U N I V E R Z A V L J U B L J A N I
Fakulteta za elektrotehniko

Vitomir Štruc, Simon Dobrišek

INFORMACIJA IN KODI

DOPOLNILNI UČBENIK Z VAJAMI

UNIVERZITETNI ŠTUDIJSKI PROGRAM II. STOPNJE
ELEKTROTEHNIKA - AVTOMATIKA IN INFORMATIKA

PRVA IZDAJA

Ljubljana, 2016

CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

519.72(075.8)(0.034.2)
681.51(075.8)(0.034.2)

ŠTRUC, Vitomir, 1981-
Informacija in kodi [Elektronski vir] : dopolnilni učbenik z vajami / Vitomir Štruc, Simon Dobrišek. -
1. izd. - El. knjiga. - Ljubljana : Založba FE, 2016

Način dostopa (URL): http://luks.fe.uni-lj.si/nluks/wp-content/uploads/2016/02/IK.pdf

ISBN 978-961-243-303-1 (pdf)
1. Dobrišek, Simon
283860224

Copyright © 2015 2016 Založba FE. All rights reserved.
Razmnoževanje (tudi fotokopiranje) dela v celoti ali po delih
brez predhodnega dovoljenja Založbe FE prepovedano.

URL: http://luks.fe.uni-lj.si/nluks/wp-content/uploads/2016/02/IK.pdf

Recenzenta: prof. dr. Andrej Košir, doc. dr. Matej Kristan
Založnik: Založba FE, Ljubljana
Izdajatelj: UL Fakulteta za elektrotehniko, Ljubljana
Urednik: prof dr. Sašo Tomažič

1. izdaja

Predgovor

Dopolnilni učbenik z vajami predstavlja dopolnilno študijsko gradivo, ki je na razpolago študentom predmeta Informacija in Kodi na podiplomskem drugostopenjskem študijskem programu Elektrotehnika na Fakulteti za elektrotehniko Univerze v Ljubljani na študijski smeri Avtomatika in informatika. Učbenik je namenjen predvsem pripravam na pisno preverjanje pri predmetu Informacija in Kodi, pripraven pa je za vsakogar, ki bi s pomočjo preprostih primerov želel poglobiti svoje znanje in razumevanje teorije informacij in kodiranja.

Dopolnilni učbenik vsebuje krajše povzetke teoretičnega ozadja snovi predmeta, pojasnila ter rešitve izbranih primerov računskih nalog. Prav tako je v njem zajet nekoliko obširnejši pregled (z ilustrativnimi primeri) osnovnih pojmov teorije verjetnosti, ki dopolnjuje vsebino učbenika predmeta:

> *N. Pavešić: Informacija in kodi, (druga, spremenjena in dopolnjena izdaja), Založba FE in FRI, 2010.*

Pričujoči učbenik je sestavljen iz desetih poglavij in enega dodatka, ki smiselno sledijo vsebini učbenika in podajajo dodatne primere skupaj z rešitvami z namenom boljšega razumevanje snovi predmeta.

Avtorja se zahvaljujeta vsem sodelavcem in preteklim izvajalcem predmeta, ki so pripomogli k nastanku tega dopolnilnega učbenika. Posebna zahvala gre strokovnima recenzentoma Mateju Kristanu in Andreju Koširju ter Darji Mihelič, ki je učbenik lektorirala.

V Ljubljani, februarja 2016 *Vitomir Štruc in Simon Dobrišek*

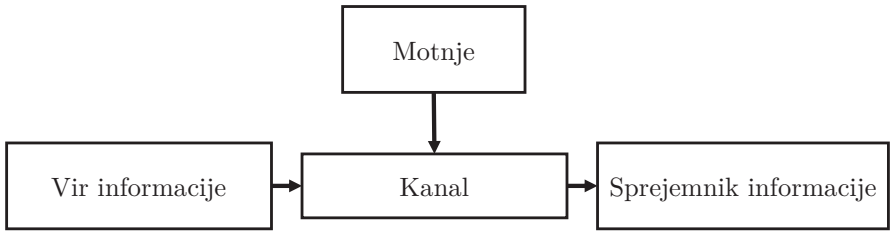
Kazalo

1 Uvod

Teorija informacij obravnava informacijsko-komunikacijski sistem, ki ga sestavljajo trije odprti, dinamični in naključni podsistemi. Ti trije sistemi so *vir informacije*, *kanal* in *sprejemnik informacije*. Sprejemnik je preko kanala, na katerega vplivajo *motnje*, povezan z virom informacije. Če je spreminjanje stanj sprejemnika odvisno od spreminjanju stanj vira, se med njima preko kanala prenaša informacija.

Teoretični model dinamičnih naključnih sistemov na splošno temelji na teoriji verjetnosti. Za tovrstne sisteme namreč predpostavljamo, da se v določenem trenutku nahajajo v enem izmed svojih možnih stanj, pri čemer je sprememba stanja sistema odvisna od *naključja*. V povezavi s tem tako razmišljamo o *verjetnosti*, da se sistem v danem trenutku nahaja v enem od možnih stanj.

Pri teoretičnem modelu komunikacijskega sistema predpostavljamo, da se stanja vira informacije preslikujejo v znake, ki se pojavljajo na vhodu kanala. Kanal vhodne znake nato preslikuje v znake na svojem izhodu, te preslikave pa so zaradi motenj v kanalu naključne. Pri sprejemniku na izhodu kanala se prejeti znaki zopet preslikajo, tokrat v spremembe stanj sprejemnika. Osnovni model predstavljenega komunikacijskega sistema je ponazorjen na sliki 1.1.



Slika 1.1: Model komunikacijskega sistema.

Med prenosom informacije se po navadi izvaja še dodatne preslikave zna-

kov med različnimi abecedami. Te dodatne preslikave se izvajajo zaradi tehniških zahtev udejanjenja komunikacijskega sistema, zaradi izboljšanja gospodarnosti pri prenosu informacije, zaradi večanja odpornosti na motnje ali zaradi zagotavljanja tajnosti prenešene informacije.

Pri obravnavi dinamičnih naključnih sistemov in naključnih preslikav med znaki se *teorija informacij* pri postavljanju teoretičnih modelov naslanja na verjetnostno teorijo in predvsem na teoretični model naključnih spremenljivk. Za boljše razumevanje podane snovi in lažje reševanje računskih nalog so v drugem poglavju zato povzeti osnovni pojmi iz verjetnostne teorije, predstavljeni so teoretični model naključnih spremenljivk ter pogosteje uporabljene diskretne verjetnostne porazdelitve. Drugo poglavje služi kot dopolnilno gradivo k snovi, podani v učbeniku predmeta.

V tretjem in četrtem poglavju sta predstavljena in s primeri ilustrirana povezana pojma entropije in informacije, ki sta ključnega pomena tako za teorijo informacij kot tudi za kodiranje, ki je naslovljeno v kasnejših poglavjih.

V petem poglavju je vpeljan pojem informacijskega vira. V tem poglavju je na primerih ponazorjen pomen različnih terminov, kot je ergodičnost, stacionarnost ali odvečnost informacijskega vira, predstavljene so različne vrste informacijskih virov ter njihove lastnosti.

Šesto poglavje se posveča kodiranju vira informacije in razloži osnovno terminologijo v zvezi kodiranjem, vpelje pogoje za obstoj različnih vrst kodov vira ter na primeru Huffmanovega koda predstavi zamisel izgradnje neenakomernih gospodarnih kodov vira. Računske naloge petega poglavja prikažejo razliko med kodiranjem posameznih simbolov in blokov simbolov ter posledice, ki jih takšen način kodiranja ima na gospodarnost končnega koda.

Sedmo poglavje se ukvarja s tajnim kodiranjem in najprej predstavi Viegenerjev kriptografski sistem ter njegove različice kot sta Cezarjev ali Vernamov kriptografski sistem. V drugem delu šestega poglavja je opisana ideja sodobnega kriptiranja z javnim ključem, ki je podrobneje ponazorjena na primeru RSA šifrirnega in dešifrirnega postopka.

V osmem poglavju je definiran matematični model komunikacijskega kanala, ki predstavlja podsistem celotnega komunikacijskega sistema in služi kot vezni člen med informacijskim virom in sprejemnikom sporočil. Na podlagi verjetnostnega modela komunikacijskega kanala so vpeljane različne vrste kanalov in poudarjene razlike med njimi. Prav tako so predstavljene osnove karakteristike kanalov, kot je njihova kapaciteta, njihov pomen pa je orisan

na primeru preprostih računskih nalog.

Deveto poglavje prikaže postopke za dekodiranje koda kanala ter osnove mehanizme odločanja pri dekodiranju, na podlagi katerih je mogoče popraviti (oz. vsaj odkriti) določene napake, ki se pojavijo v sporočilih zaradi motenj pri prenosu preko komunikacijskega kanala.

Zadnje, deseto poglavje se posveča postopkom varnega kodiranja. V tem poglavju so najprej razloženi osnovi koncepti za popravljanje in odkrivanje napak v kodih kot je preverjanje sodosti, le-ti pa so kasneje razširjeni na določanje in reševanje sistema linearno neodvisnih enačb nad Galoisovim obsegom GO(2).

Dopolnilni učbenik vsebuje dodatek, v katerem se nahajata primera dveh pisnih izpitov iz preteklih let. Pisna izpita ne vsebujeta rešitev in sta namenjena utrjevanju učne snovi.

2 Verjetnostna teorija

2.1 Verjetnostni račun

Teorija informacij predpostavlja obstoj naključnih sistemov, ki se v določenem trenutku nahajajo v enem izmed možnih stanj, ter obravnava medsebojni vpliv in delovanje tovrstnih sistemov. Teoretični model naključnega sistema vzpostavimo s pomočjo teorije verjetnosti in z uporabo teoretičnega modela naključnih spremenljivk.

Teorija informacij tudi sicer temelji na mnogih predpostavkah teorije verjetnosti. Pri nalogah se bomo tako spraševali o verjetnosti *napake* pri prenosu informacijskega znaka, o verjetnosti *zloma* kriptografskega sistema, o verjetnosti *oddaje znakov oz. nizov znakov* s strani virov informacij, ipd.

2.1.1 Verjetnostni poskus in dogodki

Osnova verjetnostnega računa je verjetnostni *poskus* ali *eksperiment*, pri katerem je rezultat odvisen od *naključja*. Elementarni rezultati verjetnostnega poskusa se imenujejo *elementarni izidi* ali *elementarni dogodki*. Verjetnostni poskus je torej določen z množico vseh možnih elementarnih izidov, ki jo po navadi označimo z Ω. Eden ali več elementarnih izidov, ki nas pri danem verjetnostnem poskusu posebej zanimajo, tvori *dogodek*. Dogodke označimo z velikimi tiskanimi črkami iz začetka abecede, npr. A, B, C, D, medtem ko verjetnostne poskuse označimo z malimi tiskanimi črkami iz začetka abecede, npr., a, b, c, itd.

Teoretični model verjetnostnega poskusa vključuje dva posebna primera dogodkov. *Nemogoč dogodek*, N, je dogodek, ki se nikoli ne zgodi in je predstavljen s prazno množico dogodkov $N = \emptyset = \{\}$. Na drugi strani je *gotov dogodek* G ki je definiran kot dogodek, ki se zgodi vedno in je zato predstavljen z univerzalno množico vseh možnih izidov poskusa. Vse preostale dogodke štejemo za *naključne dogodke*, ki se pri izvedbi poskusa lahko zgodijo ali ne [1].

***Primer* 2.1**

Klasični verjetnostni poskus je met igralne kocke. Množica vseh možnih izidov poskusa oz. elementarnih dogodkov je $\Omega = \{1, 2, 3, 4, 5, 6\}$. Primeri dogodkov, ki bi jih pri tem poskusu lahko obravnavali pa so [1]:

$$\begin{aligned} A &= \{6\} && \text{(pade šestica)} , \\ B &= \{2, 4, 6\} && \text{(pade sodo število pik)} , \\ C &= \{2, 3, 5\} && \text{(pade praštevilo pik)} , \\ D &= \{1, 2, 3, 4\} && \text{(pade manj kot pet pik)} , \\ E &= \{1, 3, 5\} && \text{(pade liho število pik)} . \end{aligned}$$

V predstavljenem primeru so dogodki A, B, C, D in E naključni dogodki, gotov dogodek pa je določen z $G = \Omega = \{1, 2, 3, 4, 5, 6\}$. Nemogoč dogodek $N = \emptyset = \{\}$ bi pri metu kocke pomenil, da ne pade nobeno število pik iz G.

Ker predstavljamo dogodke z množicami, katerih elementi so bodisi posamični elementarni izidi bodisi več elementarnih izidov skupaj, lahko v zvezi z njimi definiramo operacije. *Unija* dogodkov je dogodek, ki se zgodi, ko se zgodi eden ali drugi ali oba dogodka, in jo označimo z $\cup$. Na drugi strani predstavlja *presek* dogodkov dogodek, ki se zgodi, ko se zgodita oba dogodka. Presek po navadi označimo z $\cap$.

V povezavi z dogodki definirajmo še dva dodatna pojma. Dva dogodka sta *nezdružljiva*, če je njun presek nemogoč dogodek in sta si *nasprotna*, ko se vedno zgodi natanko eden od njiju. Unija nasprotnih dogodkov je gotov dogodek, njun presek pa je nemogoč dogodek. Za nasprotna dogodka pravimo tudi, da je eden dogodek *negacija* drugega dogodka [3].

Množica dogodkov $\{A_i\}$ pri $i = 1, ..., n$ sestavlja *popoln sistem* dogodkov, če se v danem poskusu vedno zgodi natanko eden od njih. Popoln sistem dogodkov je torej množica dogodkov, pri katerih velja [1]

$$A_i \neq N \quad \text{za vsak} \quad i = 1, ..., n \ ,$$

$$A_i \cap A_j = N \quad \text{za vsak} \quad i, j = 1, ..., n \quad \text{in} \quad i \neq j \ , \quad \text{ter}$$

$$\bigcup_{i=1}^{n} A_i = G \ .$$

Množica vseh elementarnih izidov poskusa je vedno popoln sistem dogodkov. V zgornjih enačbah pomenita i in j celoštevilska indeksa.

***Primer* 2.2**

Še enkrat se posvetimo poskusu meta kocke in dogodkom, ki so definirani v primeru 2.1. Na primeru teh dogodkov ponazorimo naslednje pojme [1]:

- *Unija dogodkov:* Unijo dogodkov B in C označimo z $B \cup C$ in je v našem primeru enaka $B \cup C = \{2, 3, 4, 5, 6\}$.
- *Presek dogodkov:* Presek dogodkov B in C označimo z $B \cap C$ in je v našem primeru enak $B \cap C = \{2\}$.
- *Nezdružljiva dogodka:* Med dogodki iz primera 2.1 predstavljata A in C primer nezdružljivih dogodkov, saj velja $A \cap C = N$.
- *Nasprotnost in negacija:* Med dogodki iz primera 2.1 sta nasprotna dogodka B in E, ker velja $B \cap E = N$ in $B \cup E = G$. Njuno medsebojno negacijo zapišemo tudi $B = \overline{E}$ in $E = \overline{B}$.
- *Popoln sistem dogodkov:* Primeri popolnih sistemov dogodkov pri metu kocke so:

$$\{A_i\} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$$
$$\{A_i\} = \{\{1,2\}, \{3,4\}, \{5,6\}\}$$
$$\{A_i\} = \{\{1,3,5\}, \{2,4,6\}\} \ .$$

2.1.2 Verjetnost dogodkov

Verjetnost je matematični model našega *verjetja* v to, ali se bo naključni dogodek zgodil ali ne. Verjetnost je določena s preslikavo, ki danemu dogodku priredi realno število, pri čemer prirejenemu realnemu številu pravimo *verjetnost dogodka*. Verjetnost dogodka A označimo s $P(A)$ [1].

Verjetnostna preslikava je določena z aksiomi Kolmogorova, ki zahtevajo [1], [2]:

1. *Nenegativnost:* $P(A) \geq 0$ za vsak A.
2. *Aditivnost:* Verjetnost dveh nezdružljivih dogodkov je enaka vsoti verjetnosti enega in drugega dogodka, torej

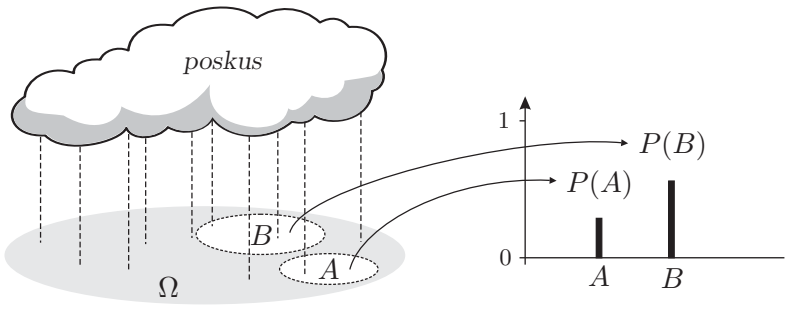
$$P(A \cup B) = P(A) + P(B) \Longleftarrow A \cup B = N \ . \tag{2.1}$$

Na splošno to velja za poljubno število medsebojno nezdružljivih dogodkov $A_1, A_2, ...$

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots . \tag{2.2}$$

3. *Normalizacijo:* Verjetnost celotne množice elementarnih dogodkov, $P(\Omega)$, je enaka 1.

