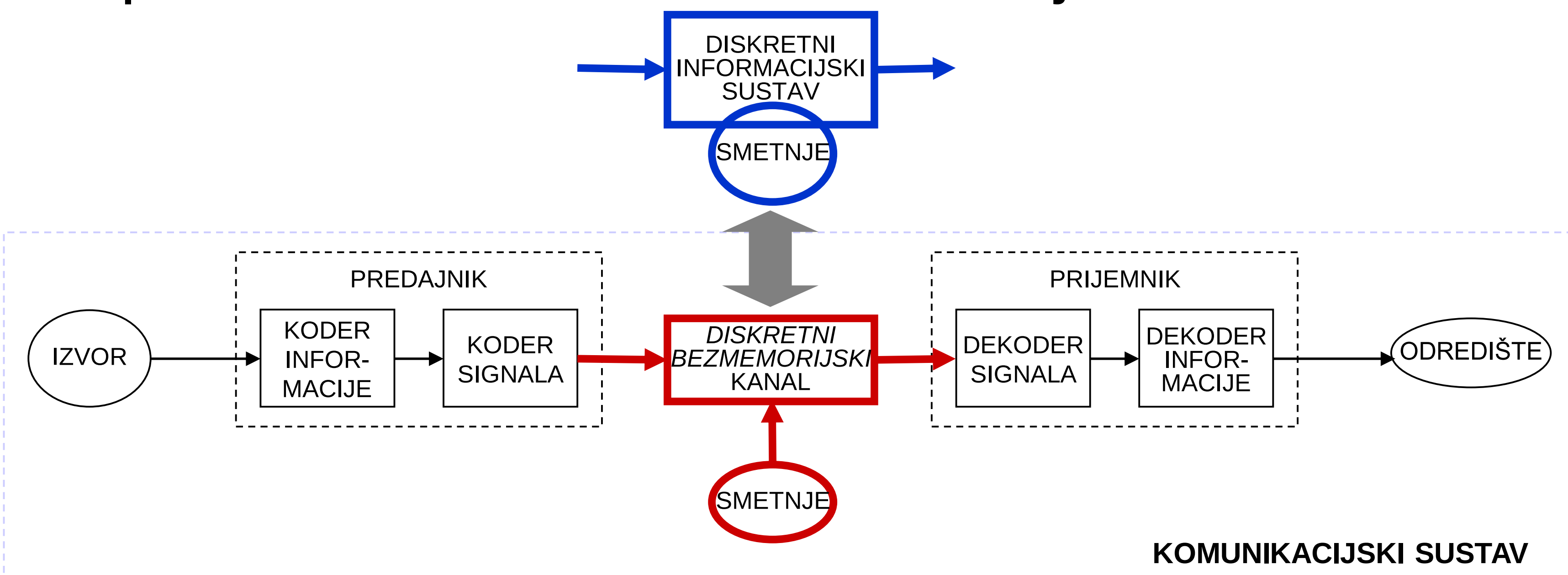
- Prosječna duljina kodne riječi:

$$L = \sum_{i=1}^{n} p_i l_i = 0.5 \cdot 1 + 0.25 \cdot 2 + 0.125 \cdot 3 + 0.125 \cdot 3 = 1.75 \left[bit / simbol \right] = H(X)$$

- Ne postoji kod sa manjom prosječnom duljinom
- **Entropija je granica kompresije bez gubitaka**

