

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

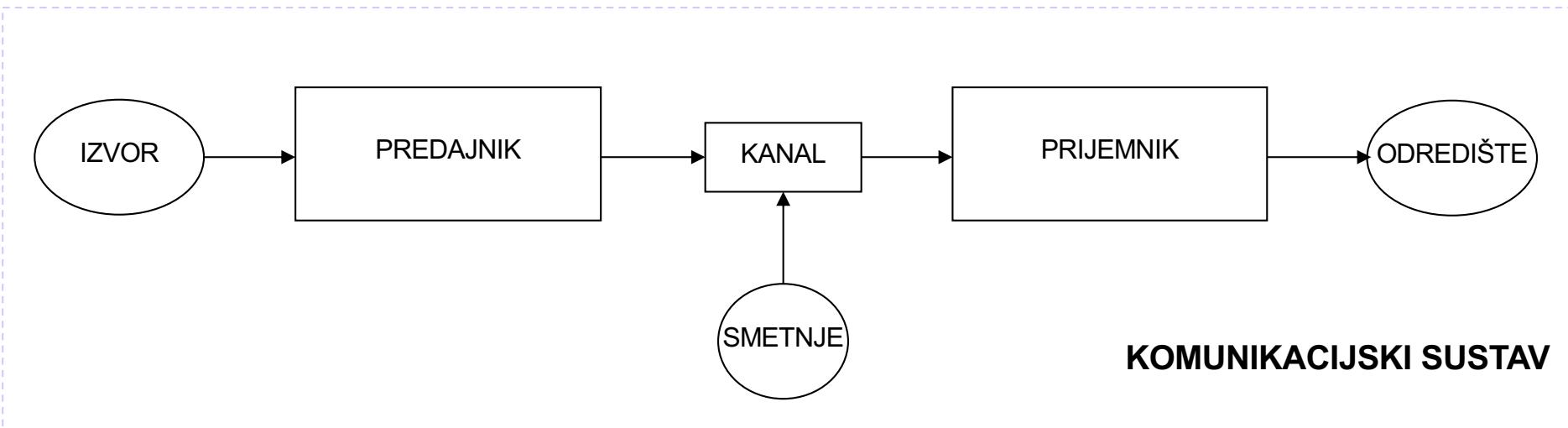
Opći model komunikacijskog sustava

DEFINICIJA



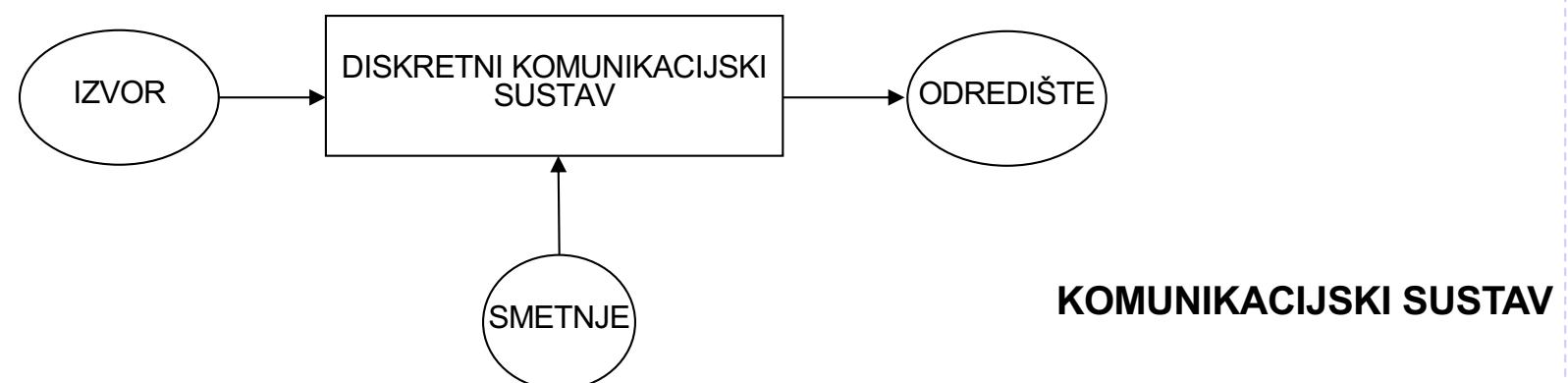
Zavod za telekomunikacije

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



Diskretni komunikacijski sustav

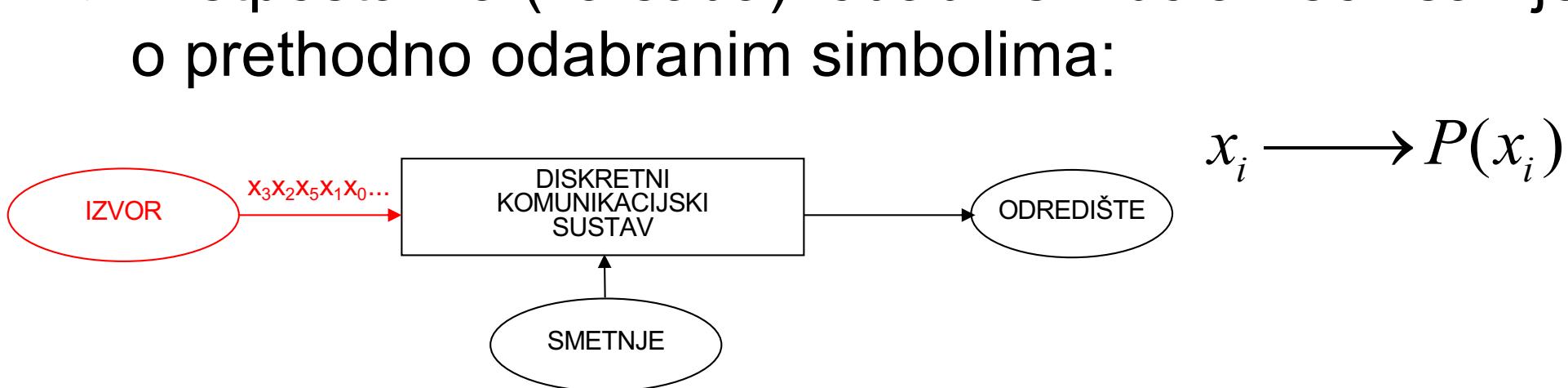
- ◆ Jednostavniji slučaj – diskretni signali
- ◆ Ključna pitanja:
 - Što je poruka?
 - Što znači prenijeti poruku?
 - Koja je mjera za količinu informacije u nekoj poruci, te informacije prenesene sustavom?



- ◆ Niz simbola odabralih iz konačne abecede X
 - Abeceda je skup elementarnih simbola

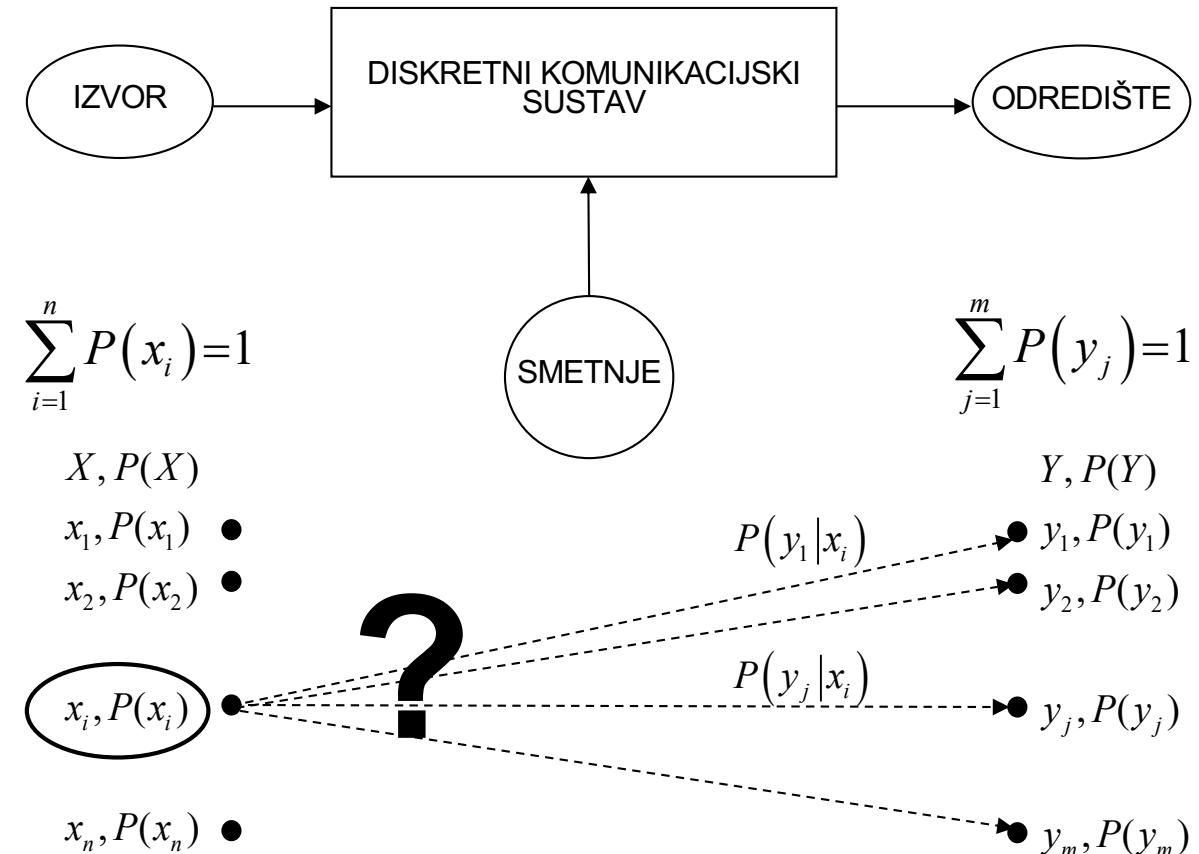
$$X = \{x_1, x_2, \dots, x_i, \dots, x_n\}$$

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



Prijenos poruke: pogled s izvora

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



Informacijski kanal

- ◆ informacijski kanal je statistički model medija kroz koji se prenosi signal
 - signal je fizikalni prikaz simbola kojeg prenosi
- ◆ cilj: proračunati koliko se informacije prenosi kroz kanal
- ◆ modeliranje kanala matricom uvjetnih vjerojatnosti

$$[P(Y|X)] = \begin{bmatrix} P(y_1|x_1) & P(y_2|x_1) & \dots & P(y_m|x_1) \\ P(y_1|x_2) & P(y_2|x_2) & \dots & P(y_m|x_2) \\ \vdots & \vdots & \vdots & \vdots \\ P(y_1|x_n) & P(y_2|x_n) & \dots & P(y_m|x_n) \end{bmatrix}$$

Informacijski kanal (2)

- ◆ zbroj članova bilo kojeg retka je 1
 - za svaki ulazni simbol x_i sigurno je da će se nešto pojaviti na izlazu
- ◆ zdržena vjerojatnost para simbola x_i i y_j je dana poznatim Bayesovim teoremom

$$\sum_{j=1}^m p_{i,j} = \sum_{j=1}^m P(y_j | x_i) = 1$$

$$P(x_i, y_j) = P(y_j, x_i) = P(y_j | x_i) \cdot P(x_i) = P(x_i | y_j) \cdot P(y_j)$$

$$\sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) = 1$$

- ◆ značenje: kad nešto uđe u kanal, sigurno će se nešto pojaviti i na izlazu kanala

Informacijski kanal (3)