Verjetnostna preslikava dogodkov je ponazorjena na sliki 2.1. Glede na to, da je gotov dogodek množica vseh elementarnih izidov poskusa, $G = \Omega$, je verjetnost gotovega dogodka $P(G) = 1$. Verjetnost nemogočega dogodka, $P(N)$, je enaka 0, kar lahko enostavno izpeljemo iz zgornjih pogojev.



Slika 2.1: Ponazoritev verjetnostne preslikave dogodkov v realno število.

Verjetnostni ali *porazdelitveni zakon* je verjetnostna preslikava vseh elementarnih dogodkov v realna števila. Iz porazdelitvenega zakona je mogoče določiti verjetnost poljubnega dogodka, ki je povezan z danim poskusom.

Porazdelitveni zakon določa verjetnostni model poskusa. Glede na to, kako je določena množica vseh možnih izidov poskusa, ločimo diskretne in zvezne verjetnostne modele.

2.1.3 Diskretni verjetnostni model

Diskretni verjetnostni model predpostavlja diskretno (števno) množico elementarnih izidov verjetnostnega poskusa. Diskretni porazdelitveni zakon, ki zadošča vsem verjetnostnim pogojem, določimo tako, da pripišemo vsakemu elementarnemu dogodku verjetnost med 0 in 1, pri čemer poskrbimo, da je vsota vseh verjetnosti enaka 1.

Če predpostavimo, da so vsi elementarni izidi danega poskusa enako verjetni, pravimo, da je porazdelitveni zakon *enakomeren*. Pri predpostavljenem enakomernem porazdelitvenem zakonu verjetnost poljubnega naključnega dogodka A določimo kot razmerje [1]

$$P(A) = \frac{m}{n} = \frac{\text{število ugodnih izidov za dogodek } A}{\text{število vseh možnih izidov poskusa}} \ .$$

Zgornji definiciji pravimo tudi klasična matematična (Laplaceova) definicija verjetnosti dogodka.

***Primer* 2.3**

Predstavljene pojme zopet ponazorimo na primeru meta kocke in dogodki iz primera 2.1. V skladu s klasično definicijo verjetnosti dogodkov so verjetnosti elementarnih dogodkov enake

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6} \ .$$

Vsota verjetnosti elementarnih dogodkov je enaka ena:

$$P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) + P(\{5\}) + P(\{6\}) = 1$$

in je zaradi nezdružljivosti elementarnih dogodkov tudi enaka verjetnosti gotovega dogodka

$$P(\{1, 2, 3, 4, 5, 6\}) = 1 \ .$$

Verjetnostna preslikava, s katero smo priredili verjetnosti elementarnih izidov našega poskusa, zadošča vsem aksiomom Kolmogorova (glej razdelek 2.1.2).

V skladu s klasično definicijo verjetnosti, lahko določimo tudi verjetnosti dogodkov, ki so za dogodke od A do E enake:

$$P(A) = \frac{1}{6} \ , \ P(B) = \frac{3}{6} \ , \ P(C) = \frac{3}{6} \ , \ P(D) = \frac{4}{6} \ , \ P(E) = \frac{3}{6} \ ,$$

kjer smo verjetnosti dogodkov B, C, D in E, ki so sestavljeni iz več elementarnih dogodkov (tj., predstavljajo unijo več elementarnih do-

godkov), določili kot

$$P(B) = P(\{2,4,6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) \ ,$$
$$P(C) = P(\{2,3,5\}) = P(\{2\}) + P(\{3\}) + P(\{5\}) \ ,$$
$$P(D) = P(\{1,2,3,4\}) = P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) \ , \text{ ter}$$
$$P(E) = P(\{1,3,5\}) = P(\{1\}) + P(\{3\}) + P(\{5\}) \ .$$

Verjetnost dogodka lahko ocenimo tudi z dejanskim izvajanjem poskusa in ugotavljanjem frekvenc dogodkov. V tem primeru lahko pridobimo statistično oceno verjetnosti dogodka, ki je določena kot [1]

$$P(A) = \frac{m}{n} = \frac{\text{število realizacij dogodka } A}{\text{število ponovitev poskusa}} \ .$$

V praksi smo se pogosto prisiljeni zatekati k statističnim ocenam verjetnosti, ker so teoretični modeli naključnih naravnih pojavov v celoti podani le izjemoma. Za razliko od klasične definicije verjetnosti, statistična definicija ne predpostavlja, da so elementarni dogodki enako verjetni.

Diskretni porazdelitveni zakon pogosto podajamo grafično v obliki histograma.

***Primer* 2.4**

Poglejmo si primer podajanja porazdelitvenega zakona za črke, ki se pojavljajo v slovenskih besedilih. V tem primeru za elementarne dogodke štejemo (naključno) izbiro posamezne črke pri nastajanju besedila, torej pri tipkanju oziroma pisanju.

Najprej moramo definirati množico vseh možnih elementarnih izidov. Denimo, da nas zanima zgolj 25 črk slovenske abecede (ne glede na velikost), da ločil ne obravnavamo ter da bomo vse preostale možne simbole (številke, posebne znake, itd.) povsem ignorirali. S tem imamo opravka s 25 možnimi elementarnimi izidi verjetnostnega poskusa, torej

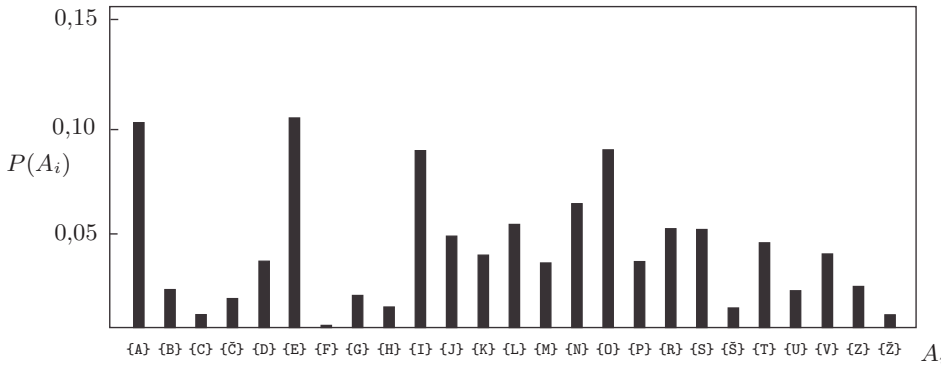
$$G = \{\{\texttt{'A'}\}, \{\texttt{'B'}\}, ..., \{\texttt{'Ž'}\}\} \ .$$

Pri oceni verjetnosti posameznih elementarnih dogodkov se moramo nujno zateči k statistični oceni verjetnosti. To izvedemo preprosto tako,

da preštejemo črke v danem besedilu in iz ugotovljenih frekvenc ocenimo verjetnosti. Denimo, da besedilo vsebuje 40000 črk in da smo našteli 4284 črk `'E'`. Dogodek pojavljanja te črke označimo z $E = \{\texttt{'E'}\}$. Ocenjena verjetnost $P(E)$ je tako

$$P(E) = P(\{\texttt{'E'}\}) = \frac{m}{n} = \frac{4284}{40000} \approx 0{,}1071 \ .$$

Enako storimo še za vse preostale črke. Z določitvijo verjetnosti vsake od črk smo določili celoten porazdelitveni zakon, ki ga lahko ponazorimo v obliki histograma kot je prikazano na sliki 2.2.



Slika 2.2: Verjetnostna porazdelitev pojavljanja črk v slovenskih besedilih v obliki histograma.

2.1.4 Zvezni verjetnostni model

Zvezni verjetnostni model predpostavlja zvezno (neštevno) množico elementarnih izidov verjetnostnega poskusa. Porazdelitvenega zakona v tem primeru ne moremo podati z verjetnostmi elementarnih izidov, zato porazdelitvenega zakona tudi ne moremo podati kot histogram verjetnosti elementarnih izidov poskusa.

Pri zveznem verjetnostnem modelu porazdelitveni zakon določamo s porazdelitvenimi funkcijami ene ali več realnih, neodvisnih spremenljivk. Porazdelitvene funkcije, ki se tipično uporabljajo pri zveznem verjetnostnem modelu, sta kumulativna porazdelitvena funkcija ali zbirna funkcija verjetnosti.

***Primer* 2.5**

Podzemna pripelje na postajo v natančno petnajst minutnih intervalih. V naključnem trenutku pride potnik na postajo. Zanima nas verjetnost, da bo čakal največ pet minut.

V predstavljenem primeru imamo opravka z zveznim verjetnostnim modelom, ker je zaradi poljubno natančnega merjenja časa (čas je poljubna realna vrednost) neskončno možnih trenutkov, v katerih bi po prihodu potnika vlak lahko pripeljal na postajo. Verjetnost, da vlak prispe na postajo točno 1,000... minuto po prihodu potnika na postajo, je enaka 0, enako pa velja tudi za vse ostale diskretne trenutke.

Verjetnost, da podzemna pripelje v manj kot petih minutah po prihodu potnika na postajo, lahko izračunamo kot razmerje med dolžino časovnega intervala čakanja, ki ga štejemo še za ugodnega za obravnavani dogodek, in dolžino celotnega časovnega intervala možnega čakanja. Verjetnost, da bo potnik čakal manj kot 5 minut je torej enaka 5 minut/15 minut oz. 1/3.

Pri zveznih verjetnostnih modelih verjetnosti dogodka ne moremo več računati z ugotavljanjem števila elementarnih dogodkov, ki so za dogodek ugodni, ker elementarnih dogodkov ne moremo več šteti. Namesto tega verjetnost računamo z integriranjem po zveznih področjih, ki so za dogodek ugodni, in z računanjem razmerja med vrednostjo tega integrala in vrednostjo integrala po celotnega možnega področja, ki pokriva vse možne izide verjetnostnega poskusa.

Zaradi večje zahtevnosti obravnave zveznih verjetnostnih modelov se bomo v nadaljevanju v glavnem omejili na diskretne verjetnostne modele.

2.1.5 Lastnosti verjetnosti

Splošni pogoj aditivnosti za verjetnost določa, da je verjetnost dveh nezdružljivih dogodkov enaka vsoti verjetnosti enega in drugega dogodka [1], [3]

$$A \cap B = N \implies P(A \cup B) = P(A) + P(B). \tag{2.3}$$

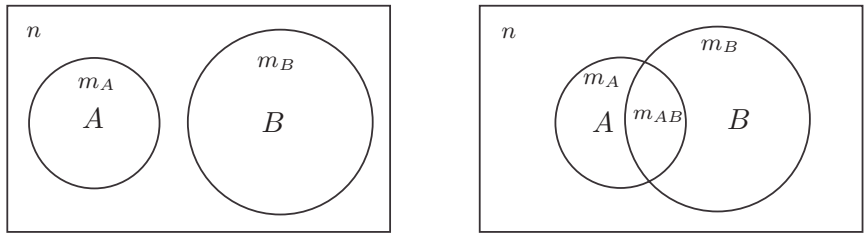
Verjetnost unije združljivih dogodkov A in B je v splošnem določena kot

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)\ . \tag{2.4}$$

Pri diskretnem verjetnostnem modelu z enako verjetnimi elementarnimi izidi poskusa lahko verjetnost unije dogodkov pridobimo iz razmerja med številom elementarnih izidov, ki so ugodni za eden (m_A), drugi (m_B) ali oba dogodka (m_{AB}), in številom vseh možnih elementarnih izidov poskusa (n):

$$P(A \cup B) = \frac{m_A + m_B - m_{AB}}{n} \ . \tag{2.5}$$

Pri nezdružljivih dogodkih je število m_{AB} enako nič. Razmerje med temi števili je razvidno tudi iz Vennovih diagramov, kot je ponazorjeno na sliki 2.3



Slika 2.3: Vennova diagrama za nezdružljiva (levo) in združljiva dogodka (desno).

2.1.6 Pogojna verjetnost

Pri pogojni verjetnosti predpostavljamo, da imamo na razpolago delno informacijo o izidu poskusa in se sprašujemo o manjkajoči informaciji. Takšen primer predstavlja situacija, ko vemo, da se je zgodil dogodek B in se sprašujemo po verjetnosti, da se zgodi tudi dogodek A.

Primeri vprašanj, pri katerih se pojavlja pogojna verjetnost, so podani v naslednjem seznamu [1].

- Pri metu kocke pade število pik, ki je sodo. Kako verjetno je, da je število pik tudi praštevilo?
- Pri igri ugibanja besed vemo, da je prva črka besede “d” in nas sedaj zanima, kako verjetno je, da je druga črka “a”.
- Kako verjetno je, da ima pacient neko bolezen, kljub temu, da je diagnostični test, ki ni povsem zanesljiv, dal negativni rezultat?
- Na radarju se je prikazala pika. Kako verjetno je, da gre za letalo?

Pri pogojni verjetnosti vemo, da se je zgodil elementarni izid, ki je ugoden za dogodek B, kako verjetno je, da je ta izid ugoden tudi za dogodek A. Prostor možnih izidov, ki nas še zanimajo in ki bi lahko bili ugodni za dogodek A, se tako skrči na možne izide, ki so ugodni za dogodek B.

Ta razmislek nas napelje na ugotovitev, da lahko pri enako verjetnih elementarnih izidih, opisano pogojno verjetnost določimo kot razmerje [1]

$$P(A \mid B) = \frac{m_{AB}}{m_A} = \frac{\text{število ugodnih izidov za dogodek } A \cap B}{\text{število ugodnih izidov za dogodek } B} \ .$$

Če imenovalec in števec podelimo z številom vseh možnih izidov verjetnostnega poskusa, dobimo zvezo

$$P(A \mid B) = \frac{m_{AB}/n}{m_A/n} = \frac{P(A \cap B)}{P(B)} \ .$$

Verjetnosti preseka dogodkov $P(A \cap B)$ pravimo tudi *vezana verjetnost* in govori o verjetnosti, da se dogodka A in B zgodita hkrati. Z upoštevanjem simetričnosti vezane verjetnosti pridemo do splošne zveze med *lastno* ($P(B)$ oz. $P(A)$), *pogojno* ($P(A \mid B)$ oz. $P(B \mid A)$) in *vezano verjetnostjo* $P(A \cap B)$:

$$P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A) \ .$$

***Primer* 2.6**

Vrnimo se k verjetnostnemu poskusu meta kocke in dogodkom iz primera 2.1. Če pri metu kocke predpostavimo, da pade število pik, ki je sodo, potem je verjetnost, da pade število pik, ki je hkrati tudi praštevilo, enaka

$$P(C \mid B) = \frac{m_{CB}}{m_B} = \frac{1}{3} \ ,$$

kjer smo upoštevali, da je med tremi sodimi števili pik eno število pik tudi praštevilo, oziroma preko zveze

$$P(C \mid B) = \frac{P(C \cap B)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3} \ ,$$

kjer smo upoštevali, da je verjetnost $P(C \cap B)$, da pade sodo praštevilo pik, enaka $1/6$ (med vsemi šestimi možnimi izidi meta je namreč takšen izid le eden, in sicer, ko je število pik enako dve) ter, da je verjetnost, da pade sodo število pik, enaka $1/2$.

2.1.7 Odvisnost in neodvisnost dogodkov

Pri pogojni verjetnosti lahko razmišljamo tudi o odvisnosti in neodvisnosti dveh dogodkov. Za dogodek A pravimo, da je neodvisen od dogodka B, če dejstvo, da se je zgodil dogodek B, ne spremeni verjetnosti, da se zgodi še dogodek A. To z drugimi besedami pomeni, da poznavanje izida dogodka B ne prinaša nobene informacije o dogodku A. V tem primeru velja [1]

$$P(A) = P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \Rightarrow \quad P(A \cap B) = P(A)P(B) \ .$$

Pri neodvisnih dogodkih je torej verjetnost preseka dogodkov enaka produktu verjetnosti posameznega dogodka.

Pri verjetnostnem poskusu, pri katerem mečemo kocko in kovanec ter en dogodek vežemo na izbrani izid meta kovanca in drugi dogodek na izbrani izid meta kocke, gre očitno za dva neodvisna dogodka, ker poznavanje enega od obeh izidov ne spremeni verjetnosti drugega dogodka.