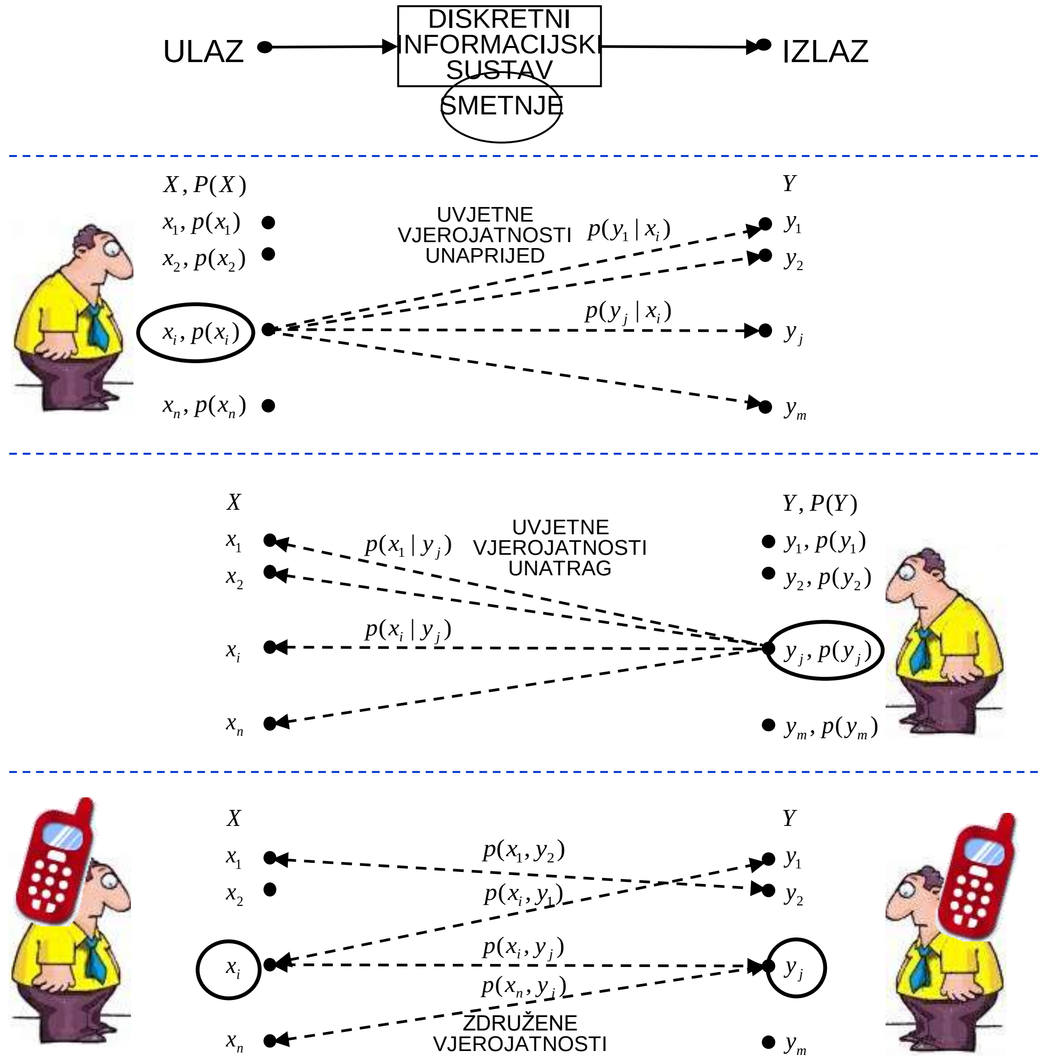
Informacijski opis komunikacijskog sustava

- Sustav bez smetnji ne postoji
  - Promatramo opći sustav uz (manja) ograničenja: diskretni bezmemorijski kanal
- Opis kanala – diskretni informacijski sustav



Vjerojatnosni opis inf. sustava (kanala)

- Opis sustava skupom vjerojatnosti
- Svaki od ova tri pogleda potpuno određuje sustav i pojave na ulazu/izlazu
- Vjerojatnosti prijelaza x→ y potpuno definiraju kanal



Primjer

- Komunikacijski kanal prenosi simbole {a, b, c}
  - $p(a) = p(b) = 2p(c)$
- Matrica uvjetnih vjerojatnosti prijelaza u kanalu:

$$\left[p\left(y_j \middle| x_i\right)\right] = \begin{bmatrix} 0,7 & 0,1 & 0,2 \\ 0,2 & 0,7 & 0,1 \\ 0,1 & 0,2 & 0,7 \end{bmatrix}$$

a) nacrtati graf prijelaza u kanalu.

b) odrediti vjerojatnost pojave pojedinog simbola na izlazu iz kanala

Odnosi vjerojatnosti u inf. sustavu (kanalu)