◆ matrica združenih vjerojatnosti

$$[P(X, Y)] = \begin{bmatrix} P(x_1, y_1) & P(x_1, y_2) & \cdots & P(x_1, y_m) \\ P(x_2, y_1) & P(x_2, y_2) & \cdots & P(x_2, y_m) \\ \vdots & \vdots & \vdots & \vdots \\ P(x_n, y_1) & P(x_n, y_2) & \cdots & P(x_n, y_m) \end{bmatrix}$$

$$[P(X, Y)] = [P(X)] \cdot [P(Y|X)]$$

$$[P(X)] = \begin{bmatrix} P(x_1) & 0 & \cdots & 0 \\ 0 & P(x_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P(x_n) \end{bmatrix}$$

$$[P(y_1), P(y_2), \dots, P(y_m)] = [P(x_1), P(x_2), \dots, P(x_n)] \cdot [P(Y|X)]$$

Informacijski kanal (4)

- ◆ dodatna svojstva:

$$\sum_{i=1}^n P(y_j | x_i) P(x_i) = \sum_{i=1}^n P(x_i, y_j) = P(y_j)$$

$$\sum_{j=1}^m P(x_i | y_j) P(y_j) = \sum_{j=1}^m P(x_i, y_j) = P(x_i)$$

$$P(x_i | y_j) = \frac{P(x_i, y_j)}{P(y_j)} = \frac{P(y_j | x_i) P(x_i)}{\sum_{i=1}^n P(y_j | x_i) P(x_i)}$$

$$\sum_{i=1}^n P(x_i | y_j) = 1$$

- ◆ značenje: za neki simbol y_j na izlazu kanala sigurno se je neki od simbola x_i pojavio na ulazu kanala

Informacijski kanal (5)

- ◆ matrica uvjetnih vjerojatnosti $P(x_i|y_j)$

$$[P(X|Y)] = \begin{bmatrix} P(x_1|y_1) & P(x_1|y_2) & \cdots & P(x_1|y_m) \\ P(x_2|y_1) & P(x_2|y_2) & \cdots & P(x_2|y_m) \\ \vdots & \vdots & \vdots & \vdots \\ P(x_n|y_1) & P(x_n|y_2) & \cdots & P(x_n|y_m) \end{bmatrix}$$

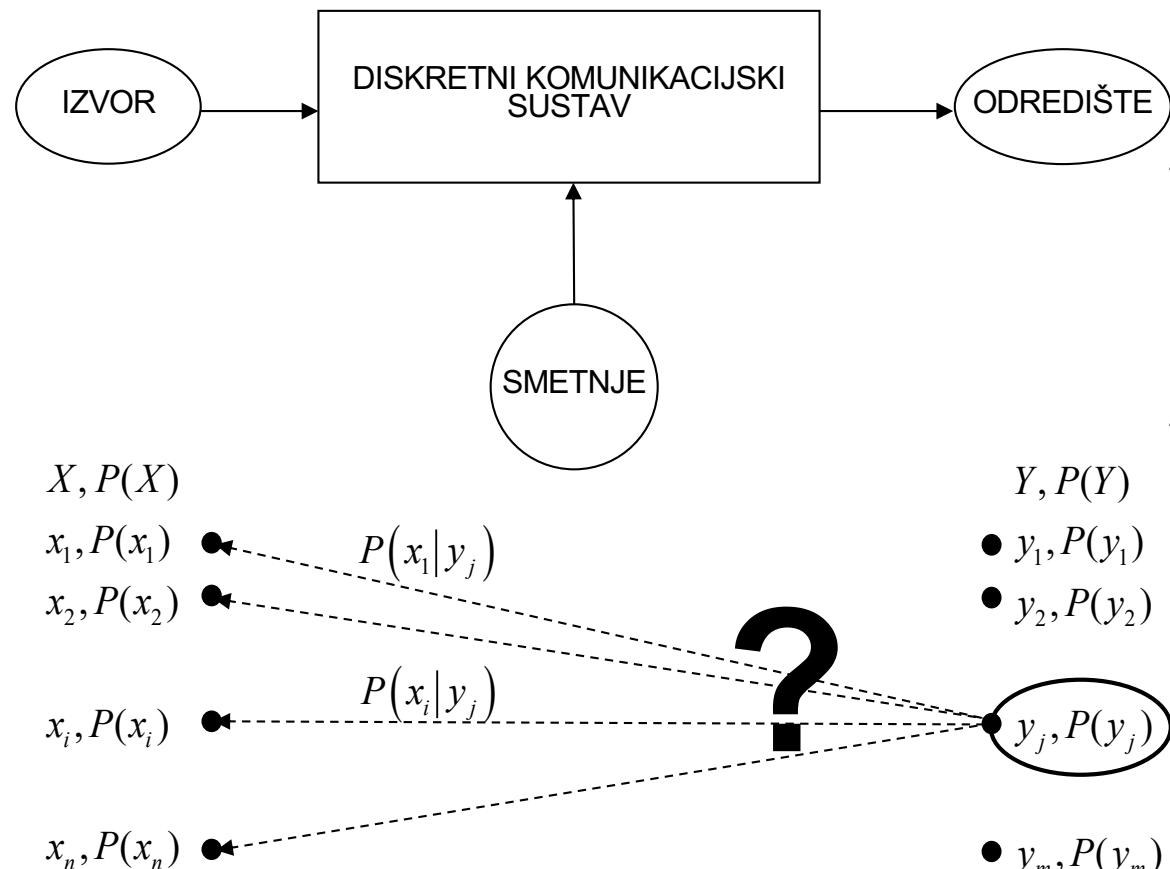
zbroj članova bilo kojeg stupca jednak je 1 – za svaki simbol y_j sigurno je da je primljen nakon što je poslan jedan od simbola x_i

$$[P(X, Y)] = [P(X|Y)] \cdot [P(Y)]$$

$$\begin{bmatrix} P(x_1) \\ \vdots \\ P(x_n) \end{bmatrix} = [P(X|Y)] \cdot \begin{bmatrix} P(y_1) \\ \vdots \\ P(y_m) \end{bmatrix}$$

$$[P(Y)] = \begin{bmatrix} P(y_1) & 0 & \cdots & 0 \\ 0 & P(y_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P(y_m) \end{bmatrix}$$

Prijenos poruke: pogled s odredišta



- ◆ Prije pojave simbola y_j na odredištu znamo vrlo malo o događajima na izvoru
- ◆ Nakon opažanja simbola y_j znamo više: primili smo **informaciju!**
- ◆ Jedan od osnovnih problema TI-a:
 - kako kvantitativno odrediti informaciju koju "nosi" određeni simbol/poruka, odnosno
 - kako matematički definirati mjeru za informaciju

Informacija kao mjera neodređenosti

- ◆ u slučaju determinističkog modela izvora točno je poznat redoslijed slanja simbola
 - mjerjenje količine informacije nepotrebno
- ◆ u statistički definiranom modelu nije unaprijed poznat točan redoslijed slanja simbola
 - postoji potreba za statističkim opisom modela
 - potraga za količinom informacije je u stvari potraga za statističkim parametrom koji je povezan s vjerojatnosnim modelom izvora
 - količina informacije je parametar koji treba iskazati relativnu mjeru **neodređenosti** koja se odnosi na pojavu svakog pojedinačnog simbola u promatranom slijedu

Veza između informacije i vjerojatnosti

- ◆ pretpostavimo abecedu od n simbola: x_1, x_2, \dots, x_n
 - razdioba vjerojatnosti: $P(x_1), P(x_2), \dots, P(x_n)$,
 - mora vrijediti: $\sum_{i=1}^n P(x_i) = 1$
- ◆ pitanje: kad primimo neki od tih simbola, koliko smo primili informacije?
 - npr. ako je $P(x_1) = 1$, a sve ostale $P(x_i) = 0$, tada nema iznenadjenja, prema tome niti informacije, jer unaprijed znamo ishod prijenosa
 - informacija je obrnuto proporcionalna vjerojatnosti pojave simbola
 - ako primimo manje vjerojatan simbol, iznenadjenje je veće

Definicija sadržaja informacije

- ◆ količina informacije ili sadržaj informacije definiran je kao

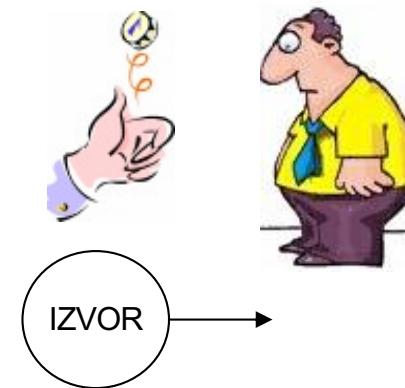
$$I(x_i) = \log_2 \left(\frac{1}{P(x_i)} \right) = -\log_2 (P(x_i)) \text{[bit]}$$

- ◆ informacija ima svojstvo aditivnosti:
 - količina informacije od dva različita i međusobno neovisna simbola jednaka je zbroju količina informacije od svakog od ta dva simbola
 - shodno tome vrijedi:

$$I(x_1, x_2) = I(x_1) + I(x_2) = \log_2 \left(\frac{1}{P(x_1)P(x_2)} \right)$$

Sadržaj informacije poruke - primjer

- ◆ Koliko informacija 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?

Primjer za ilustraciju definicije količine informacije

- ◆ iz kataloga odabiremo jedan iz skupa od n modela nekog proizvoda: $\{x_1, x_2, \dots, x_n\}$
- ◆ željena količina informacije (za kojom tragamo) pridružena odabiru jednog određenog modela x_k mora biti funkcija vjerojatnosti odabira x_k :

$$I(x_k) = f(P(x_k))$$

- ◆ ako pretpostavimo da je podjednako vjerojatno da ćemo odabrati bilo koji model, tada vrijedi

$$I_1(x_k) = f\left(\frac{1}{n}\right)$$

- ◆ nadalje, pretpostavimo da je promatrani proizvod dostupan u m boja: $\{c_1, c_2, \dots, c_m\}$

Primjer (nastavak)

- ◆ ako pretpostavimo da je odabir bilo koje boje jednako vjerojatan, tada vrijedi, po analogiji:

$$I_2(c_j) = f(P(x_j)) = f\left(\frac{1}{m}\right)$$

- ◆ pri čemu se radi o istoj funkciji kao i kod modela
- ◆ neka postoji dva načina odabira proizvoda:
 - a) odabiremo model i nakon toga boju, pri čemu su ta dva odabira međusobno neovisna
 - b) istovremeno odabiremo i model i boju i tretiramo kao jedan odabir između $n \cdot m$ ponuđenih opcija

Primjer (nastavak)

- ◆ potraga za odgovarajućom funkcijom $f(x)$ se temelji na intuitivnom odabiru koji pretpostavlja jednakost između količine informacije u oba slučaja, a i b:

$$a) I(x_k, c_j) = I_1(x_k) + I_2(c_j) = f\left(\frac{1}{n}\right) + f\left(\frac{1}{m}\right), \quad b) I(x_k \wedge c_j) = f\left(\frac{1}{nm}\right)$$

- ◆ dakle, mora vrijediti:

$$f\left(\frac{1}{n}\right) + f\left(\frac{1}{m}\right) = f\left(\frac{1}{nm}\right)$$

- ◆ ta jednakost ima više mogućih rješenja, a najvažnije od njih za definiranje količine informacije je: $f(x) = -\log(x)$

Mjera za količinu informacije

- ◆ H. Nyquist (1924.) i R. Hartley (1928.) zaključili da mjera za količinu informacije koju prenosi određeni simbol mora imati logaritamski karakter:
 - Hartley je predložio da se simbolu koji se bira iz skupa od n mogućih simbola pridruži informacija $I(n) = \log(n)$
 - razlog: informacija koju nosi poruka sastavljena od dva simbola, pri čemu se jedan bira iz skupa od m , a drugi iz skupa od n mogućih simbola, treba biti jednak zbroju informacija koju nose pojedini simboli:
 - broj parova simbola je $m \cdot n$, pa treba vrijediti $I(m \cdot n) = I(m) + I(n)$
 - » logaritamska funkcija predstavlja logičan odabir
 - $\log(n) > 0$, za svaki $n \in \mathbf{N}$ (informacija je uvijek pozitivna)
 - $\log(n)$ – mjera neizvjesnosti u vezi s izborom jedne od n mogućih alternativa
 - » $m < n \Rightarrow \log(m) < \log(n)$ (veći broj alternativa stvara veću neizvjesnost)