2.1.8 Popolna verjetnost

O popolni verjetnosti razmišljamo pri dogodkih, ki jih povežemo z nekim popolnim sistemom dogodkov. Naj bo dogodek A povezan s popolnim sistemom dogodkov $\{H_i\}$, pri $i = 1, 2, ..., n$. Pri poznavanju vseh lastnih verjetnosti dogodkov H_i in pogojnih verjetnosti dogodka A glede na posamezni dogodek H_i, torej $P(A \mid H_i)$, je popolna verjetnost dogodka A določena kot verjetnost unije vseh presekov med dogodkom A in posameznimi dogodki H_i, torej [1]

$$P(A) = \sum_{i=1}^{n} P(A \cap H_i) = \sum_{i=1}^{n} P(H_i)P(A \mid H_i) \ .$$

***Naloga* 2.1**

V pekarni so na policah hlebci kruha enakih oblik. Kljub enaki obliki, pa imajo hlebci različno težo in so narejeni iz dveh različnih vrst moke (cenejše in dražje). Od 200 hlebcev kruha, ki jih pek proda na dan je 50 težjih in 150 lažjih. Vsi težji so narejeni iz dražje moke. Med lažjimi hlebci jih je 50 narejenih iz dražje moke, preostali so narejeni iz

cenejše moke. V prodajalno pride kupec, ki naključno izbere en hlebec. Kolikšna je verjetnost, da bo kupec izbral hleb iz dražje moke?

Rešitev:

Označimo dogodka, da je hlebec težak oz. lahek z $\{H_1, H_2\}$, kjer H_1 označuje dogodek, da je kupec izbral težji hlebec, in H_2 označuje dogodek, da je kupec izbral lažji hlebec kruha. Glede na to, da je hlebec lahko bodisi težak bodisi lahek in ne oboje hkrati, imamo opravka s popolnim sistemom dogodkov.

Označimo dogodek, da je kupec izbral hlebec narejen iz dražje moke z A. Zanima nas popolna verjetnost tega dogodka v povezavi z omenjenim popolnim sistemom dogodkov. Iz podanih podatkov lahko ocenimo verjetnosti:

$$P(H_1) = \frac{50}{200} = \frac{1}{4} \ ,$$

$$P(H_2) = \frac{150}{200} = \frac{3}{4} \ ,$$

$$P(A \mid H_1) = \frac{50}{50} = 1 \ ,$$

$$P(A \mid H_2) = \frac{50}{150} = \frac{1}{3} \ .$$

Iskana popolna verjetnost je potem enaka:

$$P(A) = \sum_{i=1}^{2} P(H_i)P(A \mid H_i) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{3} = 0{,}5 \ .$$

Verjetnost dogodka, da kupec izbere hlebec kruha iz dražje moke, je torej enaka 0,5.

2.1.9 Bayesov izraz

Bayesov izraz je kombinacija zveze med lastno, pogojno in vezano verjetnostjo ter izraza za popolno verjetnost [1]

$$P(H_k \mid A) = \frac{P(H_k)P(A \mid H_k)}{P(A)} = \frac{P(H_k)P(A \mid H_k)}{\sum_{i=1}^{n} P(H_i)P(A \mid H_i)} \ ,$$

kjer množica dogodkov $\{H_i\}$ določa popoln sistem dogodkov. Bayesov izraz nam omogoča sklepanje o verjetnosti kakšnega od dogodkov iz popolnega sistema dogodkov, pri danem dogodku A.

***Primer* 2.7**

Vrnimo se k primeru o pekarni iz naloge 2.1 (glej str. 14) in se tokrat vprašajmo o verjetnosti, da je kupec izbral lažji hlebec kruha iz dražje moke. Pri tem uporabimo enake oznake za dogodke, in sicer H_1 za dogodek, da je kupec izbral težji hleb, H_2 za dogodek, da je kupec izbral lažjega, in A za dogodek, da je kupec izbral hlebrc iz dražje moke. Zanima nas verjetnost $P(H_2 \mid A)$.

S pomočjo Bayesovega izraza lahko zapišemo

$$\begin{aligned} P(H_2 \mid A) &= \frac{P(H_2)P(A \mid H_2)}{P(A)} = \\ &= \frac{P(H_2)P(A \mid H_2)}{P(H_1)P(A \mid H_1) + P(H_2)P(A \mid H_2)} = \\ &= \frac{\frac{3}{4} \cdot \frac{1}{3}}{\frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{3}} = \\ &= 0{,}5 \ . \end{aligned}$$

Verjetnost, da je kupec izbral lažji hlebec kruha iz dražje moke, je enaka 0,5.

2.2 Naključne spremenljivke

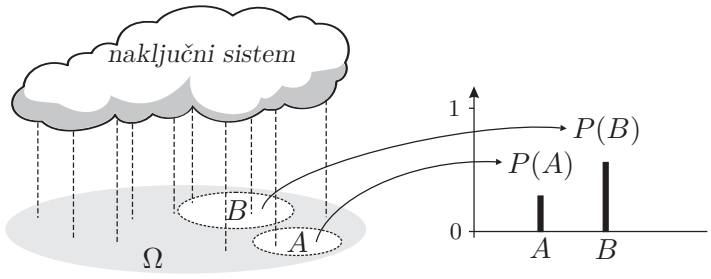
2.2.1 Diskretne naključne spremenljivke

Naključne spremenljivke so spremenljivke, katerih vrednost je odvisna od naključja. Matematično jih predstavimo kot preslikavo elementarnih izidov naključnega poskusa v realna števila (preslikava je ponazorjena na sliki 2.4), pri čemer bomo za *diskretne naključne spremenljivke* predpostavili, da je zaloga vrednosti spremenljivk diskretna. Množica realnih števil, v katera lahko preslikamo izide poskusov, je torej končna. Naključne spremenljivke po navadi označujemo z velikimi tiskanimi črkami s konca abecede, npr. X, Y, ali Z.

Zalogo vrednosti $\mathcal{Z}(X) = \{x_1, ..., x_i, ..., x_n\}$ diskretne naključne spremenljivke X običajno zapišemo skupaj z njeno porazdelitvijo verjetnosti $P_X = (p(x_1), ..., p(x_i), ..., p(x_n))$ v verjetnostni shemi [1]:

$$X \sim \begin{pmatrix} x_1, & x_2, & ..., & x_i, & ..., & x_n \\ p(x_1), & p(x_2), & ..., & p(x_i), & ..., & p(x_n) \end{pmatrix} .$$

Dejstvo, da diskretna naključna spremenljivka X zavzame prav vrednost x_i iz svoje zaloge vrednosti, predstavlja dogodek $\{X = x_i\}$, ki nastopi z verjetnostjo $p(x_i) = P(X{=}x_i) \geq 0$. Ker lahko naključna spremenljivka zavzame zgolj eno vrednost iz svoje zaloge vrednosti, so dogodki $\{X = x_i\}$ paroma nezdružljivi in tvorijo popoln sistem dogodkov, za katerega velja $\sum_{i=1}^{n} p(x_i) = 1$.



Slika 2.4: Ponazoritev naključne spremenljivke kot preslikave možnih stanj naključnega sistema v realno število.

Porazdelitveno funkcijo naključne spremenljivke lahko predstavimo bodisi kot porazdelitev verjetnosti $p(x_i) = P(X{=}x_i)$, za $i = 1, 2, ..., n$, bodisi s kumulativno funkcijo $F(x) = P(X{<}x) = \sum_{x_k<x} P(X{=}x_k)$.

V okviru teorije informacij nas bo zanimal model diskretnega naključnega sistema, ki se v danem trenutku nahaja v enem izmed n možnih stanj. Stanja bomo označili z $x_1, ..., x_i, ..., x_n$, verjetnost, da sistem preide v ta stanja v prihodnjem trenutku, pa s $p(x_1), ..., p(x_i), ..., p(x_n)$. Vidimo, da lahko teoretični model takšnega sistema predstavimo z diskretno naključno spremenljivko, pri čemer bi morali stanja naključnega sistema preslikati v realna števila, ki bi, denimo, predstavljala vrednost amplitude, frekvence, faznega zamika električnega signala ipd.

V našem primeru se bo torej namesto elementarnih izidov naključnega poskusa obravnavalo možna stanja naključnega sistema, kar teoretičnega modela naključne spremenljivke bistveno ne spremeni. Dejanska preslikava stanj sistema v realna števila nas pri *teoriji informacij* po navadi ne zanima, zato bomo zalogo vrednosti naključnih spremenljivk označevali z istimi oznakami, kot jih bomo uporabljali za stanja sistema.

Par naključnih spremenljivk

Par diskretnih naključnih spremenljivk bomo uporabljali za dva povezana naključna sistema. Zaloga vrednosti para naključnih spremenljivk je kartezični produkt zalog vrednosti obeh posameznih spremenljivk, ki smo ju vezali v par. Par diskretnih naključnih spremenljivk (X, Y) tako črpa iz zaloge vrednosti

$$\mathcal{Z}(X, Y) = \{(x_i, y_j) : \ i = 1, ..., m; \ j = 1, ..., n\}$$

v skladu z dvorazsežno vezano porazdelitvijo verjetnosti

$$P_{(X,Y)} = \left(p(x_i, y_j) \geq 0 : \ i = 1, ..., m; \ j = 1, ..., n; \ \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i, y_j) = 1\right) ,$$

kjer je $p(x_i, y_j) = P(X{=}x_i, Y{=}y_j)$ vezana verjetnost, da v danem trenutku spremenljivka X zavzame vrednost x_i in hkrati spremenljivka Y vrednost y_j. Lastne in pogojne verjetnosti zavzemanja vrednosti pri obeh naključnih spremenljivkah označimo s [4]

$$p(x_i) = P(X{=}x_i) \quad , \quad p(y_j) = P(Y{=}y_j) \ ,$$

$$p(x_i \mid y_j) = P(X{=}x_i \mid Y{=}y_j) \quad , \quad p(y_j \mid x_i) = P(Y{=}y_j \mid X{=}x_i) \ .$$

Z upoštevanjem splošne zveze med vezano, lastno in pogojno verjetnostjo ter lastnosti popolne verjetnosti [4]

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

lahko iz vezane porazdelitve verjetnosti izpeljemo *robni porazdelitvi* obeh spremenljivk, ki sta definirani z naslednjimi izrazi [4]

$$P_X = \left(p(x_i) \geq 0 : \ p(x_i) = \sum_{j=1}^{n} p(x_i, y_j);\ i = 1, ..., m;\ \sum_{i=1}^{m} p(x_i) = 1\right) \ ,$$

$$P_Y = \left(p(y_j) \geq 0 : \ p(y_j) = \sum_{i=1}^{m} p(x_i, y_j);\ j = 1, ..., n;\ \sum_{j=1}^{n} p(y_j) = 1\right)$$

ter pogojni porazdelitvi spremenljivk pri znani vrednosti druge spremenljivke

$$P_{X|Y=y_j} = \left(p(x_i \mid y_j) \geq 0 : \ p(x_i \mid y_j) = \frac{p(x_i, y_j)}{p(y_j)};\ i = 1, ..., m\right) \ ,$$

$$P_{Y|X=x_i} = \left(p(y_j \mid x_i) \geq 0 : \ p(y_j \mid x_i) = \frac{p(x_i, y_j)}{p(x_i)};\ j = 1, ..., n\right).$$

Pri tem ponovno velja $\sum_{i=1}^{m} p(x_i \mid y_j) = 1$ in $\sum_{j=1}^{n} p(y_j \mid x_i) = 1$.

***Naloga* 2.2**

Tri tovarne izdelujejo isti izdelek dveh kakovostnih razredov. Prva izdela v povprečju 10%, druga tovarna 15% in tretja samo 5% prvovrstnih izdelkov, preostali izdelki so drugega kakovostnega razreda. Trgovina ima na policah 600 izdelkov, ki izvirajo iz vseh treh tovarn, so obeh kakovostnih razredov in so pakirani v enotni embalaži. Med njimi je 100 izdelkov iz prve, 200 iz druge in 300 iz tretje tovarne. Predpostavimo, da se število prvovrstnih izdelkov pri vsaki tovarni sklada z omenjenimi verjetnostmi njihove proizvodnje. Kupec naključno izbere izdelek s police in ga zanima kakovost izdelka oziroma njegovo poreklo iz ene od treh tovarn. Predstavite ta naključni poskus s parom diskretnih naključnih spremenljivk.

Rešitev:

Iz besedila razberemo, da lahko opisani naključni poskus predstavimo s kombinacijo dveh diskretnih naključnih spremenljivk, pri čemer se ena nanaša na naključno izbiro tovarne, iz katere izvira izdelek, in druga na naključno izbiro kakovosti izdelka. Označimo njuni zalogi vrednosti z

$$\begin{aligned}\mathcal{Z}(X) &= \{x_1, x_2, x_3\} = \{1, 2, 3\} \sim \\ &\sim \{\texttt{'tovarna ena'}, \texttt{'tovarna dve'}, \texttt{'tovarna tri'}\} \,, \\ \mathcal{Z}(Y) &= \{y_1, y_2\} = \{1, 2\} \sim \\ &\sim \{\texttt{'je prvovrsten'}, \texttt{'ni prvovrsten}\} \,.\end{aligned}$$

Iz podatkov o številu izdelkov na polici lahko izračunamo lastne verjetnosti, da naključno izbrani izdelek izvira iz ene od treh tovarn

$$\begin{aligned}P_X &= (p(x_1)\,,\, p(x_2)\,,\, p(x_3)) = \left(\frac{100}{600}\,,\, \frac{200}{600}\,,\, \frac{300}{600}\right) = \\ &= \left(\frac{1}{6}\,,\, \frac{1}{3}\,,\, \frac{1}{2}\right) \approx (0{,}166\,,\, 0{,}333\,,\, 0{,}5) \,.\end{aligned}$$

Podatki o verjetnosti izdelave prvovrstnih izdelkov za posamezno tovarno nam podajajo pogojne verjetnosti za vse tri pogojne porazdelitve

$$\begin{aligned}P_{Y|X=x_1} &= (p(y_1 \mid x_1)\,,\, p(y_2 \mid x_1)) = (0{,}1\,,\, 0{,}9) \,, \\ P_{Y|X=x_2} &= (p(y_1 \mid x_2)\,,\, p(y_2 \mid x_2)) = (0{,}15\,,\, 0{,}85) \,, \\ P_{Y|X=x_3} &= (p(y_1 \mid x_3)\,,\, p(y_2 \mid x_3)) = (0{,}05\,,\, 0{,}95) \,.\end{aligned}$$

Z upoštevanjem zveze $p(x_i, y_j) = p(x_i)p(y_j \mid x_i)$ izračunamo vse vezane verjetnosti in jih podamo v vezanem porazdelitvenem zakonu

$$\begin{aligned}P_{(X,Y)} &= \begin{pmatrix} p(x_1, y_1)\,, & p(x_2, y_1)\,, & p(x_3, y_1) \\ p(x_1, y_2)\,, & p(x_2, y_2)\,, & p(x_3, y_2) \end{pmatrix} \approx \\ &\approx \begin{pmatrix} 0{,}166 \cdot 0{,}1\,, & 0{,}333 \cdot 0{,}15\,, & 0{,}5 \cdot 0{,}05 \\ 0{,}166 \cdot 0{,}9\,, & 0{,}333 \cdot 0{,}85\,, & 0{,}5 \cdot 0{,}95 \end{pmatrix} \approx \\ &\approx \begin{pmatrix} 0{,}016\,, & 0{,}049\,, & 0{,}025 \\ 0{,}149\,, & 0{,}283\,, & 0{,}475 \end{pmatrix} \,,\end{aligned}$$

kjer z znakom $\approx$ označimo, da je rezultat približen zaradi zaokroževanja.

Z upoštevanjem zveze $p(y_j) = \sum_{i=1}^{m} p(x_i, y_j)$ iz vezane porazdelitve verjetnosti določimo še drugo robno porazdelitev lastnih verjetnosti kakovosti izbranega izdelka

$$P_Y = \big(0{,}016 + 0{,}049 + 0{,}025\,,\ 0{,}149 + 0{,}283 + 0{,}475\big) \approx \\ \approx \big(0{,}0916\,,\ 0{,}9083\big)\ .$$

Z upoštevanjem zveze $p(x_i \mid y_j) = \frac{p(x_i, y_j)}{p(y_j)}$ nato izračunamo še pogojni porazdelitvi

$$P_{X|Y=y_1} \approx \Big(\frac{0{,}016}{0{,}0916}\,,\ \frac{0{,}049}{0{,}0916}\,,\ \frac{0{,}025}{0{,}0916}\Big) \approx (0{,}1818\,,\ 0{,}5454\,,\ 0{,}2727)\ ,$$

$$P_{X|Y=y_2} \approx \Big(\frac{0{,}149}{0{,}9083}\,,\ \frac{0{,}283}{0{,}9083}\,,\ \frac{0{,}475}{0{,}9083}\Big) \approx (0{,}1651\,,\ 0{,}3119\,,\ 0{,}5229)\ .$$

S tem smo določili vse verjetnostne porazdelitvene zakone, ki jih po navadi obravnavamo pri paru naključnih spremenljivk. Iz teh porazdelitvenih zakonov lahko razberemo odgovore na različna vprašanja, ki bi si jih lahko zastavili v zvezi z modeliranim verjetnostnim poskusom. Odgovor na vprašanje, "*Kako verjetno je, da je morebitni izbrani prvovrstni izdelek izdelala druga tovarna?*", tako daje verjetnost $P(X{=}x_2 \mid Y{=}y_1) = p(x_2 \mid y_1) = 0{,}5454$.