MATEMATIČKI OPIS	ZNAČENJE
$\sum_{i=1}^{n} p(x_i) = \sum_{j=1}^{m} p(y_j) = 1$	Skup simbola na ulazu je potpun; isto vrijedi i za izlaz.
$p(x_i) = \sum_{j=1}^{m} p(x_i, y_j),\ p(y_j) = \sum_{i=1}^{n} p(x_i, y_j)$	Vjerojatnost pojave simbola je zbroj vjerojatnosti pojava svih parova u kojima se taj simbol pojavljuje.
$p(x_i, y_j) = p(x_i) p(y_j \mid x_i) = p(y_j) p(x_i \mid y_j)$	Prijelazi između tri pogleda na sustav (pogled s ulaza, s izlaza ili oboje istovremeno). Veza između tri načina potpunog opisa sustava.
$p(x_i \mid y_j) = \frac{p(x_i, y_j)}{p(y_j)} = \frac{p(x_i, y_j)}{\sum_{i=1}^{n} p(x_i, y_j)} = \frac{p(x_i) p(y_j \mid x_i)}{\sum_{i=1}^{n} p(x_i) p(y_j \mid x_i)}$	Prijelaz iz apriorne u aposteriornu vjerojatnost pojave x_i. Izračun unazadnih vjerojatnosti prijelaza. Bayesova formula.

Vjerojatnosni opis → informacijski opis

- Entropija: informacijski opis slučajnih događaja

DOGAĐAJI

$$x_i \in X,\ X = \{x_1, x_2, \ldots, x_n\}$$

VJEROJATNOSNI OPIS

vjerojatnosti: $p(x_i),\ i = 1,\ldots,n$

INFORMACIJSKI OPIS

entropija: $H(X)$

$$H(X) = \sum_{i=1}^{n} p(x_i) \log \frac{1}{p(x_i)} = \sum_{i=1}^{n} \boxed{\boxed{p(x_i)}\boxed{I(x_i)}}$$

SADRŽAJ INFORMACIJE POJEDINOG DOGAĐAJA

ZBROJ INFORMACIJSKIH DOPRINOSA SVIH MOGUĆIH DOGAĐAJA

INFORMACIJSKI DOPRINOS POJEDINOG DOGAĐAJA

VJEROJATNOST POJEDINOG DOGAĐAJA

Informacijske mjere

vlastite entropije:
- $H(X)$ ♦ Entropija na ulazu sustava
- $H(Y)$ ♦ Entropija na izlazu sustava

$H(X,Y)$ ♦ Združena entropija

uvjetne entropije:
- $H(Y|X)$ ♦ Entropija šuma, irelevantnost
- $H(X|Y)$ ♦ Ekvivokacija, mnogoznačnost

$I(X;Y)$ ♦ Srednji uzajamni sadržaj informacije, transinformacija

DEFINICIJA

Entropija na ulazu, izlazu, združena entropija

Zavod za telekomunikacije

- Promatramo događaje na ulazu i izlazu odvojeno:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \qquad H(Y) = -\sum_{j=1}^{m} p(y_j) \log p(y_j)$$

- Promatramo događaje zajednički:
  - Združena entropija para slučajnih varijabli (definicija):

$$H(X,Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \log p(x_i, y_j)$$

Uvjetna entropija (općenito)

- Prosječna preostala neodređenost varijable *Y* nakon što je poznata varijabla *X*

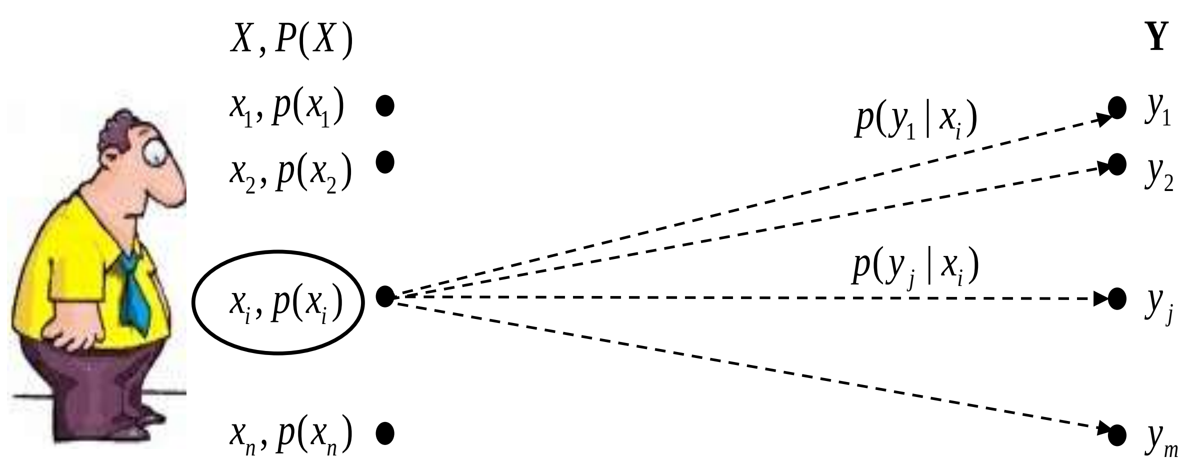
$$H(Y|X) = \sum_{i=1}^{n} p(x_i) H(Y|x = x_i)$$