Mjera za količinu informacije (2)

- ◆ mana Hartleyeve mjere je u tome što se sve poruke tretiraju ravnopravno, tj. kao jednako vjerojatne, a to nije u skladu s fizikalnom realnošću
- ◆ C. Shannon (1948.) je uočio da simboli i poruke kao i šumovi u komunikacijskim sustavima imaju statistička obilježja
 - simboli se pojavljuju u skladu s određenim razdiobama vjerojatnosti
 - simbol koji se pojavljuje s vjerojatnošću $P(x_i) = p_i > 0$, nosi informaciju $I(p_i) = \log(1/p_i) > 0$
 - temeljem toga definirana je entropija kao srednja (prosječna) količina informacije koju nosi pojedini simbol
- ◆ ako ima n mogućih simbola i vrijedi $\sum_{i=1}^n p_i = 1$
- ◆ tada je srednja količina informacije koju nosi pojedini simbol

$$H(p_1, \dots, p_n) = \sum_{i=1}^n p_i I(p_i) = - \sum_{i=1}^n p_i \log(p_i)$$

Mjera za količinu informacije (2)

- ◆ Hartleyeva definicija je poseban slučaj Shannonove za $p_i = 1/n$
- ◆ Shannonova definicija respektira činjenicu da se simboli pojavljuju u skladu s određenom razdiobom vjerojatnosti (p_1, \dots, p_n) pri čemu je
 - ◆ $p_i > 0$ i $\sum_{i=1}^n p_i = 1$
 - ◆ dakle, srednja informacija po simbolu ovisi samo o razdiobi vjerojatnosti, a ne o nekim drugim veličinama koje obilježavaju simbole

Svojstva Shannonove entropije

- ◆ $H(p_1, \dots, p_n)$ – entropija konačne razdiobe vjerojatnosti
 - neizvjesnost glede slučajnog izbora jednog od n mogućih simbola
 - sličan izraz koristio je i Boltzman za entropiju idealnog plina
- ◆ osnovna svojstva Shannonove entropije:
 - $H(1) = 0$
 - $H(p_1, \dots, p_n)$ ne ovisi o permutaciji vjerojatnosti p_1, \dots, p_n
 - $H(p_1, \dots, p_n, 0) = H(p_1, \dots, p_n)$ ($0 \cdot \log(0) = 0$)
 - $H(1/n, \dots, 1/n) = h(n)$ je rastuća funkcija od $n \in \mathbb{N}$
 - $H(p, 1 - p)$ je neprekidna funkcija za $0 < p < 1$
 - ako je $1 \leq r \leq n$, $q_1 = p_1 + \dots + p_r$ i $q_2 = p_{r+1} + \dots + p_n$
 - tada je

$$H(p_1, \dots, p_n) = H(q_1, q_2) + q_1 H\left(\frac{p_1}{q_1}, \dots, \frac{p_r}{q_1}\right) + q_2 H\left(\frac{p_{r+1}}{q_2}, \dots, \frac{p_n}{q_2}\right)$$

Teorem jedinstvenosti

- ◆ pitanje: da li je Shannonova entropija jedina moguća mjera neizvjesnosti s navedenim svojstvima?
 - odgovor daje teorem jedinstvenosti
 - ističe svojstva pomoću kojih je moguće dokazati da je mjera za neizvjesnost, kao određena funkcija od broja n i konačne razdiobe vjerojatnosti (p_1, \dots, p_n) , uz $p_i \geq 0$ i uz $\sum_{i=1}^n p_i = 1$, baš Shannonova entropija, uz proizvoljnu bazu veću od jedan
 - ◆ neka je $P_n = \left\{ (p_1, \dots, p_n) \in R^n : p_i \geq 0, \sum_{i=1}^n p_i = 1 \right\}$
 - ◆ skup svih razdioba vjerojatnosti na n -članom skupu ($\forall n, H : P_n \rightarrow \mathbf{R}$)
 - ◆ **teorem:** funkcija H , kao funkcija od $n \in \mathbf{N}$ i $(p_1, \dots, p_n) \in P_n$, koja ima svojstva označena crveno na prethodnom slajdu, nužno je oblika
- $$H(p_1, \dots, p_n) = -C \sum_{i=1}^n p_i \log_b(p_i), C > 0 \text{ i } b > 1, C, b \in \mathbf{R}$$
- ◆ mjera za neizvjesnost, kao određena funkcija prirodnog broja n i konačne razdiobe vjerojatnosti (p_1, \dots, p_n) je baš Shannonova entropija, uz proizvoljnu logaritamsku bazu veću od 1

Izbor baze logaritma

- ◆ ako se postavi i zahtjev kojim se definira jedinična neizvjesnost $H(1/2, 1/2) = h(2) = 1$
 - jediničnu neizvjesnost ima situacija s dva jednakovjerojatna stanja
- ◆ tada je $b = 2^C$
- ◆ ako se uzme $C = 1$, tada je $H(p_1, \dots, p_n) = -\sum_{i=1}^n p_i \log_2 p_i$ [bit/simbol]
- ◆ jedan bit entropije odgovara situaciji s dva jednakovjerojatna stanja
- ◆ mjera za informacijski sadržaj ovisi o bazi logaritma:
 - za bazu 2 – bit
 - za bazu e – nat ili nit
 - za bazu 10 – Hartley ili dit
- ◆ **bit kao mjeru za informacijski sadržaj ne brkati s bitom kao binarnom znamenkom!**

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)} \quad \log_2(x) = \text{ld}(x) = \frac{\ln(x)}{\ln(2)}$$

Entropija

- ◆ Entropija diskretne slučajne varijable
 - odabrana je baza logaritma 2, u nastavku: $\log = \log_2$

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

$$\log_a(b) = \frac{1}{\log_b(a)}$$

$$H_r(X) = -\sum_{i=1}^n P(x_i) \log_r(\overleftarrow{P}(x_i)) \text{[*/simbol]} = H_2(X) \log_r(2)$$

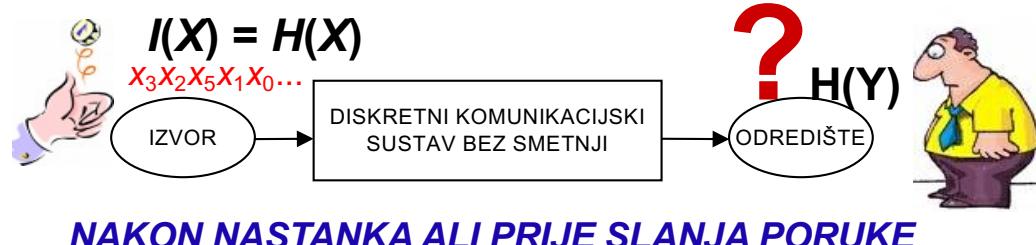
- * označava mjeru informacije (bit, nit, dit ili ...)
- ◆ Entropija daje mjeru za sadržaj informacije

$$H(X) = \sum_{i=1}^n I(x_i) P(x_i) = E(I(X))$$

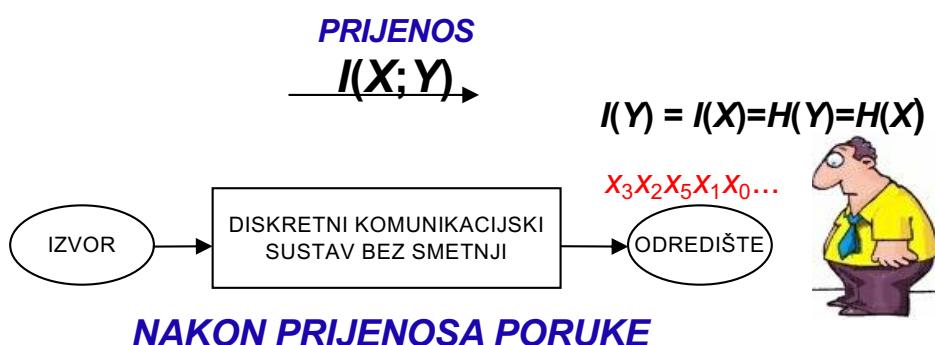
Entropija, neodređenost, sadržaj informacije u sustavu bez smetnji



- ◆ Neodređenost = entropija



- ◆ Informacija na izvoru, neodređenost na odredištu



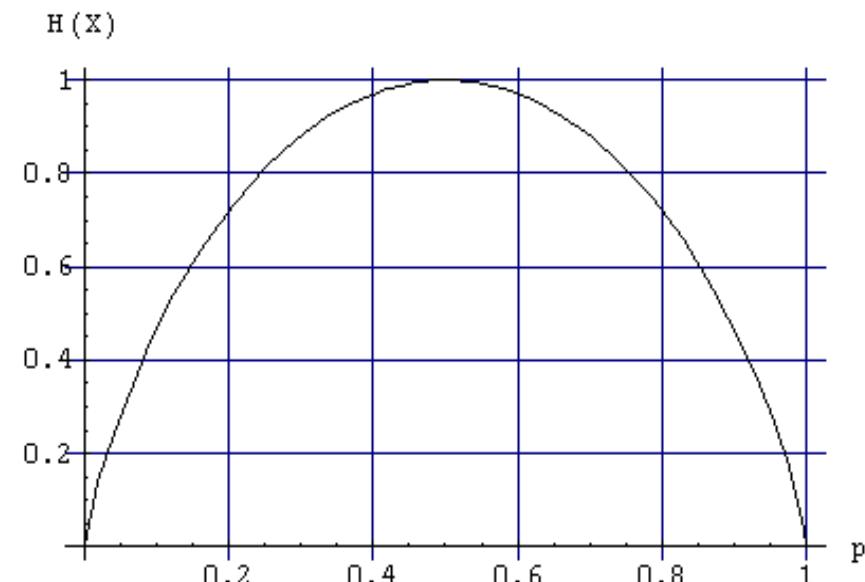
- ◆ Prijenosom poruke neodređenost je nestala

Primjer računanja entropije za dva simbola

- ◆ $X = \{x_1, x_2\}, P(x_1) = p, P(x_2) = 1 - p$

$$H(X) = p \log_2 \left(\frac{1}{p} \right) + (1-p) \log_2 \left(\frac{1}{1-p} \right)$$

$$\lim_{x \rightarrow 0} (x \log(x)) = 0$$



- ◆ $H(p, 1-p)$ je neprekidna funkcija za $0 < p < 1$
 - ◆ svojstvo entropije, slajd 22