Niz naključnih spremenljivk

V okviru teorije informacij pogosto obravnavamo nize znakov, ki jih oddaja nek vir. Vire nizov znakov pri tem predstavimo z nizi naključnih spremenljivk, ki črpajo iz iste zaloge vrednosti, ter verjetnostmi oddaje teh znakov v diskretnih časovnih trenutkih.

Niz n-tih spremenljivk označimo z

$$(X_1, ..., X_t, ..., X_{n-2}, X_{n-1}, X_n)\ ,$$

in njihovo zalogo vrednosti kot

$$\mathcal{Z}(X_t) = A = \{x_1, ..., x_a\}\ ,$$

kjer A označuje t.i. abecedo vira in a moč abecede [4].

Naj $(x_{i_1}, ..., x_{i_n})$ označuje urejeno n-terico (oz. niz) znakov, ki črpa iz abecede vseh možnih nizov dolžine n znakov A^n. Vezano porazdelitev verjetnosti lahko potem zapišemo kot [4]

$$P_{X_1,...,X_n} = \left(p(x_{i_1}, ..., x_{i_n}) \geq 0 : (x_{i_1}, ..., x_{i_n}) \in A^n\right) ,$$

pri čemer velja

$$\sum_{(x_{i_1},...,x_{i_n})\in A^n} p(x_{i_1}, ..., x_{i_n}) = 1$$

in je

$$p(x_{i_1}, ..., x_{i_n}) = P(X_1{=}x_{i_1}, ..., X_n{=}x_{i_n}) .$$

Pri takšnih nizih naključnih spremenljivk nas zanimajo pogojne porazdelitve verjetnosti znakov glede na predhodne znake. Denimo, pogojna porazdelitev verjetnosti glede na dva predhodna znaka [4]:

$$P_{X_n|(X_{n-2}=x_{i_{n-2}},X_{n-1}=x_{i_{n-1}})} = \Big(p(x_{i_n} \mid (x_{i_{n-2}}, x_{i_{n-1}})) \geq 0 : \ i = 1, ..., a\Big) ,$$

kjer

$$p(x_{i_n} \mid (x_{i_{n-2}}, x_{i_{n-1}})) = P(X_n{=}x_{i_n} \mid (X_{n-2}{=}x_{i_{n-2}}, X_{n-1}{=}x_{i_{n-1}}))$$

označuje pogojno verjetnost i-tega znaka iz abecede A glede na dana predhodna znaka v nizu $(x_{i_{n-2}}, x_{i_{n-1}})$. Ta pogojna verjetnost je določena kot [4]

$$p(x_{i_n} \mid x_{i_{n-2}}, x_{i_{n-1}}) = \frac{p(x_{i_{n-2}}, x_{i_{n-1}}, x_{i_n})}{p(x_{i_{n-2}}, x_{i_{n-1}})} ,$$

kjer

$$p(x_{i_{n-2}}, x_{i_{n-1}}, x_{i_n}) = P(X_{n-2}{=}x_{i_{n-2}}, X_{n-1}{=}x_{i_{n-1}}, X_n{=}x_{i_n})$$

predstavlja vezano verjetnost niza treh znakov $(x_{i_{n-2}}, x_{i_{n-1}}, x_{i_n})$ in

$$p(x_{i_{n-2}}, x_{i_{n-1}}) = P(X_{n-2}{=}x_{i_{n-2}}, X_{n-1}{=}x_{i_{n-1}})$$

označuje vezano verjetnost niza (predhodnih) dveh znakov $(x_{i_{n-2}}, x_{i_{n-1}})$, ki jo lahko izpeljemo tudi iz zveze

$$p(x_{i_{n-2}}, x_{i_{n-1}}) = \sum_{i_n=0}^{a} p(x_{i_{n-2}}, x_{i_{n-1}}, x_{i_n}) .$$

Na podoben način lahko izpeljemo poljubno pogojno porazdelitev verjetnosti glede na dolžino predhodnih nizov znakov. Iz zapisanega vidimo, da

statistična ocena vezane porazdelitve nizov znakov izbrane dolžine zadošča za določitev verjetnostnega modela tovrstnega niza naključnih spremenljivk.

Statistično oceno verjetnostnih porazdelitev pridobimo s štetjem posameznih n-teric znakov v danem, končno dolgem nizu oddanih znakov.

Denimo, da nas zanimajo nizi črk dolžine 3 in da imamo na razpolago računalniški zapis daljšega besedila, sestavljenega samo iz teh črk. Statistično oceno vezane porazdelitve pridobimo s programskim ugotavljanjem frekvenc nizov dolžine 3, pri čemer se po datoteki besedila premikamo po en znak naprej od začetka do konca datoteke. Glede na to, da se bo vsaka črka iz besedila pojavila v nizih na prvem, drugem in tretjem mestu, lahko pričakujemo, da bodo robne porazdelitve, ki bi jih izpeljali iz vezane porazdelitve enake za vse tri vezane naključne spremenljivke v nizu, in da bodo te tri robne porazdelitve enake porazdelitvi, ki bi jo statistično ocenili z ugotavljanjem frekvenc posameznih črk. To sicer ni povsem res, ker majhna razlika med tremi robnimi porazdelitvami nastopi kot posledica dejstva, da prvi in zadnji dve črki v besedilu ne nastopita ne vseh treh mestih v obravnavanih nizih.

***Naloga* 2.3**

Vir oddaja črke slovenske abecede. Vzemimo, da smo črke priredili realnim številom po naslednji prevedbeni tabeli

A	B	C	Č	D	E	F	G	H	I	J	K	L
0	1	2	3	4	5	6	7	8	9	10	11	12
M	N	O	P	R	S	Š	T	U	V	Z	Ž	
13	14	15	16	17	18	19	20	21	22	23	24	25

in smo statistično ocenili, da je v oddanih nizih črk verjetnost niza treh črk ('V','E','M') enaka 0,000075. Ocenili smo še, da je verjetnost niza dveh črk ('V','E') enaka 0,007275. Kolikšna je verjetnost, da črka {'M'} sledi nizu dveh črk ('V','E')?

Rešitev:

Iz naloge lahko razberemo, da je verjetnost niza črk ('V','E','M') enaka 0,000075, kar lahko s pomočjo niza naključnih spremenljivk zapišemo kot $P(X_n{=}22 \mid (X_{n-2}{=}5, X_{n-1}13{=}$, kjer smo črke ('V','E','M') prevedli

v relana števila s pomočjo podane prevedbene tabele.

Verjetnost, niza treh črk zapišemo kot $P(X_{n-2}{=}22, X_{n-1}{=}5, X_n{=}13) = 0{,}000075$ in verjetnost niza dveh črk ('V','E') kot $P(X_{n-2}{=}22, X_{n-1}{=}5) = 0{,}007275$.

Iskano verjetnost potem preprosto določimo s pomočjo naslednjega izraza

$$P(X_n{=}13 \mid (X_{n-2}{=}22, X_{n-1}{=}5)) = \frac{P(X_{n-2}{=}22, X_{n-1}{=}5, X_n{=}13)}{P(X_{n-2}{=}22, X_{n-1}{=}5)} = \\ = \frac{0{,}000075}{0{,}007275} \approx 0{,}01 \ .$$

2.2.2 Odvisnost naključnih spremenljivk

Lastnost neodvisnosti naključnih spremenljivk izpeljemo iz lastnosti verjetnosti neodvisnih dogodkov. Pri neodvisnih naključnih spremenljivkah X in Y tako velja, da so vse vezane verjetnosti enake produktu ustreznih lastnih verjetnosti robnih porazdelitev obeh spremenljivk [1]:

$$p(x_i, y_j) = p(x_i) \cdot p(y_j);\ i = 1, ..., n;\ j = 1, ..., m \ .$$

Ta lastnost je posledica dejstva, da so pri neodvisnih spremenljivkah tudi pogojne porazdelitve enake robnim porazdelitvam, torej

$$P_X = P_{X|Y=y_j};\ j = 1, ..., m \ , \quad \text{in} \quad P_Y = P_{Y|X=x_i};\ i = 1, ..., n \ .$$

***Naloga* 2.4**

Na polici je 200 izdelkov, od tega 20 iz prve 40 iz druge in 140 iz tretje tovarne, pri čemer je iz prve tovarne 8, iz druge 16 in iz tretje 56 izdelkov prvovrstnih. Kupec pride v trgovino in naključno izbere izdelek. Denimo, da ve, kakšna je verjetnost, da ima v rokah izdelek iz posamezne tovarne. Ali lahko po ugotovitvi prvovrstnosti izdelka izve kaj več o tem, iz katere tovarne bi lahko bil izdelek? In obratno, denimo, da kupec ve, kakšna je verjetnost, da je izbral prvovrsten izdelek, ali lahko po ugotovitvi tega iz katere tovarne je izdelek, izve kaj več o tem, katere vrste bi lahko bil izbrani izdelek?

Rešitev:

Pred pridobitvijo informacije o vrsti izdelka oz. tovarni porekla ima kupec neko vnaprejšnje prepričanje o verjetnosti, da je izbrani izdelek prvovrsten ali ne, oziroma, da ga je izdelala ena od tovarn. Ko kupec ugotovi vrsto izdelka ali tovarno porekla, pa se lahko to prepričanje spremeni.

Na primer, ko kupec ugotovi vrsto izdelka, se lahko njegovo prepričanje o verjetnosti, da ga je izdelala ena od tovarn, spremeni. Velja tudi obratno, ko izve, katera tovarna je izdelek izdelala, se lahko spremeni njegovo prepričanje o verjetnosti prvovrstnosti izdelka. V tem primeru pravimo, da sta naključni spremenljivki, ki ju povezujemo z vrsto izdelkov oziroma s tovarnami porekla, med seboj odvisni, saj realizacija ene spremenljivke spremeni porazdelitev verjetnosti druge spremenljivke. Pri neodvisnih spremenljivkah nam vedenje o realizaciji ene od obeh spremenljivk ne pove ničesar novega o možni realizaciji druge spremenljivke.

Iz podatkov lahko ugotovimo vezano porazdelitev dveh naključnih spremenljivk, pri čemer z eno predstavimo tovarne, kjer izdelek izdelujejo, z drugo pa prvovrstnost izdelka. Naj bo ponovno

$$\begin{aligned}\mathcal{Z}(X) &= \{x_1, x_2, x_3\} = \{1, 2, 3\} \sim \\ &\sim \{\texttt{'tovarna ena'}, \texttt{'tovarna dve'}, \texttt{'tovarna tri'}\},\\ \mathcal{Z}(Y) &= \{y_1, y_2\} = \{1, 2\} \sim \\ &\sim \{\texttt{'je prvovrsten'}, \texttt{'ni prvovrsten}\}.\end{aligned}$$

Vezano porazdelitev verjetnosti obeh spremenljivk izračunamo iz podatkov, pri čemer upoštevamo, kolikokrat je izbira izdelka ugodna za posamezne kombinacije realizacij obeh naključnih spremenljivk

$$\begin{aligned}P_{(X,Y)} &= \begin{pmatrix} p(x_1, y_1)\,, & p(x_2, y_1)\,, & p(x_3, y_1) \\ p(x_1, y_2)\,, & p(x_2, y_2)\,, & p(x_3, y_2) \end{pmatrix} = \\ &= \begin{pmatrix} \frac{8}{200}\,, & \frac{16}{200}\,, & \frac{56}{200} \\ \frac{12}{200}\,, & \frac{24}{200}\,, & \frac{84}{200} \end{pmatrix} = \\ &= \begin{pmatrix} 0{,}04\,, & 0{,}08\,, & 0{,}28 \\ 0{,}06\,, & 0{,}12\,, & 0{,}42 \end{pmatrix}.\end{aligned}$$

Iz vezane porazdelitve lahko izpeljemo obe robni porazdelitvi

$$\begin{aligned} P_X &= \big(p(x_1)\,,\, p(x_2)\,,\, p(x_3)\big) = \\ &= \big(0{,}04 + 0{,}06\,,\, 0{,}08 + 0{,}12\,,\, 0{,}28 + 0{,}42\big) = \\ &= \big(0{,}1\,,\, 0{,}2\,,\, 0{,}7\big), \end{aligned}$$

$$\begin{aligned} P_Y &= \big(p(y_1)\,,\, p(y_2)\big) = \\ &= \big(0{,}04 + 0{,}08 + 0{,}28\,,\, 0{,}06 + 0{,}12 + 0{,}42\big) = \\ &= \big(0{,}4\,,\, 0{,}6\big). \end{aligned}$$

Ugotovimo, da so vezane verjetnosti enake produktu lastnih verjetnosti, kar pomeni, da sta obe spremenljivki neodvisni. Rezultat postane še bolj očiten, če podamo še vse pogojne porazdelitve

$$\begin{aligned} P_{X|Y=y_1} &= \Big(\frac{0{,}04}{0{,}4}\,,\, \frac{0{,}08}{0{,}4}\,,\, \frac{0{,}28}{0{,}4}\Big) = \big(0{,}1\,,\, 0{,}2\,,\, 0{,}7\big), \\ P_{X|Y=y_2} &= \Big(\frac{0{,}06}{0{,}6}\,,\, \frac{0{,}12}{0{,}6}\,,\, \frac{0{,}42}{0{,}6}\Big) = \big(0{,}1\,,\, 0{,}2\,,\, 0{,}7\big), \\ P_{Y|X=x_1} &= \Big(\frac{0{,}04}{0{,}1}\,,\, \frac{0{,}06}{0{,}1}\Big) = \big(0{,}4\,,\, 0{,}6\big), \\ P_{Y|X=x_2} &= \Big(\frac{0{,}08}{0{,}2}\,,\, \frac{0{,}12}{0{,}2}\Big) = \big(0{,}4\,,\, 0{,}6\big), \\ P_{Y|X=x_3} &= \Big(\frac{0{,}28}{0{,}7}\,,\, \frac{0{,}42}{0{,}7}\Big) = \big(0{,}4\,,\, 0{,}6\big). \end{aligned}$$

Vidimo, da so pogojne porazdelitve enake robnima, torej sta obe spremenljivki vzajemno neodvisni. Kupec torej ničesar novega ne izve o tovarni, če pozna kakovost izdelka in obratno.

2.2.3 Matematično upanje naključne spremenljivke

Matematično upanje $E(X)$ oz. srednjo vrednost diskretne naključne spremenljivke

$$X \sim \begin{pmatrix} x_1, & x_2, & \ldots, & x_i, & \ldots, & x_n \\ p(x_1), & p(x_2), & \ldots, & p(x_i), & \ldots, & p(x_n) \end{pmatrix}$$

definiramo kot

$$E(X) = \bar{x} = \sum_{i=1}^{n} p(x_i) x_i \ .$$

Matematično upanje je merilo centralne tendence diskretne naključne spremenljivke in se lahko tolmači kot delni opis verjetnostne porazdelitve naključne spremenljivke.

2.2.4 Standardni odklon in varianca

Varianca $V(X)$ diskretne naključne spremenljivke X je določena kot

$$V(X) = \sigma_X^2 = \sum_{i=1}^{n} (x_i - \mu)^2 p(x_i) = E\big((X-\mu)^2\big) \ .$$

Standardni odklon je definiran kot pozitivni kvadratni koren variance,

$$\sigma_X = +\sqrt{\sigma_X^2} \ .$$

Varianca je merilo za razpršenost naključne spremenljivke in se včasih lažje izračuna po naslednjem obrazcu

$$\sigma_X^2 = \sum_{i=1}^{n} x_i^2 p(x_i) - \Big(\sum_{i=1}^{n} x_i p(x_i)\Big)^2 = E(X^2) - E(X)^2 \ .$$

***Primer* 2.8**

Ponazorimo izračun matematičnega upanja in variance na primeru meta kocke. Met kocke lahko ponazorimo z naključno spremenljivko X, ki ima naslednjo verjetnostno shemo

$$X \sim \begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6 \\ \frac{1}{6}, & \frac{1}{6}, & \frac{1}{6}, & \frac{1}{6}, & \frac{1}{6}, & \frac{1}{6} \end{pmatrix} \ .$$

Matematično upanje in varianco naključne spremenljivke izračunamo kot

$$\begin{aligned} E(X) = \bar{x} &= \sum_{i=1}^{n} p(x_i) x_i = \\ &= \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 2 + \frac{1}{6}\cdot 3 + \frac{1}{6}\cdot 4 + \frac{1}{6}\cdot 5 + \frac{1}{6}\cdot 6 = \\ &= \frac{7}{2} \ , \end{aligned}$$

$$V(X) = \sigma_X^2 = E(X^2) - E(X)^2 = \\ = \frac{1}{6} \cdot 1^2 + \frac{1}{6} \cdot 2^2 + \frac{1}{6} \cdot 3^2 + \frac{1}{6} \cdot 4^2 + \frac{1}{6} \cdot 5^2 + \frac{1}{6} \cdot 6^2 - \mu^2 = \\ = \frac{35}{12} \ .$$

Matematično upanje lahko tolmačimo kot pričakovano srednjo vrednost naključne spremenljivke v primeru, ko bi verjetnostni poskus ponavljali več krat. V predstavljenem primeru, bi srednja vrednost ponavljanja poskusa meta kocke konvergirala proti $\frac{7}{2}$. Varianca, po drugi strani, pove, kako so meti v povprečju razpršeni okrog pričakovane srednje vrednosti.