$$= -\sum_{i=1}^{n} p(x_i) \sum_{j=1}^{m} p(y_j|x_i) \log p(y_j|x_i)$$

$$= -\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \log p(y_j|x_i)$$

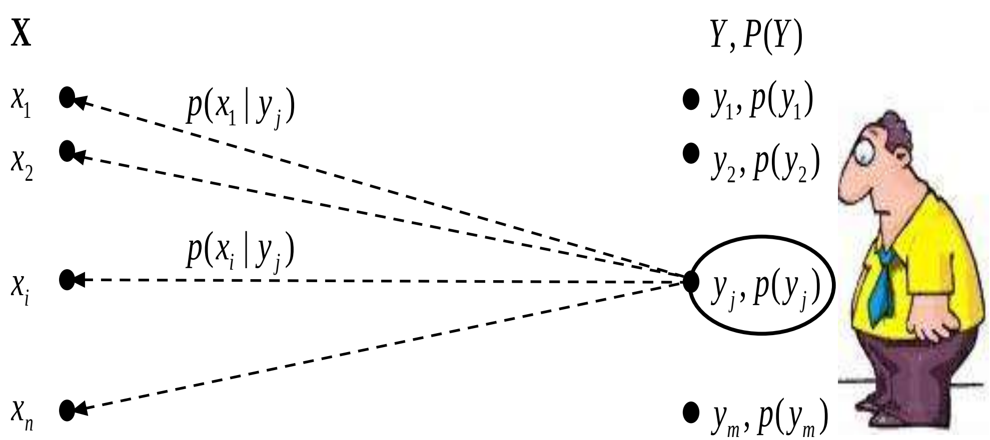
Entropija šuma ili irelevantnost

- Uvjetna entropija *H(Y|X)*
- Neodređenost simbola na izlazu nakon što je poslan simbol sa ulaza (promatrano s ulaza)
- Posljedica smetnji



Mnogoznačnost ili ekvivokacija

- Uvjetna entropija H(X|Y)
- Preostala neodređenost simbola na ulazu nakon što je primljen simbol na izlazu (promatrano s izlaza)



Relativna entropija

- Mjera udaljenosti između dviju raspodjela vjerojatnosti varijable:

$$D(p\,||\,q)=\sum_{i=1}^{n} p(x_i)\log\frac{p(x_i)}{q(x_i)}$$

- Interpretacija
  - Stvarne vjerojatnosti su p; mi pretpostavljamo q
  - Ta pogreška nosi neefikasnost; to je relativna entropija
  - Kodiranjem prema pogrešnim vjerojatnostima trošimo D(p||q) više bitova po simbolu nego što je potrebno:

$$L=\sum_{i=1}^{n} p(x_i)\log\frac{1}{q(x_i)}=\sum_{i=1}^{n} p(x_i)\log\frac{1}{p(x_i)}+\sum_{i=1}^{n} p(x_i)\log\frac{p(x_i)}{q(x_i)}=H(X)+D(p\,||\,q)$$

Srednji uzajamni sadržaj informacije (transinformacija)

DEFINICIJA

Zavod za telekomunikacije

- Definicija:

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_i,y_j)\log\frac{p(x_i,y_j)}{p(x_i)p(y_j)}$$

- Interpretacija:
  - Koliko informacije jedna varijabla pruža o drugoj
  - U kojoj mjeri su dvije varijable zavisne
    - Nezavisne: I(X;Y) = 0
    - Jednake: I(X;Y) = H(X) = H(Y)

Odnos entropije i uzajamnog sadržaja informacije

- Uzajamni sadržaj informacije *I(X;Y)* predstavlja smanjenje neodređenosti varijable *X* uzrokovano poznavanjem varijable *Y*

$$I(X;Y) = H(X) - H(X|Y)$$

- Uzajamni sadržaj informacije dviju varijabli je simetričan:

$$I(Y;X) = I(X;Y).$$

Odnos između entropije, združene entropije i uvjetne entropije

- Združena entropija (neodređenost) para varijabli jednaka je zbroju neodređenosti jedne varijable, te preostale neodređenosti druge varijable uz uvjet da je prva varijabla poznata.

$$H(X,Y) = H(X) + H(Y \mid X)$$

- Uzajamni sadržaj informacije je razlika između zbroja pojedinačnih entropija varijabli i združene entropije tih istih varijabli.