Svojstva entropije

- ◆ Sadržaj informacije ne može biti negativan
- ◆ Sadržaj informacije je 0 ako se uvijek pojavljuje samo jedan simbol
- ◆ Neodređenost i sadržaj informacije su maksimalni ako su vjerojatnosti simbola jednako raspoređene
- ◆ Svojstvo aditivnosti
 - svojstvo logaritamske funkcije

$$H(X) \geq 0$$

$$H(X) = 0 \Leftrightarrow \exists i \mid P(x_i) = 1$$

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

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

$$H(XY) = H(X) + H(Y)$$

Maksimalna vrijednost entropije

$$\log_e(x) \leq x - 1$$

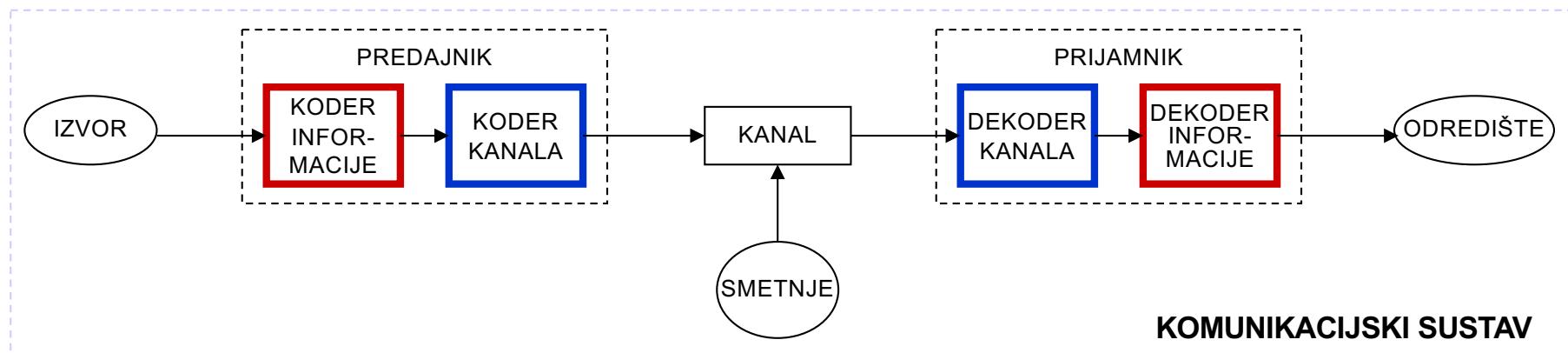
- ◆ promatrajmo:
$$H(X) - \log(n) = \sum_{i=1}^n P(x_i) \log \frac{1}{P(x_i)} - \sum_{i=1}^n P(x_i) \log(n)$$
$$= \sum_{i=1}^n P(x_i) \log \left(\frac{1}{nP(x_i)} \right) = \log(e) \sum_{i=1}^n P(x_i) \log_e \left(\frac{1}{nP(x_i)} \right)$$
$$H(X) - \log(n) \leq \log(e) \sum_{i=1}^n P(x_i) \left[\frac{1}{nP(x_i)} - 1 \right]$$
$$\leq \log(e) \left(\sum_{i=1}^n \frac{1}{n} - \sum_{i=1}^n P(x_i) \right) \leq \log(e) \cdot (1 - 1) = 0$$
- ◆ dakle: $H(X) \leq \log(n)$
 - jednakost je moguće postići samo ako su svi $P(x_i)$ jednaki i iznose $1/n$

Bit i binarna znamenka

- ◆ Teorija informacije: bit je osnovna jedinica informacije
- ◆ U većini ostalih primjena bit je binarna znamenka
- ◆ Primjer koji pomaže u razlikovanju tih dviju definicija bita:
 - bacamo “nepošteni” novčić, pismo = 1, glava = 0, koliko je ovo bita: **1111111111** ?
 - poruka duljine 10 bita (binarnih znamenaka)
 - informacijski sadržaj = 0 bita
- ◆ obično je iz konteksta jasno da li se odnosi na binarne znamenke ili sadržaj informacije

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

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

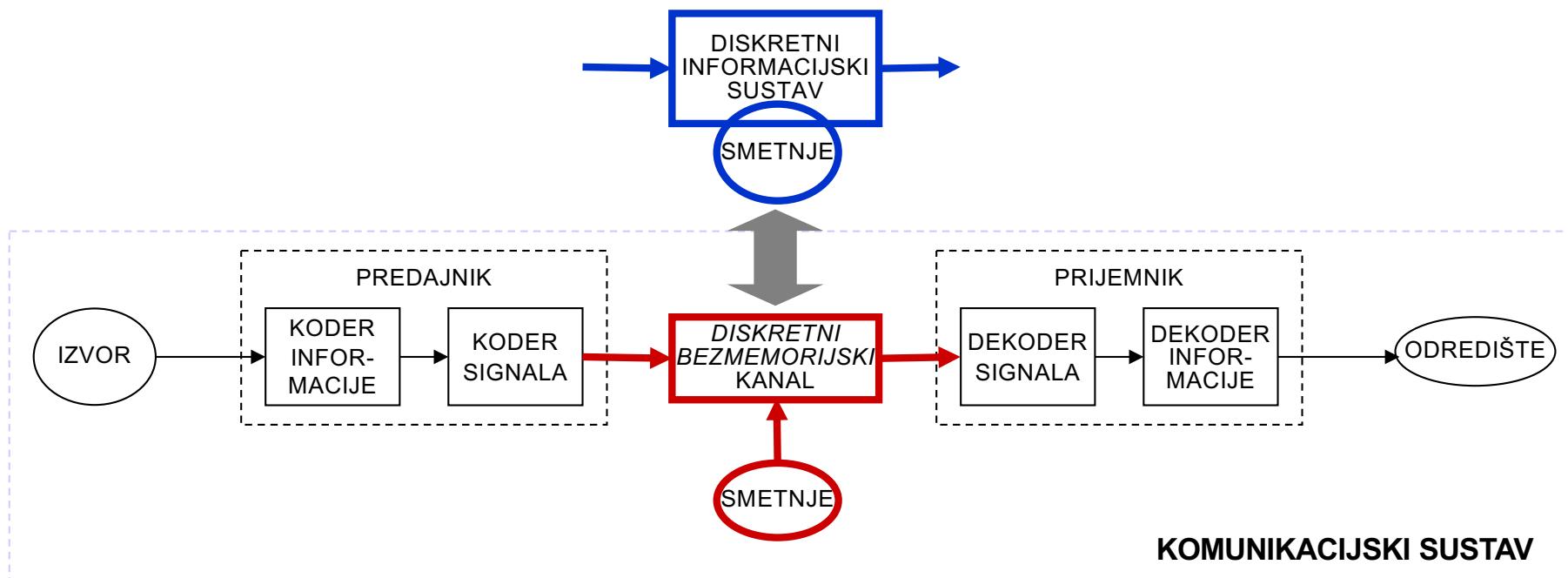
- ◆ Prosječna duljina kodne riječi:

$$L = \sum_{i=1}^n p_i l_i = 0,5 \cdot 1 + 0,25 \cdot 2 + 0,125 \cdot 3 + 0,125 \cdot 3 = 1,75 \text{ [bit/simbol]} = H(X)$$

- ◆ Ne postoji kod s manjom prosječnom duljinom kodne riječi
- ◆ **Entropija predstavlja graničnu vrijednost kompresije bez gubitaka**

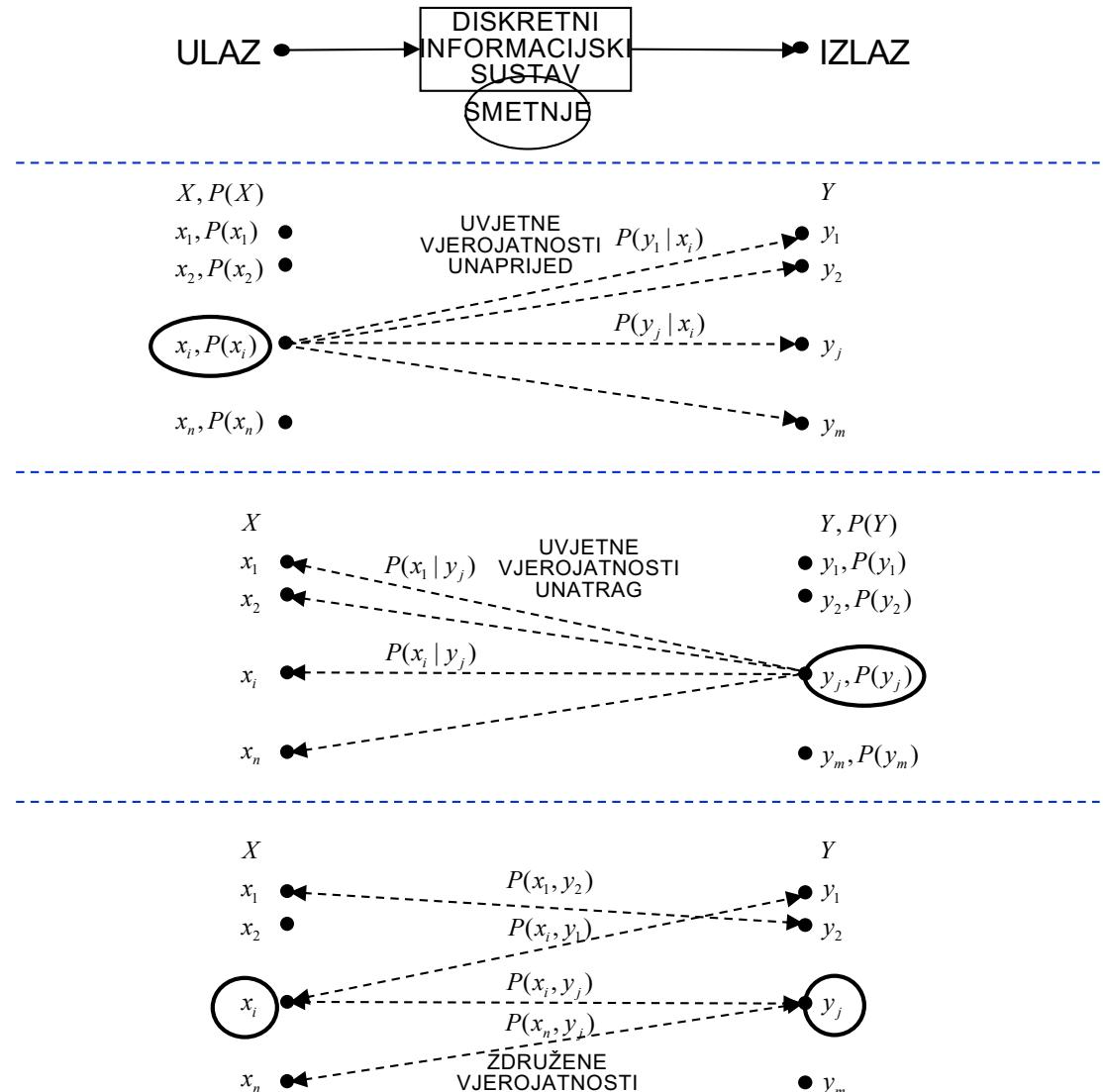
Informacijski opis komunikacijskog sustava

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



Vjerojatnosni opis informacijskog sustava (kanala)

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



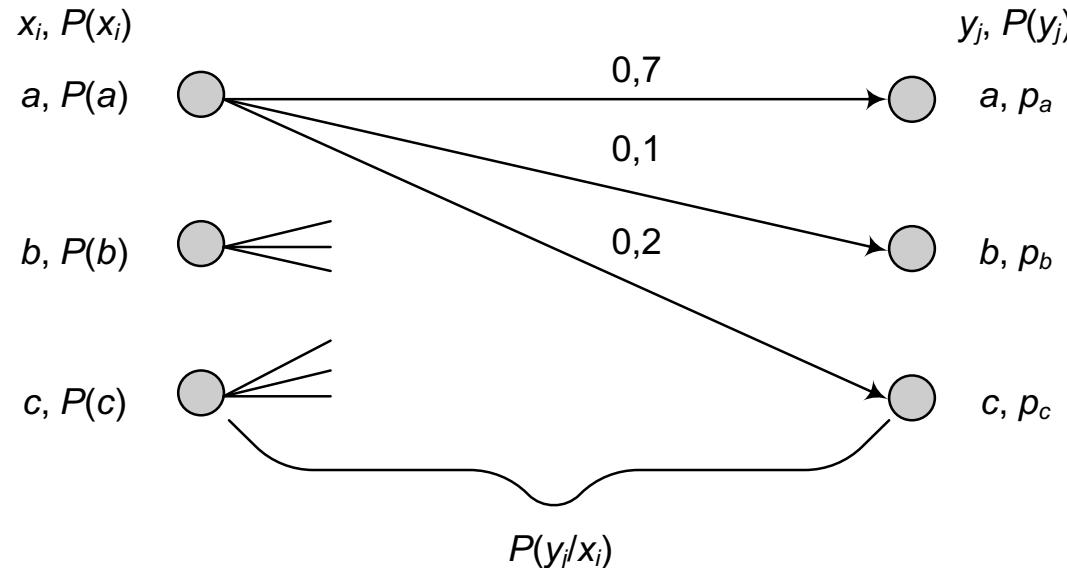
- ◆ Komunikacijski kanal prenosi simbole $\{a, b, c\}$
 - $P(a) = P(b) = 2P(c)$
- ◆ Matrica uvjetnih vjerojatnosti prijelaza u kanalu:

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

Zadatak:

- a) nacrtati graf prijelaza u kanalu
- b) odrediti vjerojatnost pojave pojedinog simbola na izlazu iz kanala

Rješenje primjera:



- ◆ a) graf prijelaza u kanalu
- ◆ b) iz uvjeta
 - $P(a) = P(b) = 2P(c)$ i
 - $P(a) + P(b) + P(c) = 1$
 - dobivamo: $P(a) = P(b) = 0,4$ i $P(c) = 0,2$

Rješenje primjera (2):

- ◆ vjerojatnosti pojave pojedinog simbola na izlazu iz kanala određujemo iz matrice zdrženih vjerojatnosti:

$$\begin{bmatrix} P(x_i, y_j) \end{bmatrix} = \begin{bmatrix} P(x_i)P(y_j | x_i) \end{bmatrix} = \begin{bmatrix} P(x_1, y_1) & P(x_1, y_2) & P(x_1, y_3) \\ P(x_2, y_1) & P(x_2, y_2) & P(x_2, y_3) \\ P(x_3, y_1) & P(x_3, y_2) & P(x_3, y_3) \end{bmatrix} = \begin{bmatrix} 0,28 & 0,04 & 0,08 \\ 0,08 & 0,28 & 0,04 \\ 0,02 & 0,04 & 0,14 \end{bmatrix}$$

- ◆ zbroj elemenata po retku je $P(x_i)$, a $\sum_{j=1}^3 P(x_i, y_j) = P(x_i)$
- ◆ zbroj elemenata po stupcu je $P(y_j)$:
- ◆ $P(y_1) = p_a = 0,38$
- ◆ $P(y_2) = p_b = 0,36$
- ◆ $P(y_3) = p_c = 0,26$
- ◆ provjera: $p_a + p_b + p_c = 1$

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.
$P(x_i) = \sum_{j=1}^m P(x_i, y_j), \quad P(y_j) = \sum_{i=1}^n P(x_i, y_j)$	Vjerojatnost pojave simbola je zbroj vjerojatnosti pojava svih parova u kojima se taj simbol pojavljuje.
$P(x_i, y_j) = P(x_i)P(y_j x_i) = P(y_j)P(x_i y_j)$	Prijelazi između tri pogleda na sustav (pogled s ulaza, s izlaza ili oboje istovremeno). Veza između tri načina potpunog opisa sustava.
$P(x_i y_j) = \frac{P(x_i, y_j)}{P(y_j)} = \frac{P(x_i, y_j)}{\sum_{i=1}^n P(x_i, y_j)} = \frac{P(x_i)P(y_j x_i)}{\sum_{i=1}^n P(x_i)P(y_j x_i)}$	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

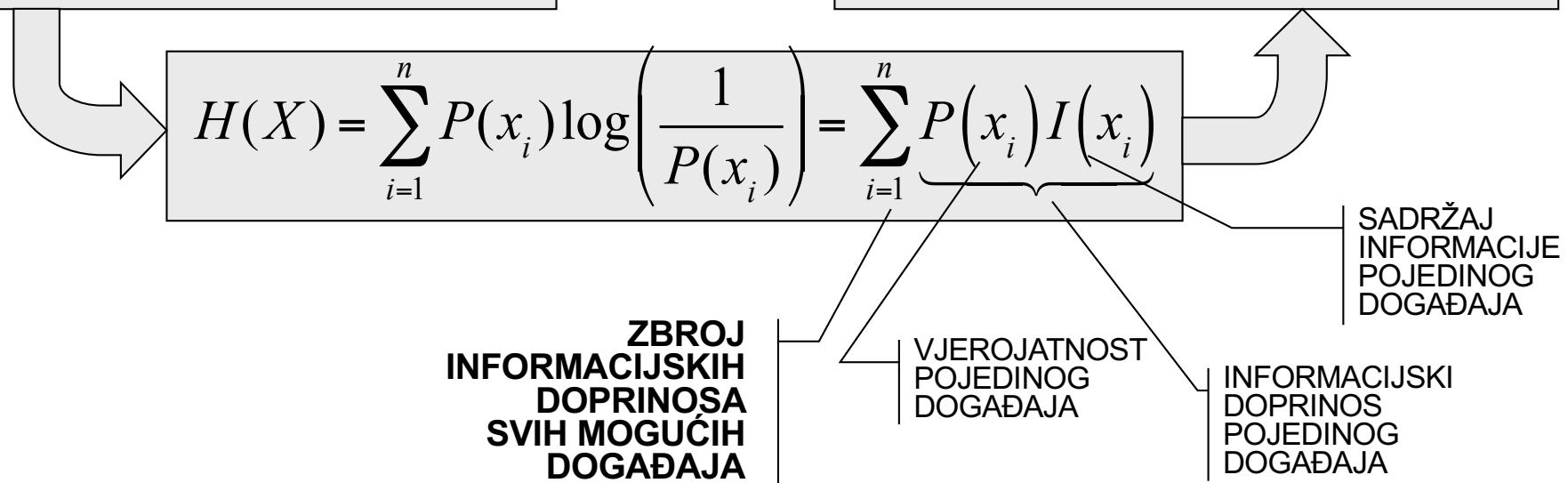
$$x_i \in X, X = \{x_1, x_2, \dots, x_n\}$$

VJEROJATNOSNI OPIS

vjerojatnosti: $P(x_i), i = 1, \dots, n$

INFORMACIJSKI OPIS

entropija: $H(X)$



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

Entropija na ulazu, izlazu, združena entropija

- ◆ Promatramo događaje na ulazu i izlazu odvojeno:

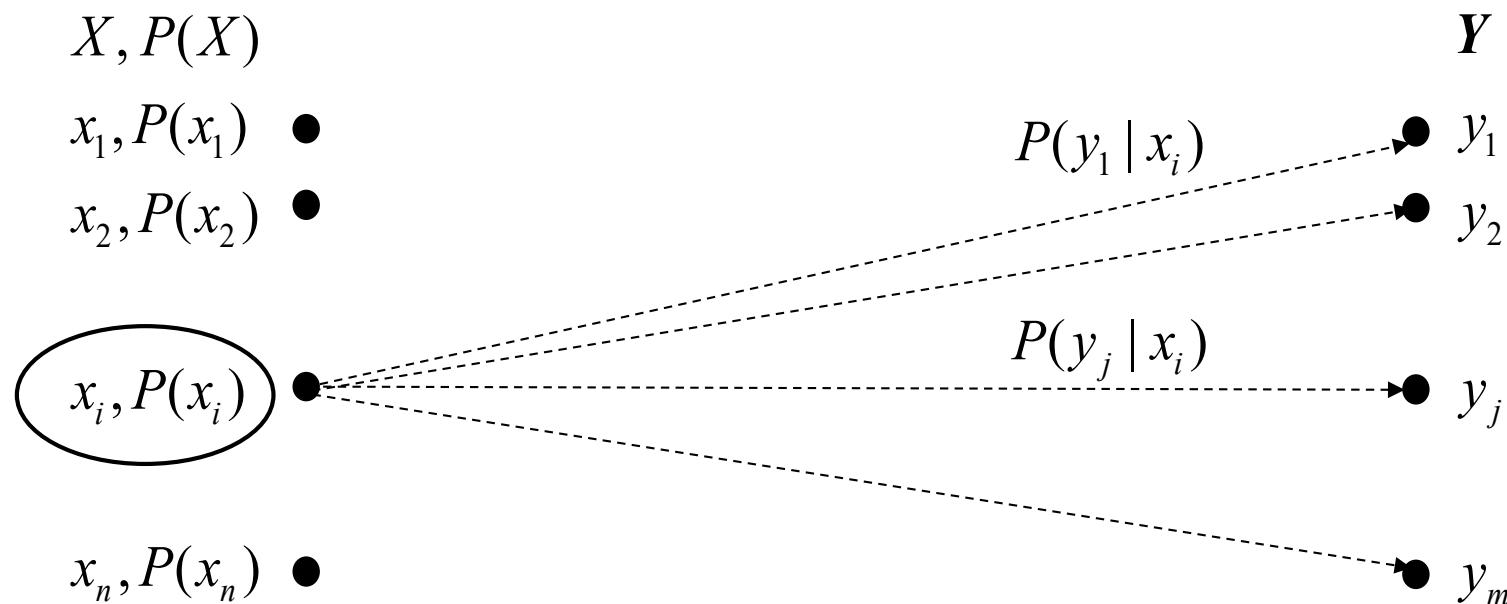
$$H(X) = -\sum_{i=1}^n P(x_i) \log(P(x_i)) \quad H(Y) = -\sum_{j=1}^m P(y_j) \log(P(y_j))$$

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

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

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



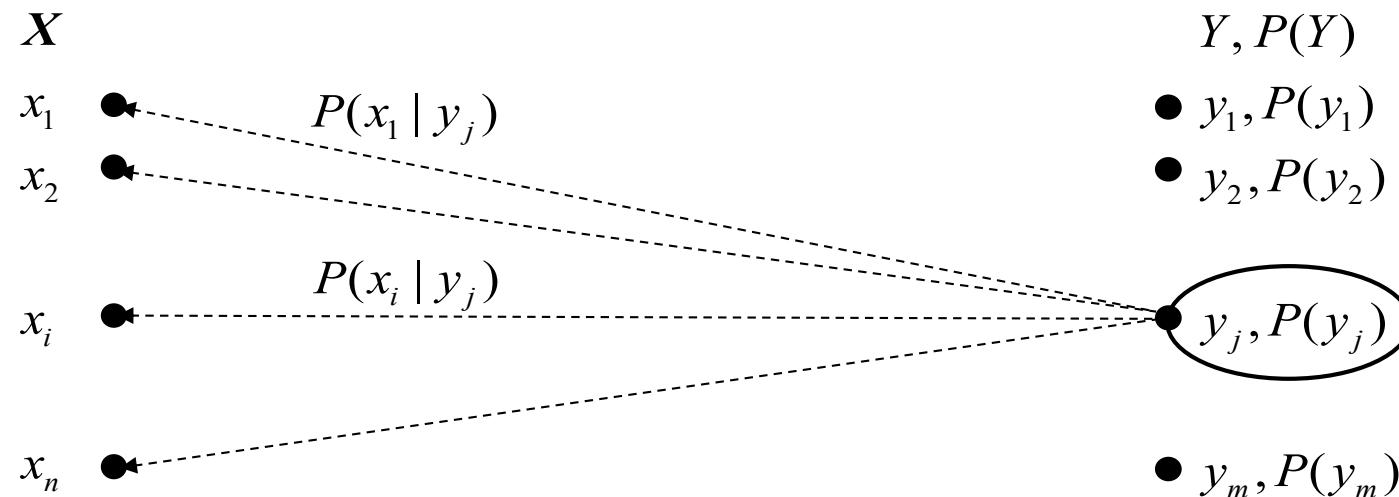
Entropija šuma (2)

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

$$\begin{aligned} H(Y | X) &= E\left(H(Y | x_i)\right) = \sum_{i=1}^n P(x_i)H(Y | x=x_i) \\ &= -\sum_{i=1}^n P(x_i) \sum_{j=1}^m P(y_j | x_i) \log(P(y_j | x_i)) \\ &= -\sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) \log(P(y_j | x_i)) \end{aligned}$$

Mnogoznačnost ili ekvivokacija

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



Ekvivokacija (2)

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

$$\begin{aligned} H(X | Y) &= E\left(H\left(X \middle| y_j\right)\right) = \sum_{j=1}^m P(y_j)H(X | y=y_j) \\ &= -\sum_{j=1}^m P(y_j) \sum_{i=1}^n P(x_i | y_j) \log(P(x_i | y_j)) \\ &= -\sum_{j=1}^m \sum_{i=1}^n P(x_i, y_j) \log(P(x_i | y_j)) \end{aligned}$$

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.