2.3 Pomembnejše verjetnostne porazdelitve

Naključne spremenljivke so določene z njihovo zalogo vrednosti ter verjetnostno porazdelitvijo, ki daje informacijo o verjetnosti, da naključna spremenljivka zavzame določeno vrednost iz svoje zaloge vrednosti. Odvisno od naključnega poskusa (oz. naključnega sistema), ki ga želimo modelirati z naključno spremenljivko, izberemo ustrezno verjetnostno porazdelitev, pri čemer izbiramo med porazdelitvami predstavljenimi v nadaljevanju.

2.3.1 Bernoullijeva porazdelitev

Bernoullijeva porazdelitev se uporablja za opis diskretnih naključnih spremenljivk, ki lahko zavzamejo zgolj dve možni vrednosti. Tipični primer takšne naključne spremenljivke predstavlja izid meta kovanca, kjer lahko naključna spremenljivka zavzame vrednost 0, če pade grb, in vrednost 1, če pade cifra (ali obratno). Bernoullijevo naključno spremenljivko X opišemo kot [1]

$$X = \begin{cases} x_1, & \text{v prvem primeru (npr. cifra)} \\ x_2, & \text{v drugem primeru (npr. grb)} \ , \end{cases}$$

njena verjetnostna porazdelitev pa je določena z

$$P_X = \begin{cases} p, & \text{ko je } X = x_1 \\ 1-p, & \text{ko je } X = x_2 \ . \end{cases}$$

Verjetnostna shema Bernoullijeve naključne spremenljivke je torej enaka

$$X \sim \begin{pmatrix} x_1, & x_2 \\ p(x_1), & p(x_2) \end{pmatrix} = \begin{pmatrix} 0, & 1 \\ p, & 1-p \end{pmatrix} .$$

2.3.2 Binomska porazdelitev

Binomsko porazdelitev uporabimo za opis n zaporednih poskusov z zgolj dvema možnima izidoma. Primer naključnega poskusa, ki ga po navadi opišemo z binomsko porazdelitvijo, predstavlja n zaporednih metov kovanca, kjer nas, denimo, zanima verjetnost, da bo v n metih, padla cifra natanko k krat. Zaloga vrednosti binomske naključne spremenljivke, s katero predstavimo naključni poskus, torej predstavlja množica $\{0, 1, ..., k, ..., n\}$, pripadajoče verjetnosti pa lahko izračunamo v skladu z naslednjo enačbo [1]

$$p(x_k) = \binom{n}{k} p^k (1-p)^{n-k},$$

kjer p označuje verjetnost, da se bo v dani ponovitvi poskusa zgodil prvi od dveh možnih izidov (npr. pade cifra), $1-p$ označuje verjetnost, da se bo v dani ponovitvi poskusa zgodil drugi od dveh možnih izidov (npr. pade grb), in pomeni $k = 0, 1, ..., n$.

Binomska porazdelitev ima dva parametra, n in p, in jo pogosto označimo kot $b(n, p)$. Matematično upanje naključne spremenljivke, ki se podreja binomski porazdelitvi, je np in njena varianca je $np(1-p)$.

Verjetnostno porazdelitev binomske naključne spremenljivke zapišemo kot

$$X \sim \begin{pmatrix} x_0, & x_1, & ..., & x_k, & ..., & x_n \\ p(x_0), & p(x_1), & ..., & p(x_k), & ..., & p(x_n) \end{pmatrix} =$$

$$= \begin{pmatrix} 0, & 1, & ..., & k, & ..., & n \\ \binom{n}{0} p^0 (1-p)^{n-0}, & \binom{n}{1} p^1 (1-p)^{n-1}, & ..., & \binom{n}{k} p^k (1-p)^{n-k}, & ..., & \binom{n}{n} p^n (1-p)^0 \end{pmatrix}.$$

2.3.3 Enakomerna porazdelitev

Enakomerno porazdelitev uporabimo za naključne spremenljivke, ki opisujejo verjetnostne poskuse z izidi, ki so enako verjetni. Primer takšnega poskusa je met kocke, kjer lahko naključna spremenljivka zavzame poljubno vrednost iz svoje zaloge vrednosti $\{1, 2, 3, 4, 5, 6\}$ z verjetnostjo $\frac{1}{6}$.

Naključna spremenljivka, ki se podreja enakomerni porazdelitvi ima v splošnem n možnih stanj, torej $\mathcal{Z}(X) = \{1, 2, ..., n\}$, ki so vsa enako verjetna $p(x_i) = P(X{=}x_i) = \frac{1}{n}$. Verjetnostno shemo enakomerno porazdeljene naključne spremenljivke zapišemo kot

$$X \sim \begin{pmatrix} x_1, & x_2, & ..., & x_n \\ p(x_1), & p(x_2), & ..., & p(x_n) \end{pmatrix} = \begin{pmatrix} 1, & 2, & ..., & n \\ \frac{1}{n}, & \frac{1}{n}, & ..., & \frac{1}{n} \end{pmatrix} .$$

Matematično upanje naključne spremenljivke je (aritmetična sredina):

$$E(X) = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \ .$$

2.3.4 Poissonova porazdelitev

Poissonova porazdelitev je sorodna binomski porazdelitvi in se uporablja za opis naključnih poskusov oz. sistemov, pri katerih nas zanima verjetnost nastopa določenega dogodka (izida, stanja) znotraj neke v naprej predpisane časovne enote.

Primer Poissonove naključne spremenljivke predstavlja število klicev v klicni center mestne taksi službe, kjer lahko naključna spremenljivka (število klicev), zavzame poljubno vrednost iz neskončne zaloge vrednosti $\mathcal{Z}(X) = \{0, 1, 2, ..., k, ...\}$ z verjetnostjo [1], [3]:

$$p(x_k) = \frac{\lambda^k e^{-\lambda}}{k!} \ ; \ \lambda > 0 \ ,$$

pri čemer je k število (relevantnih) izidov v dani časovni enoti in je λ povprečno število izidov na časovno enoto. Vidimo lahko, da je λ parameter Poissonove porazdelitve, ki hkrati predstavlja tudi matematično upanje in varianco Poissonove naključne spremenljivke.

Verjetnostno shemo Poissonove naključne spremenljivke zapišemo kot

$$X \sim \begin{pmatrix} x_0, & x_1, & ..., & x_k, & ... \\ p(x_0), & p(x_1), & ..., & p(x_k), & ... \end{pmatrix} =$$
$$= \begin{pmatrix} 0, & 1, & ..., & k, & ... \\ p(x_0) = \frac{\lambda^0 e^{-\lambda}}{0!}, & p(x_1) = \frac{\lambda^1 e^{-\lambda}}{1!}, & ..., & p(x_k) = \frac{\lambda^k e^{-\lambda}}{k!}, & ... \end{pmatrix} .$$

2.3.5 Geometrijska porazdelitev

Geometrijska porazdelitev se uporablja za opis naključnih poskusov, pri katerem sta možna dogodka A in $\overline{A}$ z verjetnostma p in $1-p$ in nas zanima število potrebnih ponovitev poskusa k, da se zgodi dogodek A. Primer takšnega poskusa zopet predstavlja met kovanca, pri čemer se sprašujemo o verjetnosti, da v k-ti ponovitvi poskusa pade cifra. Naključna spremenljivka, ki se podreja geometrijski porazdelitvi, zopet črpa iz neskončne zaloge vrednosti $\mathcal{Z}(X) = \{0, 1, 2, ..., k, ...\}$, tokrat z verjetnostjo [3]:

$$p(x_k) = p(1-p)^{k-1} \ .$$

Verjetnostno shemo geometrijske naključne spremenljivke zapišemo kot

$$\begin{aligned} X \sim & \begin{pmatrix} x_1, & x_2, & ..., & x_k, & ... \\ p(x_1), & p(x_2), & ..., & p(x_k), & ... \end{pmatrix} = \\ = & \begin{pmatrix} 1, & 2, & ..., & k, & ... \\ p(1-p)^0, & p(1-p)^1, & ..., & p(1-p)^{k-1}, & ... \end{pmatrix} \ . \end{aligned}$$

***Naloga* 2.5**

Informacijski vir oddaja simbole v simetrični dvojiški komunikacijski kanal. Verjetnost, da pride pri prenosu sporočila do napake, na kakšnem od oddanih simbolov, je enaka $p_n = 0{,}1$, pri čemer napako predstavlja sprememba simbola iz $0 \to 1$ ali $1 \to 0$. Vsak simbol, ki ga oddamo v kanal, ponovimo (oz. prenesemo) trikrat,

$$\begin{aligned} 0 \to x_1 = 000 \ , \\ 1 \to x_2 = 111 \ . \end{aligned}$$

Zaradi morebitnih napak med prenosom, dobimo na sprejemni strani poljubno zaporedje treh dvojiških simbolov $y_1, ..., y_8$, ki jih sprejemnik

dekodira po naslednjih pravilih:

$$\begin{aligned} y_1 &= 000 \to 0 \ , \\ y_2 &= 001 \to 0 \ , \\ y_3 &= 010 \to 0 \ , \\ y_4 &= 011 \to 1 \ , \\ y_5 &= 100 \to 0 \ , \\ y_6 &= 101 \to 1 \ , \\ y_7 &= 110 \to 1 \ , \\ y_8 &= 111 \to 1 \ . \end{aligned}$$

Kolikšna je verjetnost, da bo sprejemnik napačno dekodiral prejeto zaporedje treh simbolov?

Rešitev:

Iz naloge lahko razberemo, da bo sprejemnik oddani simbol pravilno dekodiral, če med prenosom pride le do ene napake na oddanih simbolih. V primeru dveh oz. treh napak je dekodiranje napačno. Verjetnost napačnega dekodiranja je torej enaka

$$p_{napake} = p_2 + p_3 \ ,$$

kjer sta p_2 in p_3 verjetnosti nastopa dveh oz. treh napak na simbolih pri prenosu skozi komunikacijski kanal. Verjetnosti lahko izračunamo po binomskem porazdelitvenem zakonu:

$$p_2 = \binom{3}{2} p_n^2 (1-p_n)^1 = \binom{3}{2} 0{,}1^2 \cdot 0{,}9^1 = 3 \cdot 0{,}009 = 0{,}027 \ ,$$

$$p_3 = \binom{3}{3} p_n^3 (1-p_n)^0 = \binom{3}{2} 0{,}1^3 \cdot 0{,}9^0 = 3 \cdot 1 = 0{,}001 \ ,$$

kar nam da verjetnost napake

$$p_{napake} = p_2 + p_3 = 0{,}027 + 0{,}001 = 0{,}028 \ .$$

***Naloga* 2.6**

Informacijski vir oddaja sporočila v komunikacijski kanal. Kodiranje sporočil omogoča, da sprejemnik ugotovi, ali je prišlo med sprejemom sporočila do napake. Naj bo verjetnost, da se pri prenosu sporočila zgodi napaka, enaka $p_n = 0{,}1$. Najmanj koliko-krat je potrebno ponoviti prenos sporočila, da bo verjetnost uspešnega prenosa sporočila večja od 0,999?

Rešitev:

Iz naloge lahko razberemo, da imamo pri prenosu sporočila preko komunikacijskega kanala opravka z dvema možnima izidoma. Lahko se zgodi napaka in sprejemnik prejme napačno sporočilo ali se napaka ne zgodi in sprejemnik prejme pravilno sporočilo. Označimo prvi dogodek z A in drugi z $\overline{A}$ ter pripadajoči verjetnosti s $p = 1-p_n$ ter $(1-p) = p_n$. Ugotovimo lahko, da mora biti verjetnost pravilnega prenosa sporočila enaka

$$0{,}999 \leq p_1 + p_2 + p_3 + \cdots + p_i = p_{skupna} \ ,$$

kjer so $p_1, p_2, p_3, ..., p_i$ verjetnosti, da je prišlo do pravilnega prenosa sporočila v prvi, drugi, tretji, ... i-ti oddaji sporočila, in nas zanima, pri katerem i bo verjetnost uspešnega prenosa p_{skupna} večja ali enaka 0,999.

Vidimo, da predstavljajo verjetnosti $p_1, p_2, p_3, ..., p_i$ verjetnosti naključne spremenljivke, ki se podreja geometrijski porazdelitvi in zato izračunamo:

- verjetnost uspešnega prenosa v $k = 1$ ponovitvi: $p_1 = p(1-p)^0 = p = 0{,}9$, kar nam da $p_{skupna} = p_1 = 0{,}9$,
- verjetnost uspešnega prenosa v $k = 2$ ponovitvi: $p_2 = p(1-p)^1 = 0{,}9 \cdot 0{,}1 = 0{,}09$, kar nam da $p_{skupna} = p_1 + p_2 = 0{,}99$,
- verjetnost uspešnega prenosa v $k = 3$ ponovitvi: $p_3 = p(1-p)^2 = 0{,}9 \cdot 0{,}01 = 0{,}09$, kar nam da $p_{skupna} = p_1 + p_2 + p_3 = 0{,}999$.

Prenos sporočila je torej potrebno ponoviti 3 krat.

3 Entropija

3.1 Entropija diskretnih naključnih spremenljivk

Če vzamemo, da ima dinamičen, diskretni, naključni sistem n stanj, ki jih označimo z znaki x_i, verjetnosti, da se sistem nahaja v teh stanjih, pa s $p(x_i)$, za $i = 1,2,...,n$, lahko takšen sistem opišemo z naključno diskretno spremenljivko. Opis dinamičnega, diskretnega, naključnega sistema z naključno diskretno spremenljivko velja le za določen časovni trenutek. Zanimiv je predvsem zato, ker omogoča obravnavo entropije (oz. nedoločenost) sistema ločeno od njenega fizikalnega ozadja.

Entropija (sistema) je funkcija naključne diskretne spremenljivke, ki opisuje sistem. Entropija naključne diskretne spremenljivke X, ki črpa vrednosti iz končne zaloge vrednosti $\mathcal{Z}(X) = \{x_1, x_2, ..., x_n\}$ v skladu s porazdelitvijo verjetnosti $P_X = \big(p(x_1), p(x_2), ..., p(x_n)\big)$, kjer je $p(x_i) \geq 0$ in $\sum_{i=1}^{n} p(x_i) = 1$, je

$$H(X) = -K \sum_{i=1}^{n} p(x_i) \log_d p(x_i) \ ,$$

kjer sta $K > 0$ poljubna konstanta in $d > 1$ osnova logaritma [4].

Kadar za osnovo logaritma d izberemo vrednost $d = 2$, predstavlja enoto entropije *bit*[1], če za osnovo logaritma izberemo $d = e$ predstavlja enoto entropije *nat*. Entropijo torej podajamo kot *bit* ali *nat* na stanje naključne spremenljivke, pri čemer stanje lahko predstavlja oddani znak, oddano sporočilo, niz znakov ipd.

Entropija $H(X)$ naključne spremenljivke X je mera nedoločenosti dogodka, da zavzame naključna spremenljivka določeno vrednost iz svoje zaloge vrednosti. Vidimo tudi, da lahko pišemo [4]

$$H(X) = H\big(p(x_1), ..., p(x_n)\big) \ .$$

[1] Na mesto enote *bit* je pri izbrani osnovi logaritma $d = 2$ pogosto v uporabi tudi enota *Shannon*.

Za $H(X)$ velja [4]:

- $H(X) = H\Big(p(x_1), ..., p(x_n)\Big) \geq 0$ in je enaka 0 le v primeru, ko je nek $p(x_i) = 1$, za $i \in \{1, 2, ..., n\}$.
- $H(X) = H\Big(p(x_1), ..., p(x_n)\Big)$ je maksimalna kadar so vsa stanja naključne spremenljivke enako verjetna: $p(x_1) = \cdots = p(x_n) = 1/n$.
- $H(1/n, ..., 1/n) > H(1/m, ..., 1/m)$, ko je $n > m$.
- $H(X) = H\Big(p(x_1), ..., p(x_n)\Big)$ je neodvisna od permutacij $p(x_1), ..., p(x_n)$.
- Stanja z verjetnostjo 0 ne spremenijo vrednosti entropije:

$$H\Big(p(x_1), ..., p(x_n), 0\Big) = H\Big(p(x_1), ..., p(x_n)\Big) \ .$$

- $H\Big(p(x_1), ..., p(x_n)\Big)$ je zvezna funkcija.
- Skupna nedoločenost dveh neodvisnih naključnih spremenljivk (oz. sistemov) z m in n stanji je enaka vsoti nedoločenosti posameznih naključnih spremenljivk ($m, n \in \mathbb{N}$):

$$H\left(\frac{1}{mn}, ..., \frac{1}{mn}\right) = H\left(\frac{1}{m}, ..., \frac{1}{m}\right) + H\left(\frac{1}{n}, ..., \frac{1}{n}\right) \ .$$

- Če lahko stanja naključne spremenljivke razdelimo v dva razreda z m in n stanji, lahko entropijo računamo po naslednjem obrazcu:

$$H(p_1, ..., p_m, r_1, ..., r_n) = H(p\,,\, r) + pH\left(\frac{p_1}{p}, ..., \frac{p_m}{p}\right) + rH\left(\frac{r_1}{r}, ..., \frac{r_n}{r}\right) \ ,$$

  kjer $m, n \in \mathbb{N}$, $p = p_1 + \cdots + p_m$, $r = r_1 + \cdots + r_n$, $p + r = 1$, $p_i \geq 0$, $r_j \geq 0$ in velja $p_1 = p(x_1), p_m = p(x_m), r_1 = p(x_{m+1})$ ter $r_n = p(x_{m+n})$.