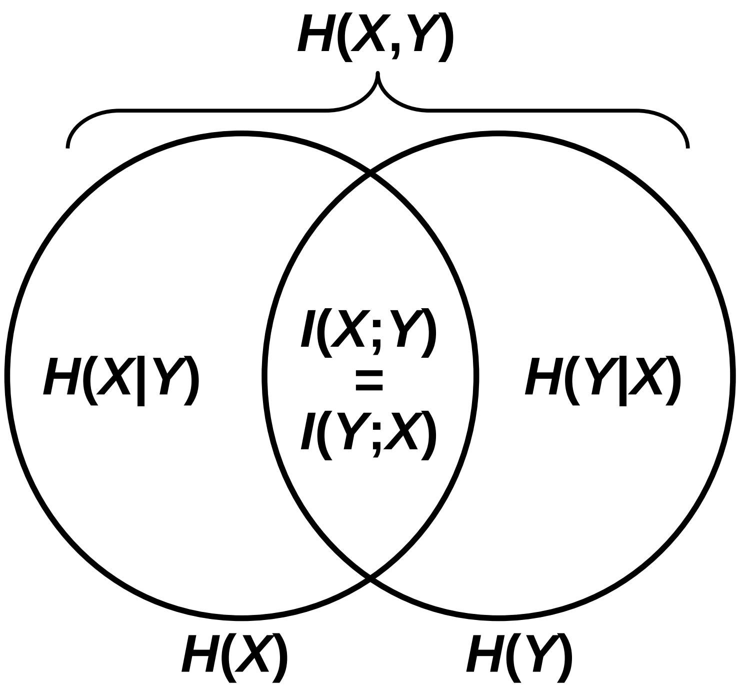
$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

Vlastiti sadržaj informacije

- Uzajamni sadržaj informacije jedne varijable same sa sobom naziva se vlastiti sadržaj informacije.
- Vlastiti sadržaj informacije slučajne varijable je upravo njena entropija:

$$I(X;X) = H(X) - H(X|X) = H(X)$$

Odnosi i svojstva informacijskih mjera



$$I(X;Y) = H(X) - H(X|Y)$$

$$I(X;Y) = H(Y) - H(Y|X)$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$$H(X,Y) = H(X) + H(Y|X)$$