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

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

Uzajamni sadržaj informacije

- ♦ omjer aposteriorne i apriorne vjerojatnosti

$$I(x_i; y_j) = \log \left(\frac{P(x_i | y_j)}{P(x_i)} \right) = \log \left(\frac{P(x_i, y_j)}{P(x_i)P(y_j)} \right)$$

$$I(x_i; y_j) = I(y_j; x_i) = \log \left(\frac{P(y_j | x_i)}{P(y_j)} \right) = \log \left(\frac{P(x_i, y_j)}{P(x_i)P(y_j)} \right)$$

$$I(x_i; x_i) = \log \left(\frac{P(x_i | x_i)}{P(x_i)} \right) = \log \left(\frac{1}{P(x_i)} \right) = I(x_i)$$

$$I(x_i; y_j) \leq I(x_i)$$

$$I(x_i; y_j) \leq I(y_j)$$

◆ Definicija:

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

◆ Interpretacija:

- Koliko informacije jedna varijabla pruža o drugoj
- U kojoj mjeri su dvije varijable ovisne
 - Neovisne: $I(X;Y) = 0$
 - 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

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

- ako $H(X)$ predstavlja neizvjesnost X -a prije nego je poznat Y , a $H(X|Y)$ predstavlja neizvjesnost X -a nakon što je poznat Y , tada razlika $H(X) - H(X|Y)$ predstavlja količinu informacije koju Y pruža o X -u

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

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

Odnos entropije i uzajamnog sadržaja informacije - izvod

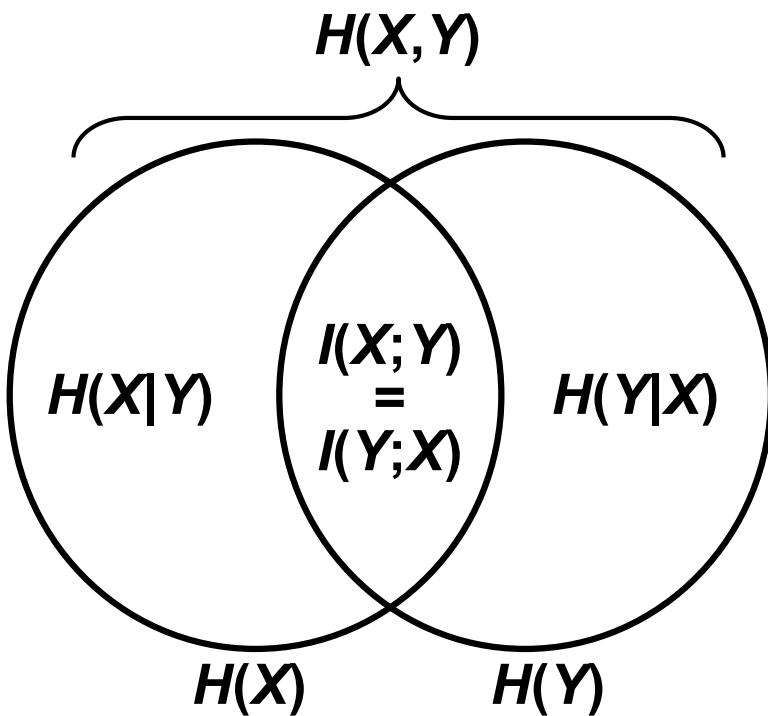
$$\begin{aligned} I(X;Y) &= \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log\left(\frac{p(x_i, y_j)}{p(x_i)p(y_j)}\right) = \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log\left(\frac{p(x_i | y_j)}{p(x_j)}\right) \\ &= \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log\left(\frac{1}{p(x_i)}\right) + \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log(p(x_i | y_j)) \\ &= \sum_{i=1}^n p(x_i) \log\left(\frac{1}{p(x_i)}\right) - \left(-\sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log(p(x_i | y_j))\right) \\ &= H(X) - H(X | Y) \end{aligned}$$

Vlastiti sadržaj informacije

- ◆ Uzajamni sadržaj informacije dviju varijabli je simetričan: $I(Y;X) = I(X;Y)$
- ◆ Uzajamni sadržaj informacije jedne varijable same sa sobom naziva se vlastiti sadržaj informacije
- ◆ Vlastiti sadržaj informacije slučajne varijable je u stvari njena entropija:

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

Odnosi i svojstva informacijskih mjera



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

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

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

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

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

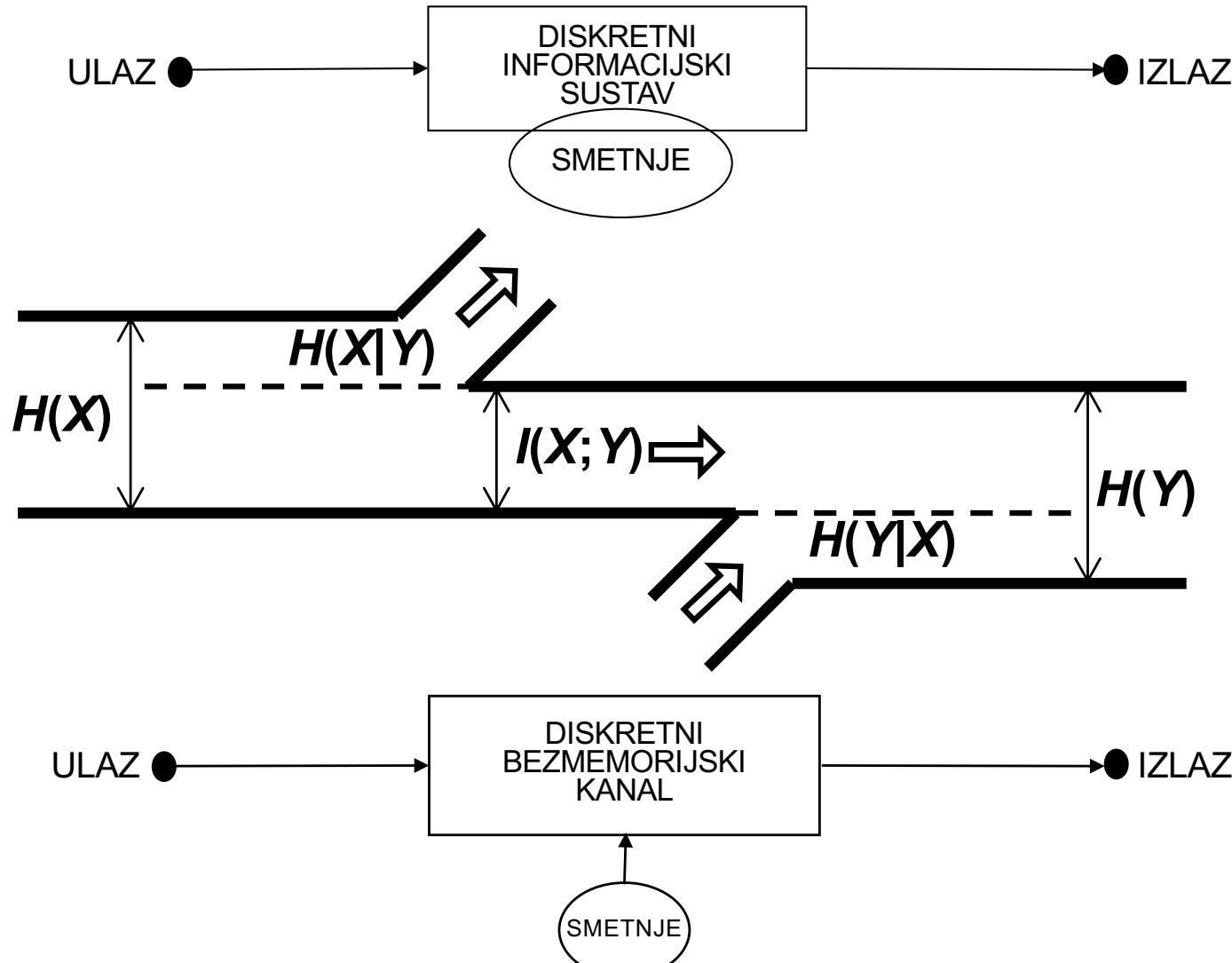
$$I(X; Y) = I(Y; X)$$

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

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

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

Prijenos informacije i informacijske mjere



Prijenos informacije i informacijske mjere

- ◆ Količina informacije koja se prenosi kanalom ovisi isključivo o:
 - karakteristikama kanala ($P(y_j|x_i)$)
 - karakteristikama ulaza ($P(x_i)$)

$$\begin{aligned} I(X;Y) &= \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log\left(\frac{p(x_i, y_j)}{p(x_i)p(y_j)}\right) = \sum_{i=1}^n \sum_{j=1}^m p(x_i)p(y_j | x_i) \log\left(\frac{p(y_j | x_i)}{p(y_j)}\right) \\ &= \sum_{i=1}^n \sum_{j=1}^m p(x_i)p(y_j | x_i) \log\left(\frac{p(y_j | x_i)}{\sum_{i=1}^n p(x_i, y_j)}\right) = \sum_{i=1}^n \sum_{j=1}^m p(x_i)p(y_j | x_i) \log\left(\frac{p(y_j | x_i)}{\sum_{i=1}^n p(x_i)p(y_j | x_i)}\right) \end{aligned}$$

- ◆ 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 $H(X, Y)$.

Rješenje primjera

- ◆ a) entropija ulaznog skupa simbola

$$H(X) = -\sum_{i=1}^3 P(x_i) \log(P(x_i)) = 1,522 \text{ [bit/simbol]}$$

- ◆ entropija izlaznog skupa simbola

$$H(Y) = -\sum_{i=1}^3 P(y_j) \log(P(y_j)) = 1,566 \text{ [bit/simbol]}$$

- ◆ b) šum

$$H(Y | X) = -\sum_{i=1}^3 \sum_{j=1}^3 P(x_i, y_j) \log(P(y_j | x_i)) = 1,157 \text{ [bit/simbol]}$$

Rješenje primjera (2)

- ekvivokacija

$$H(X | Y) = - \sum_{j=1}^3 \sum_{i=1}^3 P(x_i, y_j) \log(P(x_i | y_j))$$

- potrebno je odrediti vjerojatnosti $P(x_i | y_j)$

$$\left[P(x_i | y_j) \right] = \left[\frac{P(x_i, y_j)}{P(y_j)} \right] = \begin{bmatrix} P(x_1 | y_1) & P(x_1 | y_2) & P(x_1 | y_3) \\ P(x_2 | y_1) & P(x_2 | y_2) & P(x_2 | y_3) \\ P(x_3 | y_1) & P(x_3 | y_2) & P(x_3 | y_3) \end{bmatrix} = \begin{bmatrix} 0,7368 & 0,1111 & 0,3077 \\ 0,2105 & 0,7778 & 0,1538 \\ 0,0526 & 0,1111 & 0,5385 \end{bmatrix}$$

- provjera: zbroj elemenata po stupcu mora biti 1
 - zbroj u prvom stupcu je 0,9999 zbog zaokruživanja
- $H(X | Y) = 1,124$ [bit/simbol]

Rješenje primjera (3)