***Naloga* 3.1**

Izid meta kovanca lahko opišemo z naključno spremenljivko X z zalogo vrednosti $\mathcal{Z}(X) = \{x_1, x_2\} = \{0, 1\} \sim \{\texttt{'grb'}, \texttt{'cifra'}\}$ in porazdelitvijo verjetnosti $P_X = (p(x_1)\,,\, p(x_2)) = (0{,}5\,,\, 0{,}5)$ Določite entropijo naključne spremenljivke X.

Rešitev:

Entropijo naključne spremenljivke (na stanje) izračunamo kot:

$$H(X) = H\big(p(x_1)\,,\,p(x_2)\big) = H\left(\frac{1}{2}\,,\,\frac{1}{2}\right) = \log_2 2 = 1\ \ [bit]\,.$$

Za izračun entropije smo izbrali osnovo logaritma $d = 2$, rezultat pa je zato podan v bitih na stanje. Dobljeni rezultat lahko tolmačimo kot minimalno število vprašanj (z dvema možnima odgovoroma, npr. *d*a in *n*e), ki jih je potrebno zastaviti, da enolično določimo stanje naključne spremenljivke. V našem primeru bi zadostovalo že eno samo vprašanje ($H(X) = 1$ *bit*), npr. *Ali je padla cifra?* Odgovor na vprašanje nam popolnoma določi stanje naše naključne spremenljivke oz. izid meta kovanca s podano verjetnostno porazdelitvijo.

***Naloga* 3.2**

Denimo, da izid meta kovanca zopet predstavimo z naključno spremenljivko X z zalogo vrednosti $\mathcal{Z}(X) = \{x_1, x_2\} = \{0, 1\} \sim \{\texttt{'grb'}, \texttt{'cifra'}\}$, le da imamo tokrat opravka z nepoštenim kovancem. Verjetnostna porazdelitev je sedaj podana z $P_X = \big(p(x_1)\,,\,p(x_2)\big) = (0{,}9\,,\,0{,}1)$. Izračunajmo entropijo podane naključne spremenljivke!

Rešitev:

Podobno kot v prejšnjem primeru izračunamo entropijo kot

$$\begin{aligned} H(X) &= H\big(p(x_1)\,,\,p(x_2)\big) = H\big(0{,}9\,,\,0{,}1\big) = \\ &= -0{,}9\log_2 0{,}9 - 0{,}1\log_2 0{,}1 \approx 0{,}469\ \ [bitov]\,. \end{aligned}$$

Rezultat izračuna tokrat ni celo število, a ga lahko še zmeraj tolmačimo kot minimalno število vprašanj, ki jih je potrebno zastaviti, da enolično določimo stanje naključne spremenljivke. Drugače kot pri prejšnji nalogi moramo tokrat razmišljati o množici izidov meta kovanca in ne zgolj o izidu enega meta. Rezultat nam pove, da bi s 469 vprašanji lahko določili stanje 1000 izidov meta kovanca, če bi jih postavljali na optimalen način.

***Naloga* 3.3**

Priredimo izbiranju kart iz svežnja 32 kart naključno spremenljivko Y z zalogo vrednosti $\mathcal{Z}(Y) = \{y_1, y_2, ..., y_{32}\}$. Predpostavimo še, da je verjetnost izbire za vse karte enaka $\frac{1}{32}$. Izračunajmo entropijo tako podane naključne spremenljivke.

Rešitev:

Entropijo naključne spremenljivke Y določimo kot:

$$\begin{aligned} H(Y) &= H\big(p(y_1)\,,\, p(y_2)\,,\, ...\,,\, p(y_{32})\big) = \\ &= H\left(\frac{1}{32}\,,\, ...\,,\, \frac{1}{32}\right) = \log_2 32 = 5 \;\; [bitov]\,. \end{aligned}$$

Ker so verjetnosti vseh stanj naključne spremenljivke enake, predstavlja izračunana entropija maksimalno možno entropijo naključne spremenljivke z 32 možnimi stanji.

***Naloga* 3.4**

Izračunajmo entropijo dvojiškega informacijskega vira V, ki oddaja dva simbola v_1 in v_2. Pri tem oddaja simbol v_1 z verjetnostjo $p(v_1) = 0{,}25$ in simbol v_2 z verjetnostjo $p(v_2) = 0{,}75$.

Rešitev:

Naključno spremenljivko V (tj., dvojiški informacijski vir) lahko opišemo z naslednjo verjetnostno shemo

$$V \sim \begin{pmatrix} v_1\,, & v_2 \\ 0{,}25\,, & 0{,}75 \end{pmatrix}.$$

Entropijo vira nato izračunamo kot:

$$H(V) = H\big(p(v_1)\,,\, p(v_2)\big) = H\left(\frac{1}{4}\,,\, \frac{3}{4}\right) \approx 0{,}811 \;\; [bitov]\,.$$

***Naloga* 3.5**

Informacijski vir V ima naslednjo verjetnostno shemo

$$V \sim \begin{pmatrix} v_1\,, & v_2\,, & v_3\,, & v_4\,, & v_5 \\ 0{,}30\,, & 0{,}25\,, & 0{,}20\,, & 0{,}15\,, & 0{,}1 \end{pmatrix}.$$

Denimo, da simbole vira V zakodiramo dvojiško (tj., z dvojiškima simboloma) po sledečem predpisu

$$v_1 \to w_1 = 10\ ,$$

$$v_2 \to w_2 = 11\ ,$$

$$v_3 \to w_3 = 00\ ,$$

$$v_4 \to w_4 = 010\ ,$$

$$v_5 \to w_5 = 011\ .$$

Tako kodirane simbole pošljemo po dvojiškemu kanalu brez motenj do sprejemnika. Izračunajte entropijo vira potem, ko sprejemnik sprejme prvi kodni znak, ki je ali 0 ali 1. Kolikšna je nedoločenost vira po sprejemu drugega znaka?

Rešitev:

Preden sprejemnik sprejme prvi kodni znak, je zanj entropija vira V enaka

$$\begin{aligned} H(V) &= H\big(0{,}3\,,\ 0{,}25\,,\ 0{,}2\,,\ 0{,}15\,,\ 0{,}1\big) = \\ &= -0{,}3\log_2 0{,}3 - 0{,}25\log_2 0{,}25 - \cdots - 0{,}1\log_2 0{,}1 \approx \\ &\approx 2{,}24\ \ [bitov]\ . \end{aligned}$$

Ko sprejemnik sprejme prvi kodni znak, se razmere zanj spremenijo. V primeru, da je prvi prejeti znak enak 1, lahko sklepamo, da je vir oddal ali simbol v_1 ali simbol v_2. Še zmeraj pa ostaja preostala negotovost, ki je posledica dejstva, da ne vemo natančno, kateri od obeh simbolov je bil dejansko oddan. Podobno lahko v primeru, ko je prvi prejeti znak enak 0, sklepamo, da je vir oddal enega od simbolov v_3, v_4 oz. v_5. Tudi tokrat obstaja preostala negotovost, ki je zopet posledica dejstva, da ne vemo natančno, kateri od teh treh simbolov je bil dejansko oddan. Opazimo lahko, da se zaradi sprejema prvega znaka entropija vira

zmanjša, pri čemer je preostala entropija H_{pr} posledica negotovosti, za kateri simbol znotraj obeh skupin gre.

Na podlagi prvega oddanega kodnega znaka je torej mogoče simbole informacijskega vira V razdeliti v dve skupini (oz. razreda) z $m = 2$ in $n = 3$ stanji. Če označimo verjetnosti $p(v_1) = p_1$, $p(v_2) = p_2$, $p(v_3) = r_1$, $p(v_4) = r_2$ in $p(v_5) = r_3$, lahko s pomočjo lastnosti *utežene vsote* (glej lastnosti na strani 35) entropijo vira zapišemo kot

$$H(V) = H(p\,,\,r) + pH\left(\frac{p_1}{p}\,,\,\frac{p_2}{p}\right) + rH\left(\frac{r_1}{r}\,,\,\frac{r_2}{r}\,,\,\frac{r_3}{r}\right) = H(p\,,\,r) + H_{pr}\ ,$$

kjer sta $p = p_1 + p_2 = 0{,}55$ in $r = r_1 + r_2 + r_3 = 0{,}45$.

Prvi člen v zgornji enačbi predstavlja negotovost zaradi izbire med obema skupinama simbolov, naslednja dva člena pa predstavljata preostalo negotovost zaradi izbire med simboli znotraj obeh skupin.

Vidimo, da lahko preostalo negotovost izračunamo kot:

$$\begin{aligned} H_{pr} &= H(V) - H(p\,,\,r) = \\ &= 2{,}24 - H(0{,}55\,,\,0{,}45) = 2{,}24 - 0{,}99 = 1{,}25\ \ [bitov]\ . \end{aligned}$$

Do enekega rezultata lahko pridemo tudi po drugi poti, in sicer

$$\begin{aligned} H_{pr} &= pH\left(\frac{p_1}{p}\,,\,\frac{p_2}{p}\right) + rH\left(\frac{r_1}{r}\,,\,\frac{r_2}{r}\,,\,\frac{r_3}{r}\right) = \\ &= 0{,}55 \cdot H\left(\frac{0{,}3}{0{,}55}\,,\,\frac{0{,}25}{0{,}55}\right) + 0{,}45 \cdot H\left(\frac{0{,}2}{0{,}45}\,,\,\frac{0{,}15}{0{,}45}\,,\,\frac{0{,}1}{0{,}45}\right) \approx \\ &\approx 1{,}25\ \ [bitov]\ . \end{aligned}$$

Po sprejemu še drugega kodnega znaka, so simboli v_1, v_2 in v_3 določeni, preostala negotovost pa je zgolj posledica negotovosti izbire med $v4$ in v_5. Oddane simbole lahko tokrat razdelimo v štiri skupine (oz. razrede). Če označimo verjetnosti vseh štirih razredov z

$$\begin{aligned} r_1 &= p(v_1) = 0{,}3\ , \\ r_2 &= p(v_2) = 0{,}25\ , \\ r_3 &= p(v_3) = 0{,}2\ , \\ r_4 &= p(v_4) + p(v_5) = 0{,}25\ , \end{aligned}$$

potem lahko tudi v tem primeru entropijo vira izračunamo kot

$$H(V) =$$

$$= H(r_1, ..., r_4) + r_1 H\Big(\frac{r_1}{r_1}\Big) + r_2 H\Big(\frac{r_2}{r_2}\Big) + r_3 H\Big(\frac{r_3}{r_3}\Big) + r_4 H\Big(\frac{p(v_4)}{r_4}\,,\, \frac{p(v_5)}{r_4}\Big) =$$

$$= H(r_1\,,\, r_2\,,\, r_3\,,\, r_4) + H_{pr}\ .$$

Preostala entropija po sprejemu še drugega kodnega znaka je torej

$$\begin{aligned} H_{pr} &= H(V) - H(0{,}3\,,\, 0{,}25\,,\, 0{,}2\,,\, 0{,}25) = \\ &= 2{,}24 - 1{,}99 \approx \\ &\approx 0{,}24\ \ [bitov]\ . \end{aligned}$$

Še nazornejši je izračun po drugi poti

$$\begin{aligned} H_{pr} &= r_1 H\Big(\frac{r_1}{r_1}\Big) + r_2 H\Big(\frac{r_2}{r_2}\Big) + r_3 H\Big(\frac{r_3}{r_3}\Big) + r_4 H\Big(\frac{p(v_4)}{r_4}\,,\, \frac{p(v_5)}{r_4}\Big) = \\ &= 0{,}3 \cdot H(1) + 0{,}25 \cdot H(1) + 0{,}2 \cdot H(1) + 0{,}25 \cdot H\Big(\frac{0{,}15}{0{,}25}\,,\, \frac{0{,}1}{0{,}25}\Big) = \\ &= 0 + 0 + 0 + 0{,}25 \cdot \big(-0{,}6 \log_2 0{,}6 - 0{,}4 \log_2 0{,}4\big) \approx \\ &\approx 0{,}24\ \ [bitov]\ , \end{aligned}$$

kjer se opazi, da je entropija povezana z negotovostjo izbire med v_1, v_2 in v_3 po prejetju drugega kodnega znaka enaka 0.

3.1.1 Entropija para naključnih diskretnih spremenljivk

Par naključnih diskretnih spremenljivk (X, Y) črpa vrednosti iz zaloge vrednosti

$$\mathcal{Z}(X,Y) = \{(x_i, y_j) :\ \ i = 1{,}2, ..., m;\ \ j = 1{,}2, ..., n\}$$

v skladu z dvorazsežno porazdelitvijo verjetnosti

$$P_{(X,Y)} = \big(p(x_i, y_j) = P(X{=}x_i, Y{=}y_j) \geq 0 \colon\ i = 1, ..., m;\ j = 1, ..., n\big),$$

kjer velja $\sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i, y_j) = 1$.

Množico vseh možnih dvorazsežnih porazdelitev verjetnosti nad množico $\mathcal{Z}(X,Y)$ označimo z Δ_{mn}. $P_{(X,Y)}$ naj bo poljubna dvorazsežna porazdelitev verjetnosti iz Δ_{mn} [4] :

$$P_{(X,Y)} = \big(p(x_i,y_j) \geq 0 : \;\; i = 1,2,...,m; \;\; j = 1,2,...,n; \;\; \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i,y_j) = 1\big).$$

DEFINICIJA 2.1 *Entropija para naključnih diskretnih spremenljivk je*

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

kjer sta $K > 0$ *poljubna konstanta in* $d > 1$ *osnova logaritma* [4].

Entropija para naključnih spremenljivk izraža nedoločenost pojavljanja urejenih parov (x_i, y_j).

V zvezi s poljubno dvorazsežno porazdelitvijo $P_{(X,Y)}$ obravnavamo njuni robni porazdelitvi

$$P_X = \big(p(x_i) = P(X{=}x_i) \geq 0 : \;\; p(x_i) = \sum_{j=1}^{n} p(x_i,y_j); \;\; i = 1,2,...,m\big)$$

ter

$$P_Y = \big(p(y_j) = P(Y{=}y_j) \geq 0 : \;\; p(y_j) = \sum_{i=1}^{m} p(x_i,y_j); \;\; j = 1,2,...,n\big)$$

in pogojni porazdelitvi

$$P_{X|Y} = \big(p(x_i \mid y_j) = P(X{=}x_i \mid Y{=}y_j) \geq 0 \;\; i = 1,2,...,m\big)$$

za $j = 1,2,...,n$ ter

$$P_{Y|X} = \big(p(y_j \mid x_i) = P(Y{=}y_j \mid X{=}x_i) \geq 0 : \;\; j = 1,2,...,n\big)$$

za $i = 1,2,...,m$, kjer velja

$$p(x_i \mid y_j) = \frac{p(x_i,y_j)}{p(y_j)} \;\; \text{in} \;\; p(y_j \mid x_i) = \frac{p(x_i,y_j)}{p(x_i)} \; .$$

Iz robnih porazdelitev verjetnosti dobimo entropijo spremenljivke X, to je

$$H(P_X) = -K \sum_i p(x_i) \log_d p(x_i) = H(X) \tag{3.1}$$

in entropijo spremenljivke Y, to je

$$H(P_Y) = -K \sum_j p(y_j) \log_d p(y_j) = H(Y) \ , \tag{3.2}$$

kjer sta $K > 0$ poljubna konstanta in $d > 1$ osnova logaritma [4].

S pomočjo pogojnih porazdelitev verjetnosti lahko definiramo še naslednje entropije para naključnih spremenljivk:

DEFINICIJA 2.2 *Entropija spremenljivke X pri vrednosti spremenljivke $Y = y_j$ je*

$$H(P_{X|Y=y_j}) = -K \sum_i p(x_i \mid y_j) \log_d p(x_i \mid y_j) = H(X \mid Y = y_j) \ , \tag{3.3}$$

kjer je $j = 1, 2, ..., n$ ter sta $K > 0$ poljubna konstanta in $d > 1$ osnova logaritma [4].