$$H(X,Y) = H(Y) + H(X|Y)$$

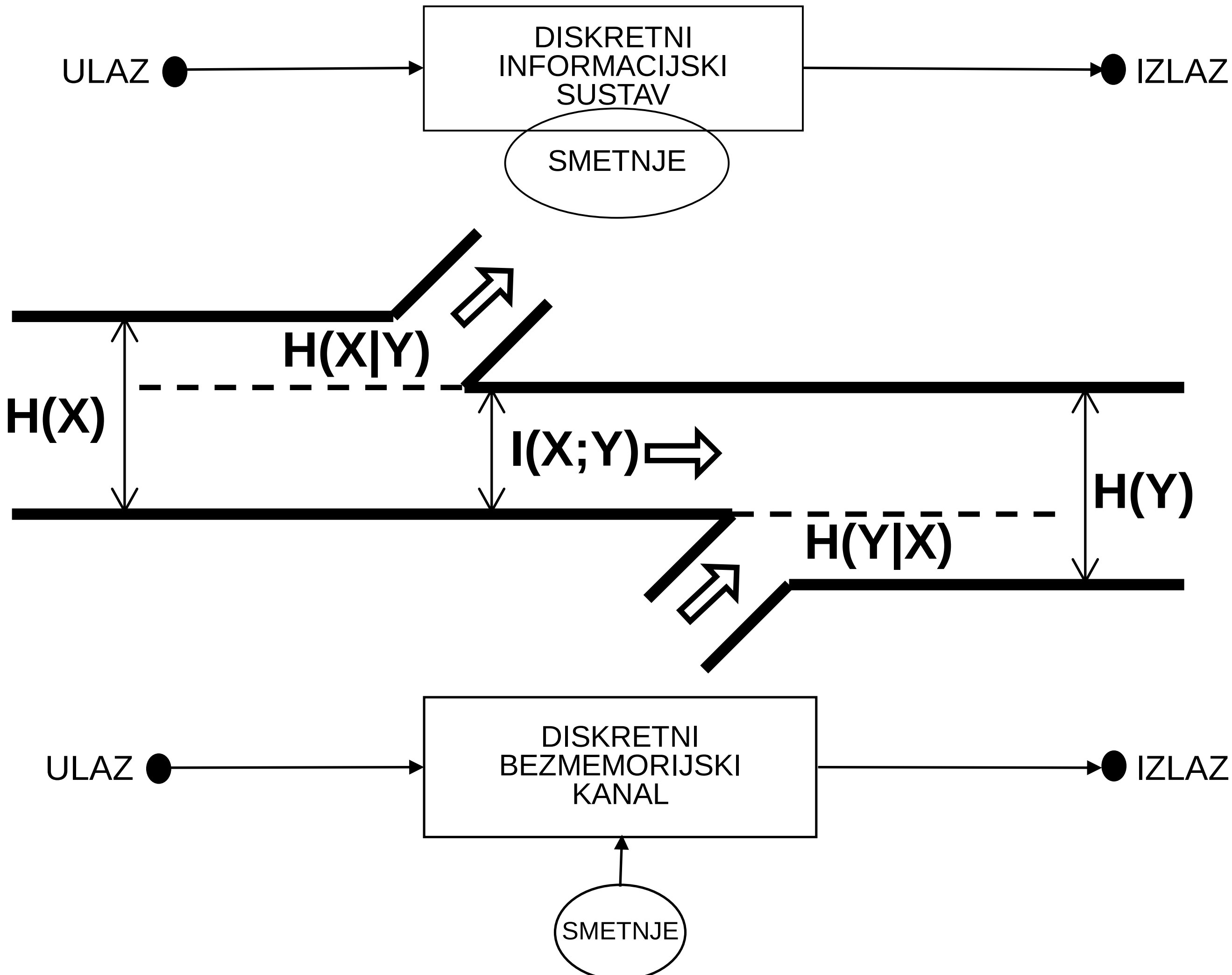
$$I(X;Y) = I(Y;X)$$

$$I(X;X) = H(X)$$

$$I(X;Y) \geq 0$$

$$H(X|Y) \leq H(X)$$

Prijenos informacije i informacijske mjere



Primjer

- Za komunikacijski sustav zadan u prethodnom primjeru matricom uvjetnih vjerojatnosti potrebno je odrediti:
  - a) entropiju ulaznog i izlaznog skupa simbola, tj. $H(X)$ i $H(Y)$;
  - b) uvjetne entropije $H(X|Y)$ i $H(Y|X)$;
  - c) uzajamni sadržaj informacije $I(X; Y)$;
  - d) združenu entropiju para varijabli $H(X, Y)$.

DEFINICIJA

Kapacitet kanala

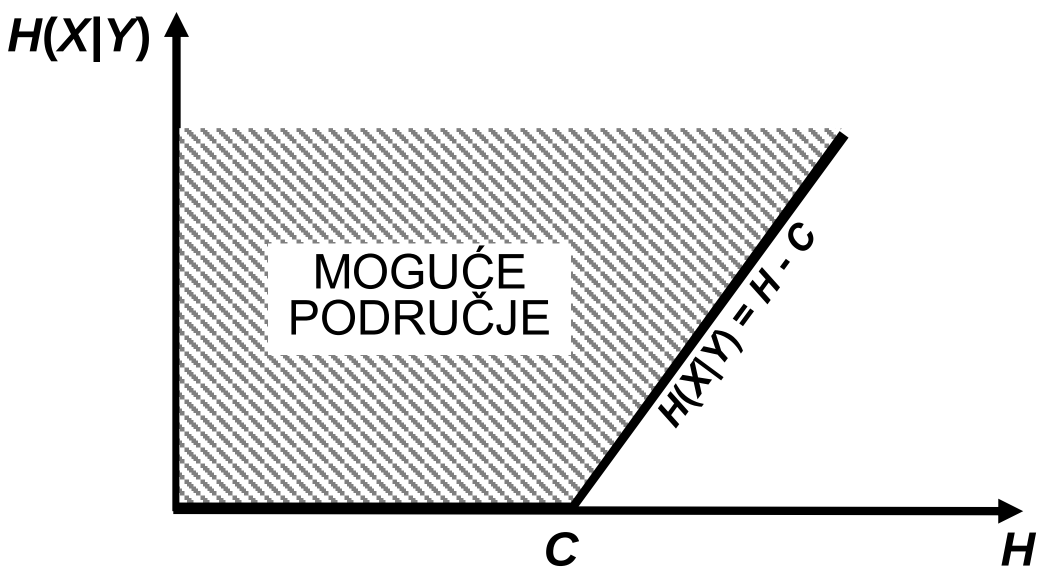
- Promatramo prijenos informacije kom. kanalom
- Simboli na ulazu s vjerojatnosima $p(x_i)$
- Kapacitet kanala je definiran kao:

$$C = \max_{\{p(x_i)\}} I(X;Y) \text{ [bit/simbol]}$$

Kapacitet kanala je maksimalna količina informacije po simbolu koja se u prosjeku može prenijeti kanalom

Temeljni teorem kanala sa smetnjama

- Kanal kapaciteta *C* [bit/simbol]
- Izvor entropije *H* [bit/simbol]
- Ako je $H \leq C$, mogući proizvoljno mali gubici
- Ako je $H > C$, nemoguć prijenos bez gubitaka



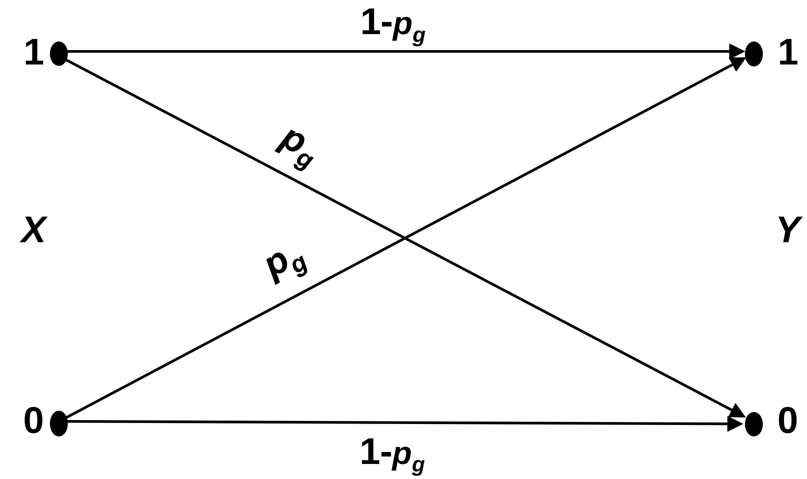
Primjer: kapacitet simetričnog binarnog kanala

$$C = \max_{\{p(x_i)\}} I(X;Y)$$

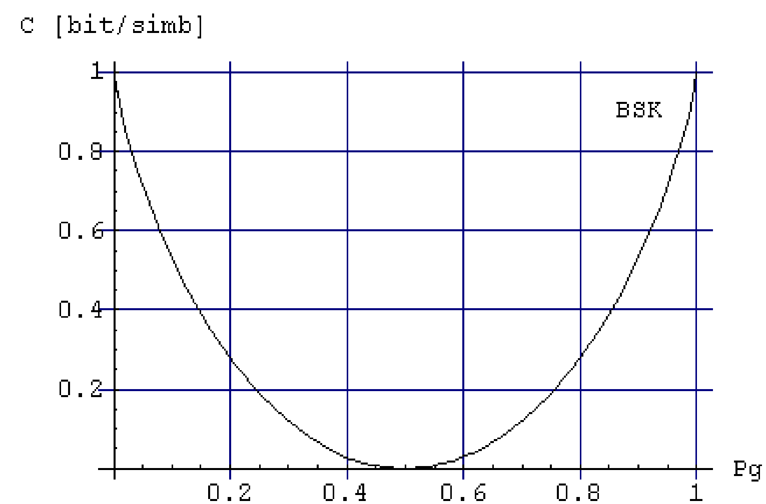
$$= \max_{\{p(x_i)\}} \left[H(Y) - H(Y \mid X) \right]$$

max. za
p(0)=p(1)=0.5

neovisno
o $p(x_i)$



$$C = 1 + p_g \log p_g + (1 - p_g)\log(1 - p_g) \left[bit / s \right]$$



Prijenos informacije komunikacijskim sustavom