- ◆ c) uzajamni sadržaj informacije, tj. iznos informacije po simbolu koji se od izvora do odredišta prenosi neizmijenjen:
 - $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = 0,4095$ [bit/simbol]
 - ili
- ◆ d) $H(X,Y) = H(X) + H(Y) - I(X;Y) = 2,6788$ [bit/simbol]

$$I(X;Y) = \sum_{i=1}^3 \sum_{j=1}^3 P(x_i, y_j) \log \left(\frac{P(x_i, y_j)}{P(x_i)P(y_j)} \right)$$

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

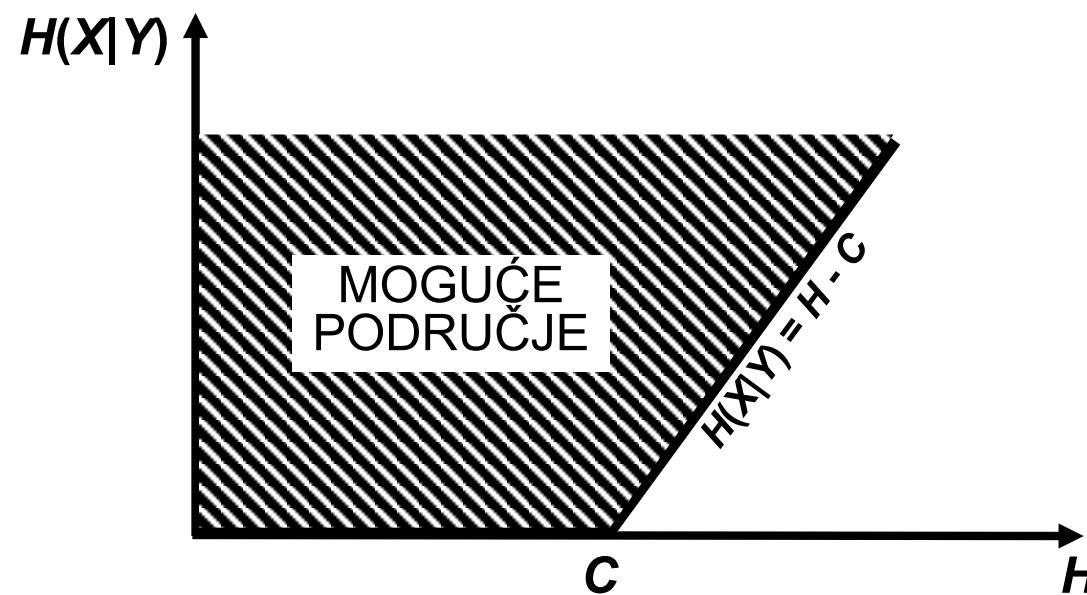
$$C = \max_{\{p(x_i)\}} I(X;Y) \text{ [bit/simbol]}$$

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

Temeljni teorem kanala sa smetnjama

Zavod za telekomunikacije

- ◆ 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

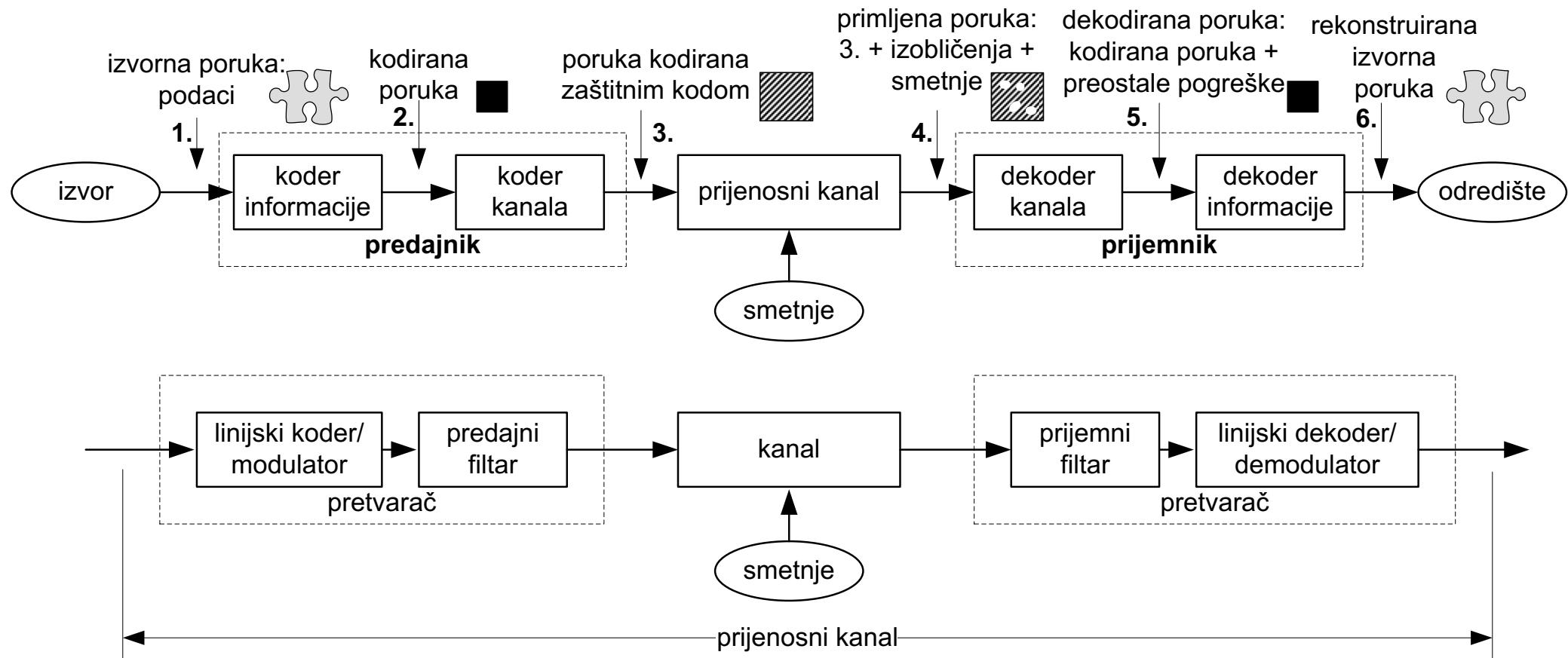


Kapacitet kanala i prijenosna brzina

$$C_T = \frac{1}{T} C \text{[bit/s]}$$

- ◆ T – trajanje simbola [s/simbol]
- ◆ apsolutna redundancija kanala: $C = I(X; Y)$
- ◆ učinkovitost kanala: $I(X; Y)/C$
- ◆ u praksi kapacitet kanala je uvijek veći od korištene prijenosne brzine
- ◆ kad bi prijenosna brzina bila veća od kapaciteta kanala sigurno bi imali pogreške u prijemu neovisno o odabranom kodu

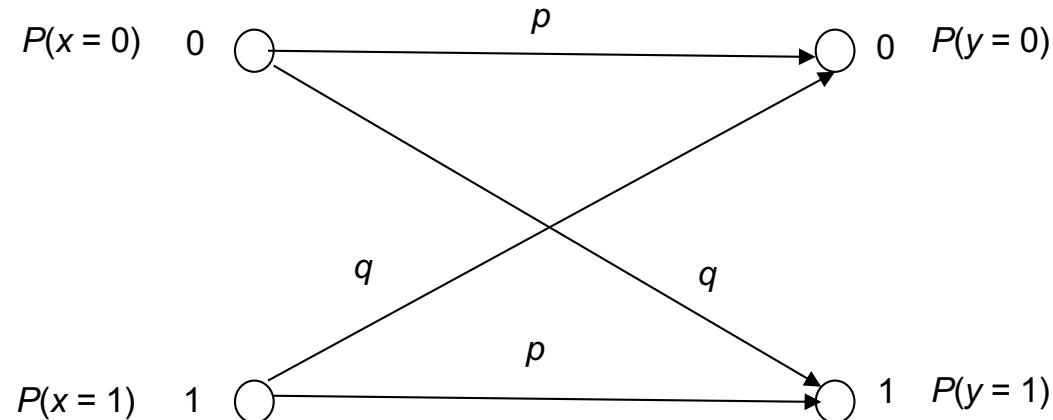
Prijenos informacije komunikacijskim sustavom



Teorija informacije

Osnovni pojmovi teorije informacije –
primjeri kanala

Binarni simetrični kanal (BSC)



$$\left[P(Y|X) \right] = \begin{bmatrix} p & q \\ q & p \end{bmatrix}$$

◆ $P(x = 0) = \alpha \quad P(x = 1) = 1 - \alpha$

◆ $P(x = 0) + P(x = 1) = 1$

◆ $P(y = 0|x = 0) = P(y = 1|x = 1) = p$

◆ $P(y = 0|x = 1) = P(y = 1|x = 0) = q$

◆ $p + q = 1$

Entropija šuma u BSC-u

$$H(X) = -\alpha \log(\alpha) - (1-\alpha) \log(1-\alpha)$$

$$H(Y|X) = - \sum_{i=1}^2 \sum_{j=1}^2 P(x_i, y_j) \log(P(y_j|x_i))$$

$$[P(X, Y)] = [P(X)] \cdot [P(Y|X)] = \begin{bmatrix} \alpha & 0 \\ 0 & 1-\alpha \end{bmatrix} \cdot \begin{bmatrix} p & q \\ q & p \end{bmatrix}$$

$$[P(X, Y)] = [P(y_j|x_i)] \cdot [P(x_i)] = \begin{bmatrix} p\alpha & q\alpha \\ q(1-\alpha) & p(1-\alpha) \end{bmatrix}$$

- ◆ zbroj po stupcu daje vjerojatnosti $P(y_j)$

$$P(y=0) = p\alpha + q(1-\alpha)$$

$$P(y=1) = q\alpha + p(1-\alpha)$$

- ◆ ako je $\alpha = 1 - \alpha = 1/2$, tada vrijedi $P(y=0) = P(y=1) = \frac{1}{2}$

Entropija šuma i transinformacija u BSC-u

◆ entropija šuma u BSC-u

$$H(Y|X) = - \sum_{i=1}^2 \sum_{j=1}^2 P(x_i, y_j) \log(P(y_j|x_i)) = \begin{bmatrix} P(X,Y) \end{bmatrix} = \begin{bmatrix} p\alpha & q\alpha \\ q(1-\alpha) & p(1-\alpha) \end{bmatrix}$$

$$= -p\alpha \log(p) - q\alpha \log(q) -$$

$$-q(1-\alpha) \log(q) - p(1-\alpha) \log(p) =$$

$$- \log(p)(p\alpha + p - p\alpha) - \log(q)(q\alpha + q - q\alpha) =$$

$$= -[p \log(p) + q \log(q)]$$

$$\begin{bmatrix} P(Y|X) \end{bmatrix} = \begin{bmatrix} p & q \\ q & p \end{bmatrix}$$

◆ transinformacija u BSC-u

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) = \\ &= H(Y) + p \log(p) + q \log(q) \end{aligned}$$

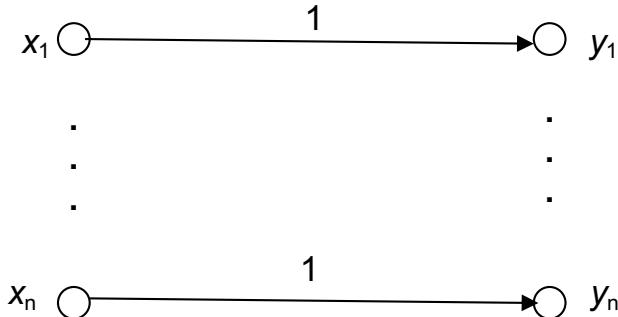
Kapacitet BSC-a

- ◆ kapacitet je jednak maksimalnom iznosu transinformacije

$$\begin{aligned} C &= \max_{P(x_i)} (I(X;Y)) = \max_{P(x_i)} (H(Y) - H(Y|X)) = \\ &= \max_{P(x_i)} (H(Y) + p \log(p) + q \log(q)) = \\ &= 1 + p \log(p) + q \log(q) \end{aligned}$$

- ◆ kapacitet BSC-a ovisi samo o p i q , a ne o vjerojatnostima $P(x_i)$
- ◆ razdioba?