DEFINICIJA 2.3 *Entropija spremenljivke Y pri vrednosti spremenljivke $X = x_i$ je*

$$H(P_{Y|X=x_i}) = -K \sum_j p(y_j \mid x_i) \log_d p(y_j \mid x_i) = H(Y \mid X = x_i) \ , \tag{3.4}$$

kjer je $i = 1, 2, ..., m$ ter sta $K > 0$ poljubna konstanta in $d > 1$ osnova logaritma [4].

DEFINICIJA 2.4 *Entropija spremenljivke X, ko poznamo spremenljivko Y, je*

$$H(X \mid Y) = \sum_j p(y_j) H(X \mid Y = y_j) = -K \sum_i \sum_j p(x_i, y_j) \log_d p(x_i \mid y_j) \ , \tag{3.5}$$

kjer sta $K > 0$ poljubna konstanta in $d > 1$ osnova logaritma [4].

DEFINICIJA 2.5 *Entropija spremenljivke Y, ko poznamo spremenljivko X, je*

$$H(Y \mid X) = \sum_i p(x_i) H(Y \mid X = x_i) = -K \sum_i \sum_j p(x_i, y_j) \log_d p(y_j \mid x_i) \ , \tag{3.6}$$

kjer sta $K > 0$ poljubna konstanta in $d > 1$ osnova logaritma [4].

$H(X \mid Y)$ imenujemo tudi *pogojna entropija* X-a glede na Y, $H(Y \mid X)$ *pogojna entropija* Y-a glede na X, $H(X, Y)$ oziroma $H(Y, X)$ pa *vezana entropija* X-a in Y-a oziroma Y-a in X-a.

Medsebojne zveze med količinami $H(X)$, $H(Y)$, $H(X, Y)$, $H(X \mid Y)$ in $H(Y \mid X)$ vzpostavimo z izrekom.

IZREK 2.1 *Veljajo naslednje zveze:*

$$H(X \mid Y) \leq H(X) \ , \tag{3.7}$$

$$H(Y \mid X) \leq H(Y) \ , \tag{3.8}$$

$$H(X, Y) = H(X) + H(Y \mid X) = H(Y) + H(X \mid Y) \ , \tag{3.9}$$

$$H(X, Y) \leq H(X) + H(Y) \ , \tag{3.10}$$

kjer velja enakost v (3.7), (3.8) *in* (3.10) *tedaj in le tedaj, če sta X in Y neodvisni naključni spremenljivki* [4].

***Naloga* 3.6**

Denimo, da smo za naključni spremenljivki X in Y izračunali njuni entropiji in pri tem dobili rezultat $H(X) = 0{,}5$ *bitov* ter $H(Y) = 0{,}9$ *bitov*. Določite vezano entropijo naključnih spremenljivk $H(X, Y)$ ob predpostavki, da sta spremenljivki neodvisni!

Rešitev:

Za neodvisne naključne spremenljivke velja

$$H(Y \mid X) = H(Y) \text{ oz. } H(X \mid Y) = H(X) \ .$$

Izraz za izračun vezane entropije se zato poenostavi v $H(X,Y) = H(X) + H(Y)$, kar da rezultat

$$H(X,Y) = H(X) + H(Y) = 0{,}5 + 0{,}9 = 1{,}4 \ \ [bitov] \ .$$

***Naloga* 3.7**

Predpostavite, da je vezana entropija dveh naključnih spremenljivk X in Y enaka $H(X,Y) = 2$ *bita* in je entropija $H(X) = 0{,}7$ *bitov*. Izračunajte pogojno entropijo $H(Y \mid X)$! Kaj lahko poveste o naključnih spremenljivkah X in Y, če veste, da je lastna entropija naključne spremenljivke Y enaka $H(Y) = 1{,}3$ *bitov*?

Rešitev:

V prvem koraku izračunamo pogojno entropijo $H(Y \mid X)$, kar storimo s pomočjo enačbe (3.9):

$$H(Y \mid X) = H(X,Y) - H(X) = 2 - 0{,}7 = 1{,}3\,[bitov] \ .$$

Opazimo, da je pogojna entropija $H(Y \mid X)$ enaka $H(Y)$ in da velja:

$$H(X \mid Y) = H(X,Y) - H(Y) = 2 - 1{,}3 = 0{,}7\,[bitov] = H(X) \ .$$

Naključni spremenljivki X in Y sta torej neodvisni.

3.1.2 Entropija n naključnih diskretnih spremenljivk

Entropijo poljubne končne množice naključnih diskretnih spremenljivk definiramo kot:

DEFINICIJA 2.6 *Entropija n–terice ($n \in \mathbb{N}$) naključnih diskretnih spremenljivk je*

$$H(X_1, X_2, ..., X_n) = -K \sum_{i_1} \cdots \sum_{i_n} p(x_{i_1}, ..., x_{i_n}) \log_d \, p(x_{i_1}, ..., x_{i_n}) \ ,$$

kjer so $K > 0$ poljubna konstanta, $d > 1$ osnova logaritma in $p(x_{i_1}, ..., x_{i_n}) = P(X_1 = x_{i_1}, X_2 = x_{i_2}, ..., X_n = x_{i_n})$ n-razsežna porazdelitev verjetnosti [4].

Med entropijami posameznih spremenljivk, pogojnimi entropijami in vezano entropijo lahko vzpostavimo naslednji zvezi:

IZREK 2.2 *Veljata zvezi:*

$$\begin{aligned} H(X_1,...,X_n) &= H(X_1)+H(X_2 \mid X_1)+H(X_3 \mid X_2,X_1)+\cdots+ \\ &\quad +H(X_n \mid X_{n-1},...,X_1) \\ &= \sum_{i=1}^{n} H(X_i \mid X_{i-1},...,X_1) \end{aligned} \tag{3.11}$$

in

$$H(X_1,...,X_n) \leq H(X_1)+\cdots+H(X_n)=\sum_{i=1}^{n} H(X_i)\ , \tag{3.12}$$

kjer velja enakost tedaj in le tedaj, če so X_i neodvisne naključne spremenljivke [4].

3.2 Entropija zveznih naključnih spremenljivk

Naj bo X naključna spremenljivka s porazdelitveno funkcijo $F(x) = P(X \leq x)$. Naključni spremenljivki pravimo, da je *zvezna*, če je njena porazdelitvena funkcija $F(x)$ zvezna nad množico vseh realnih števil $\mathbb{R}$. Naj bo $f_X(x) = F'(x)$ na območju, kjer odvod obstaja. Če je $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$, pravimo funkciji $f_X(x)$ *funkcija verjetnostne gostote* spremenljivke X. Množico vseh $x \in \mathbb{R}$, za katere je $f_X(x) > 0$, imenujemo *nosilec* naključne spremenljivke X [2], [4].

Entropijo naključne zvezne spremenljivke vpeljemo z definicijo:

DEFINICIJA 2.7 *Entropija naključne zvezne spremenljivke*[2] *X, ki je dana s funkcijo gostote verjetnosti $f_X(x)$, $x \in \mathbb{R}$, je*

$$h(X) = -K \int_{\mathcal{P}} f_X(x) \log_d f_X(x)\,dx, \tag{3.13}$$

pod pogojem, da integral (3.13) *konvergira. Pri tem so: $\mathcal{P}$ nosilec naključne spremenljivke, $K > 0$ poljubna konstanta in $d > 1$ osnova logaritma* [4].

---

[2] Anglosaški izraz *Differential Entropy*.

4 Informacija

4.1 Lastna informacija

Informacija lahko nastane le med sporazumevanjem (komunikacijo), le-to pa poteka v sistemu, ki ga sestavljajo vir informacije, kanal, ki prevaja informacijo, in sprejemnik informacije. V uvodnem poglavju smo že povedali, da takšnim sistemom pravimo komunikacijski sistemi ali bolj splošno *informacijski sistemi*.

Naj bo X diskretna naključna spremenljivka z zalogo vrednosti $\mathcal{Z}(X) = \{x_1, ..., x_i, ..., x_n\}$ in porazdelitvijo verjetnosti $p(x_i) = P(X{=}x_i) \geq 0$, kjer velja $\sum_{i=1}^{n} p(x_i) = 1$. Če zalogo vrednosti naključne spremenljivke X razumemo kot zalogo znakov, iz katere vir informacije črpa znake, rečemo, da znak x_i oddan z verjetnostjo $p(x_i)$ nosi informacijo $I(x_i)$ [4].

V znaku je zajeta informacija

$$I(x_i) = -K \log_d p(x_i), \text{ za } i = 1,2,...,n,$$

kjer sta $K > 0$ poljubna konstanta in $d > 1$ osnova logaritma. $I(x_i)$ predstavlja *lastno informacijo* znaka x_i [4].

Povprečna vrednost lastne informacije naključne diskretne spremenljivke X je tako enaka

$$E\big(I(x_i)\big) = \sum_{i=1}^{n} p(x_i) I(x_i) = -K \sum_{i=1}^{n} p(x_i) \log_d p(x_i) = H(X) \ ,$$

kjer $E(\cdot)$ pomeni operator matematičnega upanja.

Opazimo, da je povprečna lastna informacija naključne spremenljivke povezana z entropijo naključne spremenljivke. Sprejemnik torej prejme toliko informacije, za kolikor se je zmanjšala njegova nedoločenost. Ker je informacija enaka entropiji, se tudi meri v enakih enotah, tj., v *bitih*, kadar je $K = 1$ in $d = 2$ oz. v *natih*, ko je $K = 1$ in $d = e$.

***Naloga* 4.1**

Predpostavite dva informacijska vira, ki oddajata sporočila, ki jih predstavimo z diskretnima naključnima spremenljivkama X in Y. Prvi vir odda enega od 5 sporočil $x_1, x_2, x_3, x_4,$ in x_5 z verjetnostmi $(p(x_1), p(x_2), ..., p(x_3)) = (\frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{3}{8}, \frac{1}{12})$ in drugi odda sporočila $y_1, y_2, y_3, y_4,$ in y_5 z enakimi verjetnostmi $(p(y_1), p(y_2), ..., p(y_3)) = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$. Kateri vir je bolj informativen (ima večjo povprečno lastno informacijo)?

Rešitev:

Informativnost vira določimo kot povprečno lastno informacijo vira. V prvem koraku moramo izračunati lastno informacijo znakov (oz. sporočil) za oba vira. Če izberemo $K = 1$ in $d = 2$, za prvi vir določimo

$$\begin{aligned}
I(x_1) &= -\log_2 p(x_1) = \log_2 \frac{1}{4} = 2 \ \ [bitov] \ , \\
I(x_2) &= -\log_2 p(x_2) = \log_2 \frac{1}{6} \approx 2{,}59 \ \ [bitov] \ , \\
I(x_3) &= -\log_2 p(x_3) = \log_2 \frac{1}{8} = 3 \ \ [bitov] \ , \\
I(x_4) &= -\log_2 p(x_4) = \log_2 \frac{3}{8} \approx 1{,}4 \ \ [bitov] \ , \\
I(x_5) &= -\log_2 p(x_5) = \log_2 \frac{1}{12} \approx 3{,}59 \ \ [bitov] \ ,
\end{aligned}$$

kar nam da povprečno lastno informacijo prvega vira

$$\begin{aligned}
E\big(I(x_i)\big) &= \sum_{i=1}^{5} p(x_i) I(x_i) \approx \\
&\approx \frac{1}{4} \cdot 2 + \frac{1}{6} \cdot 2{,}59 + \frac{1}{8} \cdot 3 + \frac{3}{8} \cdot 1{,}4 + \frac{1}{12} \cdot 3{,}59 = 2{,}13 \ \ [bitov] \ .
\end{aligned}$$

Na podoben način izračunamo še lastno informacijo drugega vira

$$\begin{aligned}
E(I(y_i)) &= \sum_{i=1}^{5} p(y_i) I(y_i) = -\sum_{i=1}^{5} p(y_i) \log_d p(y_i) = \\
&= -\frac{1}{5} \log_2 \frac{1}{5} - \cdots - \frac{1}{5} \log_2 \frac{1}{5} = \log_2 5 \approx 2{,}32 \ \ [bitov] \ .
\end{aligned}$$

Bolj informativen je torej drugi vir. Takšen rezultat je tudi pričakovan, saj smo že pri obravnavi entropije omenili, da je le-ta maksimalna, kadar so vsa stanja naključne spremenljivke enako verjetna. V kontekstu

obravnave naključnih spremenljivk z vidika informacije opazimo, da prejmemo več informacije od naključnih spremenljivk (oz. informacijskih virov), katerih stanja imajo podobne verjetnosti nastopa, kot od naključnih spremenljivk, kjer ima določeno stanje izrazito verjetnost nastopa v primerjavi z ostalimi. V tem primeru namreč moremo z veliko gotovostjo napovedati naslednje stanje naključne spremenljivke (oz. informacijskega vira), količina informacije, ki jo pri tem prejmemo pa je majhna.

4.2 Vzajemna informacija

Pri obravnavi komunikacijskih sistemov nas pogosto poleg lastne informacije zanima *vzajemna informacija*, ki jo ena spremenljivka vsebuje o drugi. V komunikacijskih sistemih, se pojem vzajemne informacije navezuje na kanal, po katerem se prenaša informacija, in se uporablja za opis informacije, ki jo vsebuje znak na izhodu kanala o znaku na vhodu kanala.

Naj ima naključna diskretna spremenljivka X entropijo $H(X)$ in naj ima naključna diskretna spremenljivka Y entropijo $H(Y)$. Ko določimo stanje Y, se entropija spremenljivke zmanjša na $H(Y) = 0$. V primeru, da sta spremenljivki X in Y odvisni, se zmanjša tudi entropija naključne spremenljivke X, in sicer na $H(X \mid Y)$. Razlika predstavlja vzajemno informacijo [4]:

$$I(X,Y) = H(X) - H(X \mid Y) \ .$$

Vzajemna informacija je simetrična funkcija, tj., $I(X,Y) = I(Y,X)$, zato zapišemo še

$$I(X,Y) = H(Y) - H(Y \mid X)$$

in

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \ .$$

Pokažemo še, da je vzajemna informacija omejena [4]:

$$0 \leq I(X,Y) \leq \min\{H(X), H(Y)\} \ .$$

***Naloga* 4.2**

Imejmo dva kovanca, prirejenega in regularnega. Pri metu prirejenega kovanca vedno pade grb, medtem ko pri metu regularnega pade z enako

verjetnostjo bodisi cifra bodisi grb. Naključno izberemo enega od obeh kovancev in ga vržemo dvakrat zapored. Zanima nas koliko nam dvakratni met kovanca pove o tem, katerega od obeh kovancev smo izbrali?

Rešitev:

Predstavljeni problem opišemo z dvema naključnima spremenljivkama X in Y, pri čemer nas zanima vzajemna informacija $I(X,Y)$ med obema spremenljivkama. Naključna spremenljivka X opisuje postopek izbire kovanca $\{x_1, x_2\} \sim \{\texttt{'0 (regularni)'}, \texttt{'1 (prirejeni)'}\}$ in naključna spremenljivka Y rezultat dva–kratnega meta kovanca z možnimi izidi $\{y_1, y_2, y_3\} \sim \{\texttt{'0 (grbov)'}, \texttt{'1 (grb)'}, \texttt{'2 (grba)'}\}$. Spremenljivki opišemo z naslednjima verjetnostnima shemama:

$$X \sim \begin{pmatrix} x_1\,, & x_2 \\ p(x_1)\,, & p(x_2) \end{pmatrix} ,$$

$$Y \sim \begin{pmatrix} y_1\,, & y_2\,, & y_3 \\ p(y_1)\,, & p(y_2)\,, & p(y_3) \end{pmatrix} ,$$

kjer je $p(x_1) = p(x_2) = 0{,}5$ in je verjetnosti $p(y_1), p(y_2)$ in $p(y_3)$ še potrebno določiti. Ker poznamo lastne verjetnosti spremenljivke X, potrebujemo za izračun vzajemne informacije le še vse pogojne verjetnosti $p(x_i \mid y_j)$, za $i = 1,2$ in $j = 1,2,3$, ki jih določimo iz vezanega porazdelitvenega zakona $P_{(X,Y)}$.