Diskretan bešumni kanal



$$[P(Y|X)] = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

- za proračun kapaciteta kanala možemo iskoristiti matricu parova vjerojatnosti

$$P(x_i, y_j) = P(y_j | x_i) P(x_i) = \begin{cases} P(x_i) & \text{za } i = j \\ 0 & \text{za } i \neq j \end{cases}$$

$$P(x_i, y_j) = P(x_i | y_j) P(y_j) = \begin{cases} P(y_j) & \text{za } i = j \\ 0 & \text{za } i \neq j \end{cases}$$

$$P(x_i) = P(y_j) \text{ za } i = j$$

$$[P(X, Y)] = \begin{bmatrix} P(x_1, y_1) & 0 & \dots & 0 \\ 0 & P(x_2, y_2) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & P(x_n, y_n) \end{bmatrix}$$

Entropije u diskretnom bešumnom kanalu

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

$$= - \sum_{i=1}^n P(x_i, y_i) \log(P(x_i, y_i))$$

$$= H(X) = H(Y)$$

$$H(Y|X) = H(X, Y) - H(X) = 0$$

$$H(X|Y) = H(X, Y) - H(Y) = 0$$

- ◆ dakle u bešumnom kanalu i ekvivokacija je jednaka nuli

Transinformacija i kapacitet bešumnog kanala

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

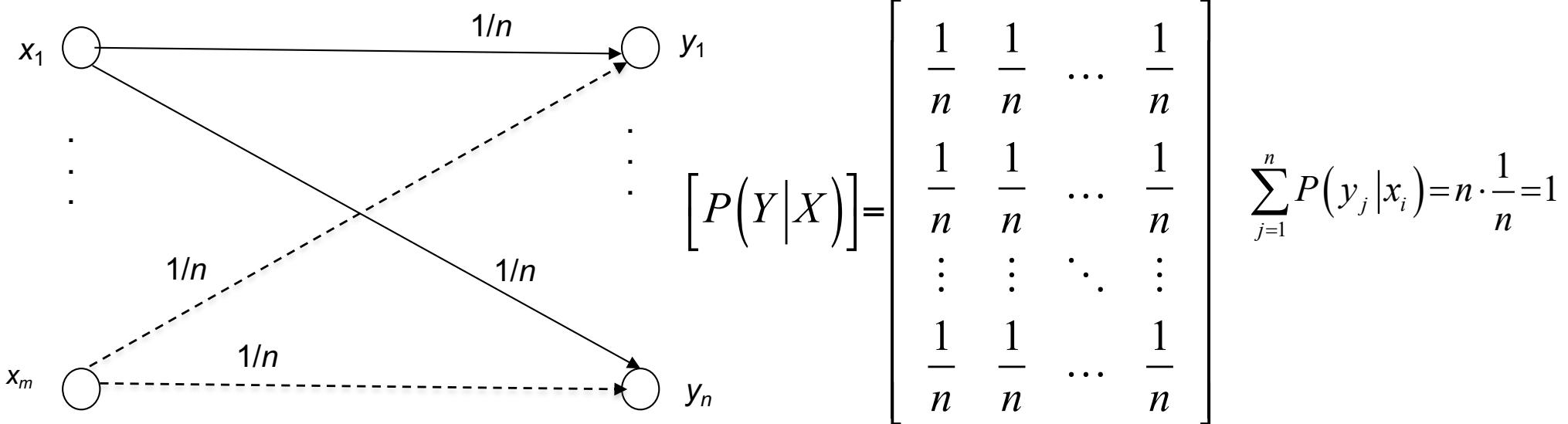
$$I(X;Y) = H(Y) - H(Y|X) = H(Y)$$

$$C = \max_{P(x_i)} (I(X;Y)) = \max_{P(x_i)} (H(X)) = \log(n)$$

- ◆ kapacitet bešumnog kanala jednak je entropiji izvora za slučaj kad su svi simboli izvora međusobno jednako vjerojatni
 - sva se informacija prenese od izvora do odredišta
 - nema gubitaka, šum i ekvivokacija jednaki su nuli

Diskretni kanal s međusobno neovisnim ulazom i izlazom

- ◆ nema korelacije između ulaznih i izlaznih simbola
- ◆ za svaki simbol na ulazu, x_i , vrijedi da može na izlazu biti primljen kao bilo koji od simbola y_j , pri čemu su svi prijelazi nekog simbola x_i u bilo koji simbol y_j međusobno jednakovjerojatni



Diskretni kanal s međusobno neovisnim ulazom i izlazom (2)

- ◆ matrica združenih vjerojatnosti prelaza

$$[P(X,Y)] = [P(X)] \cdot [P(Y|X)] =$$

$$= \begin{bmatrix} P(x_1) & 0 & \dots & 0 \\ 0 & P(x_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P(x_m) \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{n}P(x_1) & \frac{1}{n}P(x_1) & \dots & \frac{1}{n}P(x_1) \\ \frac{1}{n}P(x_2) & \frac{1}{n}P(x_2) & \dots & \frac{1}{n}P(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n}P(x_n) & \frac{1}{n}P(x_n) & \dots & \frac{1}{n}P(x_n) \end{bmatrix}$$

$$\sum_{i=1}^m P(x_i, y_j) = \frac{1}{n} \sum_{i=1}^m P(x_i) = \frac{1}{n} = P(y_j)$$

$$\sum_{j=1}^n P(x_i, y_j) = n \frac{1}{n} P(x_i) = P(x_i)$$

$$P(x_i, y_j) = P(x_i)P(y_j) = \frac{1}{n} P(x_i)$$

Entropije diskretnog kanala s međusobno neovisnim ulazom i izlazom

$$H(X) = - \sum_{i=1}^m P(x_i) \log(P(x_i))$$

$$H(Y) = - \sum_{i=1}^n P(y_j) \log(P(y_j)) = - \sum_{i=1}^n \frac{1}{n} \log\left(\frac{1}{n}\right) = n \frac{1}{n} \log(n) = \log(n)$$

$$\begin{aligned} H(Y|X) &= - \sum_{i=1}^m \sum_{j=1}^n P(x_i, y_j) \log(P(y_j|x_i)) = - \sum_{i=1}^m \sum_{j=1}^n \frac{1}{n} P(x_i) \log\left(\frac{1}{n}\right) = \\ &= - \sum_{i=1}^m n \frac{1}{n} P(x_i) \log\left(\frac{1}{n}\right) = \log(n) \end{aligned}$$

- ◆ transinformacija i kapacitet
 - s izvora do odredišta se ne prenosi informacija

$$I(X;Y) = H(Y) - H(Y|X) = 0$$

$$C = 0 \frac{\text{bit}}{\text{simbol}}$$

Slabo simetričan kanal (WSC)

- ◆ u općenitom slučaju proračun kapaciteta kanala je problem optimizacije nelinearne funkcije
 - moguće rješenje: Lagrangeovi multiplikatori
- ◆ samo u slučaju **simetričnih** i **slabo simetričnih** kanala kapacitet je moguće odrediti eksplicitno
- ◆ definicije:
 - kanal je **simetričan** ako su reci i stupci matrice kanala, $P[Y|X]$, permutacije jedni drugih
 - kanal je **slabo simetričan** ako su reci matrice kanala permutacije jedni drugih i zbroj vjerojatnosti po stupcu je jednak po svim stupcima $s_j = \sum_{x \in X} P(y_j | x_i), s_j = s_k \forall j, k \in \{1, \dots, m\}$

- ◆ X – skup ulaznih simbola, $\{x_1, \dots, x_n\}$

- ◆ Y – skup izlaznih simbola, $\{y_1, \dots, y_m\}$

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) = H(Y) - \sum_{x \in X} P(x_i) \sum_{y \in Y} P(y|x_i) \log \left(\frac{1}{P(y|x_i)} \right) \leq \\ &\leq \log(\text{card}(Y)) - \sum_{x \in X} P(x_i) \underbrace{H(Y|x_i)}_{\text{po } i\text{-tom retku}} \end{aligned}$$

- ◆ s obzirom da su svi reci permutacije npr. prvog,

vrijedi $H(Y|x_1) = H(Y|x_2) = \dots = H(Y|x_n) \equiv H(Y|x), n = \text{card}(X)$

$$\sum_{x \in X} P(x_i) H(Y|x_i) = \sum_{x \in X} P(x_i) H(Y|x) = H(Y|x)$$

$$I(X;Y) \leq \log(\text{card}(Y)) - H(Y|x)$$

Kapacitet WSC-a

- ◆ gornja granica od $I(X; Y)$ je kapacitet WSC-a ako ju je moguće postići odgovarajućom razdiobom ulaznih simbola
- ◆ neka je $P(x_i) = 1/n$, tada vrijedi

$$P(y_j) = \sum_{x \in X} P(y_j | x_i) P(x_i) = \frac{1}{n} \sum_{x \in X} P(y_j | x_i) = \frac{s_j}{n}$$

- ◆ s obzirom da je kanal slabo simetričan, vrijedi

$$\forall j \in \{1, \dots, m\} s_j = s$$

$$P(y_j) = \frac{s}{n}$$

Teorem o kapacitetu WSC-a

- ◆ budući da se maksimalna entropija postiže kad su svi simboli jednakovjerojatni, tada vrijedi da je $H(Y) = \log(\text{card}(Y))$ kad je $P(x_i) = 1/n$ za svaki i
- ◆ Teorem: za kapacitet simetričnog ili slabo simetričnog kanala vrijedi sljedeća relacija

$$C = \log(\text{card}(Y)) - H(Y|x)$$

- pri čemu je

$$H(Y|x) = \sum_{y \in Y} P(y|x_i) \log\left(\frac{1}{P(y|x_i)}\right)$$

- moguće izračunati za bilo koji redak i
- ◆ kapacitet kanala se postiže kad su ulazni simboli jednoliko raspodijeljeni, tj. $P(x_i) = 1/\text{card}(X)$

Primjer SC-a

- ◆ matrica binarnog simetričnog kanala (BSC-a)

$$[P(Y|X)] = \begin{bmatrix} p & q \\ q & p \end{bmatrix}$$

- ◆ temeljem teorema vrijedi

$$\log(\text{card}(Y)) = \log(2) = 1$$

$$H(Y|x) = -p \log(p) - q \log(q)$$

$$C = 1 + p \log(p) + q \log(q)$$

uz $P(x_i) = 1/2$

Primjer WSC-a

- ◆ razmatrajmo kanal čija je matrica zadana na sljedeći način:

$$[P(Y|X)] = \begin{bmatrix} \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & \frac{1}{6} \\ \frac{4}{1} & \frac{3}{4} & \frac{4}{1} & \frac{6}{1} \\ \frac{1}{4} & \frac{1}{6} & \frac{1}{4} & \frac{1}{3} \end{bmatrix}$$

- ◆ kanal je WSC jer su reci permutacija jedan drugog i zbroj vjerojatnosti po svim stupcima je jednak i iznosi $1/2$
- ◆ kapacitet kanala jednak je

$$C = \log(4) - H(Y|x) = 2 - 1,9591 = 0,0409 \frac{\text{bit}}{\text{simbol}}$$

Primjer kanala koji nije SC niti WSC

- ◆ razmotrimo kanal čija je matrica zadana izrazom

$$[P(Y|X)] = \begin{bmatrix} \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{3}{6} & \frac{6}{6} & \frac{6}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix}$$

- ◆ donji je redak permutacija prvog retka
- ◆ ali zbroj vjerojatnosti po stupcima nije jednak
 - u prvom stupcu $5/6$
 - u drugom stupcu $2/6$, tj. $1/3$
 - u trećem stupcu $5/6$
- ◆ dakle, kanal nije SC niti WSC