Če označimo verjetnost, da pri metu regularnega kovanca pade grb, s $p_n = 0{,}5$ in verjetnost, da pade cifra, z $(1 - p_n) = 0{,}5$, pogojno verjetnost, da ne bo padel nobeden grb, če smo izbrali regularni kovanec, zapišemo kot

$$p(y_1 \mid x_1) = \binom{2}{0} p_n^0 (1-p_n)^2 = \frac{2!}{2! \cdot 0!} \cdot 0{,}5^0 \cdot 0{,}5^2 = 0{,}25 \ .$$

Podobno storimo še za preostali dve pogojni verjetnosti:

$$p(y_2 \mid x_1) = \binom{2}{1} p_n^1 (1-p_n)^1 = \frac{2!}{1! \cdot 1!} \cdot 0{,}5^1 \cdot 0{,}5^1 = 2 \cdot 0{,}25 = 0{,}5 \ ,$$

$$p(y_3 \mid x_1) = \binom{2}{2} p_n^2 (1-p_n)^0 = \frac{2!}{0! \cdot 2!} \cdot 0{,}5^0 \cdot 0{,}5^2 = 0{,}25 \ .$$

V primeru, da smo pri izbiranju kovanca izbrali prirejen kovanec, je verjetnost, da bosta pri dva–kratnem metu padla dva grba enaka ena, preostali pogojni verjetnosti pa sta enaki nič:

$$p(y_1 \mid x_2) = 0 \ ,$$
$$p(y_2 \mid x_2) = 0 \ ,$$
$$p(y_3 \mid x_2) = 1 \ .$$

Porazdelitveni zakon $P_{(Y|X)}$ je torej enak

$$P_{(Y|X)} = \begin{pmatrix} p(y_1 \mid x_1)\,, & p(y_2 \mid x_1)\,, & p(y_3 \mid x_1) \\ p(y_1 \mid x_2)\,, & p(y_2 \mid x_2)\,, & p(y_3 \mid x_2) \end{pmatrix} = \begin{pmatrix} \frac{1}{4}\,, & \frac{1}{2}\,, & \frac{1}{4} \\ 0\,, & 0\,, & 1 \end{pmatrix}.$$

Z upoštevanjem zveze $p(x_i, y_j) = p(x_i)\, p(y_j \mid x_i)$ izračunamo vse vezane verjetnosti in jih podamo v vezanem porazdelitvenem zakonu

$$P_{(X,Y)} = \begin{pmatrix} p(x_1, y_1)\,, & p(x_2, y_1) \\ p(x_1, y_2)\,, & p(x_2, y_2) \\ p(x_1, y_3)\,, & p(x_2, y_3) \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \cdot 0{,}5\,, & 0 \cdot 0{,}5 \\ \frac{1}{2} \cdot 0{,}5\,, & 0 \cdot 0{,}5 \\ \frac{1}{4} \cdot 0{,}5\,, & 1 \cdot 0{,}5 \end{pmatrix} = \begin{pmatrix} \frac{1}{8}\,, & 0 \\ \frac{1}{4}\,, & 0 \\ \frac{1}{8}\,, & \frac{1}{2} \end{pmatrix}.$$

Z upoštevanjem zvez $p(y_j) = \sum_{i=1}^{m} p(x_i, y_j)$ in $p(x_i \mid y) = \frac{p(x_i, y_j)}{p(y_j)}$ iz vezane porazdelitve verjetnosti določimo še drugo robno porazdelitev lastnih verjetnosti P_Y in drugo pogojno porazdelitev $P_{X|Y}$:

$$\begin{aligned} P_Y &= (p(y_1)\,, p(y_2)\,, p(y_3)) = \\ &= \left(\ \tfrac{1}{8} + 0\,, \ \tfrac{1}{4} + 0\,, \ \tfrac{1}{8} + \tfrac{1}{2} \ \right) = \left(\ \tfrac{1}{8}\,, \ \tfrac{1}{4}\,, \ \tfrac{5}{8} \ \right) \end{aligned}$$

in

$$P_{(X|Y)} = \begin{pmatrix} p(x_1 \mid y_1)\,, & p(x_2 \mid y_1) \\ p(x_1 \mid y_2)\,, & p(x_2 \mid y_2) \\ p(x_1 \mid y_3)\,, & p(x_2 \mid y_3) \end{pmatrix} = \begin{pmatrix} \frac{1}{8}/\frac{1}{8}\,, & 0/\frac{1}{8} \\ \frac{1}{4}/\frac{1}{4}\,, & 0/\frac{1}{4} \\ \frac{1}{8}/\frac{5}{8}\,, & \frac{1}{2}/\frac{5}{8} \end{pmatrix} = \begin{pmatrix} 1\,, & 0 \\ 1\,, & 0 \\ \frac{1}{5}\,, & \frac{4}{5} \end{pmatrix}.$$

Ker vzajemno informacijo izračunamo kot $I(X, Y) = H(X) - H(X \mid Y)$ in poznamo verjetnostno porazdelitev spremenljivke X, najprej določimo

$$H(X) = H(p(x_1)\,, p(x_2)) = H\left(\frac{1}{2}\,, \frac{1}{2}\right) = \log_2 2 = 1 \ \ [bit] \ ,$$

nato izračunamo pogojne entropije

$$H_1 = H\big(p(x_1 \mid y_1)\,,\, p(x_2 \mid y_1)\big) = H(1\,,\,0) = 0 \ \ [bitov] \ ,$$

$$H_2 = H\big(p(x_1 \mid y_2)\,,\, p(x_2 \mid y_2)\big) = H(1\,,\,0) = 0 \ \ [bitov] \ ,$$

$$H_3 = H\big(p(x_1 \mid y_3)\,,\, p(x_2 \mid y_3)\big) = H\left(\frac{1}{5}\,,\,\frac{4}{5}\right) \approx 0{,}72 \ \ [bitov] \ ,$$

ter

$$H(X \mid Y) = p(y_1)H_1 + p(y_2)H_2 + p(y_3)H_3 \approx \frac{5}{8} \cdot 0{,}72 \approx 0{,}45 \ \ [bitov] \ .$$

Vzajemna informacija je potem enaka

$$I(X,Y) = H(X) - H(X \mid Y) \approx 1 - 0{,}45 = 0{,}55 \ \ [bitov] \ .$$

Do enakega rezultata bi prišli tudi z uporabo enačbe $I(X,Y) = H(Y) - H(Y \mid X)$.

***Naloga* 4.3**

Množico ljudi razdelimo na dve podmnožici $\mathcal{A}$ in $\mathcal{B}$. V podmnožici $\mathcal{A}$ je $\frac{1}{2}$ ljudi s kostanjevimi lasmi, $\frac{3}{10}$ jih ima temne lase in $\frac{2}{10}$ ljudi ima lase svetle barve. V podmnožici $\mathcal{B}$ je $\frac{3}{10}$ ljudi s kostanjevimi lasmi, $\frac{1}{2}$ jih ima temne lase in $\frac{2}{10}$ je svetlolasih. Verjetnost, da srečamo osebo iz podmnožice $\mathcal{A}$ naj bo $p(x_1)$. Označimo z I informacijo o pripadnosti osebe eni od skupin, ki jo dobimo na podlagi barve las. Določite izraz za informacijo I kot funkcijo verjetnosti $p(x_1)$, tj., $I\big(p(x_1)\big)$, in ugotovite pri katerem $p(x_1)$ je informacija maksimalna ter koliko le-ta znaša?

Rešitev:

Iz besedila naloge lahko razberemo, da imamo opravka z dvema naključnima spremenljivkama, kjer prva opisuje pripadnost k podmnožicama $\mathcal{A}$ in $\mathcal{B}$ in druga barvo las. Če pripadnost k podmnožicama $\mathcal{A}$ in $\mathcal{B}$ opišemo z diskretno naključno spremenljivko X z zalogo vrednosti $\{x_1, x_2\} = \{0, 1\} \sim \{$`'Iz podmnožice A'`, `'Iz podmnožice B'`$\}$ in barvo las z diskretno naključno spremenljivko Y z zalogo vrednosti $\{y_1, y_2, y_3\} = \{1, 2, 3\} \sim \{$`'Kostanjevi lasje'`, `'Temni lasje'`, `'Svetli lasje'`$\}$, potem zapišemo verjetnostni shemi naključnih spremenljivk kot

$$X \sim \begin{pmatrix} x_1\,, & x_2 \\ p(x_1)\,, & p(x_2) \end{pmatrix} = \begin{pmatrix} x_1\,, & x_2 \\ p(x_1)\,, & 1 - p(x_1) \end{pmatrix}, \text{ in}$$

$$Y \sim \begin{pmatrix} y_1\,, & y_2\,, & y_3 \\ p(y_1)\,, & p(y_2)\,, & p(y_3) \end{pmatrix},$$

kjer smo upoštevali, da lahko oseba pripada bodisi podmnožici $\mathcal{A}$ bodisi podmnožici $\mathcal{B}$ in je zato $p(x_2) = 1 - p(x_1)$.

Iz naloge razberemo še pogojne verjetnosti

$$P_{Y|X=x_1} = \big(p(y_1 \mid x_1)\,,\, p(y_2 \mid x_1)\,,\, p(y_3 \mid x_1)\big) = \left(\frac{1}{2}\,,\, \frac{3}{10}\,,\, \frac{2}{10}\right)\,, \text{ in}$$
$$P_{Y|X=x_2} = \big(p(y_1 \mid x_2)\,,\, p(y_2 \mid x_2)\,,\, p(y_3 \mid x_2)\big) = \left(\frac{3}{10}\,,\, \frac{1}{2}\,,\, \frac{2}{10}\right)\,.$$

Zanima nas vzajemna informacija med naključnima spremenljivkama X in Y, ki jo določimo kot

$$\begin{aligned} I(X,Y) &= H(Y) - H(Y \mid X) = \\ &= H(Y) - \big(p(x_1)H_{x_1}(Y \mid X = x_1) + p(x_2)H_{x_2}(Y \mid X = x_2)\big)\,. \end{aligned}$$

Ker je pogojna porazdelitev $P_{Y|X}$ znana, potrebujemo za izračun vzajemne informacije le še verjetnostno porazdelitev P_Y, ki jo lahko določimo z upoštevanjem zveze $p(y_j) = \sum_i p(x_i)p(y_j \mid x_i)$:

$$\begin{aligned} p(y_1) &= p(x_1) \cdot 0{,}5 + \big(1 - p(x_1)\big) \cdot 0{,}3 = \\ &= 0{,}2 \cdot p(x_1) + 0{,}3\,, \end{aligned}$$

$$\begin{aligned} p(y_2) &= p(x_1) \cdot 0{,}3 + \big(1 - p(x_1)\big) \cdot 0{,}5 = \\ &= -0{,}2 \cdot p(x_1) + 0{,}5\,, \end{aligned}$$

$$\begin{aligned} p(y_3) &= p(x_1) \cdot 0{,}2 + \big(1 - p(x_1)\big) \cdot 0{,}2 = \\ &= 0{,}2\,, \end{aligned}$$

kar nam da

$$P(Y) = \big(p(y_1)\,,\, p(y_2)\,,\, p(y_3)\big) = \big(0{,}2{\cdot}p(x_1){+}0{,}3\,,\, -0{,}2{\cdot}p(x_1){+}0{,}5\,,\, 0{,}2\big).$$

Na podlagi izračunanih verjetnosti določimo vse potrebne entropije:

$$H(Y \mid X = x_1) = H_{x_1}(0{,}5\,,\, 0{,}3\,,\, 0{,}2) \approx 1{,}49 \;\; [bitov]\;\;,$$

$$H(Y \mid X = x_2) = H_{x_2}(0{,}3\,,\, 0{,}5\,,\, 0{,}2) \approx 1{,}49 \;\; [bitov]$$

ter

$$H(Y) = H\big(0{,}2 \cdot p(x_1) + 0{,}3\,,\ -0{,}2 \cdot p(x_1) + 0{,}5\,,\ 0{,}2\big)$$

in zapišemo vzajemno informacijo kot funkcijo verjetnosti $p(x_1)$

$$\begin{aligned} I(X,Y) = I\big(p(x_1)\big) &= H(Y) - \big(p(x_1)H_{x_1} + p(x_2)H_{x_2}\big) = \\ &= H\big(0{,}2 \cdot p(x_1) + 0{,}3\,,\ -0{,}2 \cdot p(x_1) + 0{,}5\,,\ 0{,}2\big) - H_{x_1}, \end{aligned}$$

kjer smo upoštevali

$$p(x_1)H_{x_1} + p(x_2)H_{x_2} = p(x_1)H_{x_1} + \big(1 - p(x_1)\big)H_{x_2} = H_{x_1} = H_{x_2}.$$

Informacija bo maksimalna, ko bo parcialni odvod informacije $\nabla_{p(x_1)} I(p(x_1))$ po $p(x_1)$ enak nič, torej

$$\nabla_{p(x_1)} I(p(x_1)) = \frac{\partial H(Y)}{\partial p(y_1)} \cdot \frac{dp(y_1)}{dp(x_1)} + \frac{\partial H(Y)}{\partial p(y_2)} \cdot \frac{dp(y_2)}{dp(x_1)} + \frac{\partial H(Y)}{\partial p(y_3)} \cdot \frac{dp(y_3)}{dp(x_1)} = 0.$$

Ugotovimo, da je

$$\frac{\partial H(y)}{\partial p(y_j)} = -\left(\log_2 p(y_j) + \frac{1}{\ln 2} \right)$$

in

$$\begin{aligned} \frac{dp(y_1)}{dp(x_1)} &= \frac{d}{dp(x_1)}\big(0{,}2 \cdot p(x_1) + 0{,}3\big) = 0{,}2\ , \\ \frac{dp(y_2)}{dp(x_1)} &= \frac{d}{dp(x_1)}\big(-0{,}2 \cdot p(x_1) + 0{,}5\big) = -0{,}2\ , \\ \frac{dp(y_3)}{dp(x_1)} &= \frac{d}{dp(x_1)}(0{,}2) = 0\ , \end{aligned}$$

kjer smo uporabili zveze $\frac{d}{dx} \ln x = \frac{1}{x}$, $\log_2 x = \frac{\ln x}{\ln 2}$, $\frac{d}{dx} \log_2 x = \frac{1}{x \ln 2}$ za poenostavitev izraza

$$\frac{d}{dp}(-p \log_2 p) = -p \frac{\ln p}{\ln 2} = -\frac{\ln p}{\ln 2} - p \frac{1}{p \ln 2} = -\left(\log_2 p + \frac{1}{\ln 2} \right).$$

Izračunane odvode vstavimo v izraz za parcialni odvod informacije po spremenljivki $p(x_1)$ in dobimo

$$-0{,}2 \cdot \left(\log_2 p(y_1) + \frac{1}{\ln 2} \right) + 0{,}2 \cdot \left(\log_2 p(y_2) + \frac{1}{\ln 2} \right) - 0 = 0\ ,$$

oz. po poenostavitvi

$$\log_2 p(y_1) = \log_2 p(y_2) \ ,$$

od koder sledi

$$p(y_1) = p(y_2) \ .$$

V zgornjo enačbo vstavimo izraza za verjetnosti $p(y_1)$ in $p(y_2)$ (v odvisnosti od $p(x_1)$):

$$0{,}2 \cdot p(x_1) + 0{,}3 = -0{,}2 \cdot p(x_1) + 0{,}5$$

ter izraz poenostavimo

$$0{,}4 \cdot p(x_1) = 0{,}2 \ ,$$

kar nam da rezultat $p(x_1) = 0{,}5$.

Informacija bo torej maksimalna pri $p_{max}(x_1) = 0{,}5$ in bo znašala

$$\begin{aligned} I_{max} &= H\big(0{,}2 \cdot p_{max}(x_1) + 0{,}3\,,\ -0{,}2 \cdot p_{max}(x_1) + 0{,}5\,,\ 0{,}2\big) - H_{x_1} = \\ &= H\big(0{,}4\,,\ 0{,}4\,,\ 0{,}2\big) - H_{x1}\big(0{,}5\,,\ 0{,}3\,,\ 0{,}2\big) \approx \\ &\approx 0{,}036 \ \ [bitov] \ . \end{aligned}$$

5 Diskretni viri informacije

Diskretni vir informacije predstavlja podsistem komunikacijskega (oz. informacijskega) sistema, ki oddaja znake v diskretnih časovnih trenutkih. Pri tem črpa znake iz množice znakov oz. abecede vira $A = \{x_1, x_2, ..., x_a\}$ z močjo $a \in \mathbb{N}$. Diskretni vir informacije matematično predstavimo z [4]:

- nizom (oz. zaporedjem) n medsebojno odvisnih diskretnih naključnih spremenljivk $(X_1, X_2, ..., X_n)$, ki črpajo iz abecede vira A in kjer je $A = \{x_1, x_2, ...x_a\}$, ter
- vezano porazdelitvijo verjetnosti niza dolžine n znakov $(X_1, X_2, ..., X_n)$

  $$P(X_1{=}x_{i_1}, ..., X_n{=}x_{i_n}) = \left(p(x_{i_1}, ..., x_{i_n}) \geq 0 : (x_{i_1}, ..., x_{i_n}) \in A^n\right) ,$$

  kjer velja $\sum_{(x_{i_1},...,x_{i_n})\in A^n} p(x_{i_1}, ..., x_{i_n}) = 1$ in predstavlja množica A^n množico vseh urejenih n-teric (oz. nizov dolžine n znakov), ki jih vir lahko odda

  $$A^n = \{(x_{i_1}, ..., x_{i_n}) : x_i \in A^n \ ; \ i = 1{,}2, ..., n\} .$$

Pri obravnavi diskretnih virov informacije bomo posebno pozornost namenjali *stacionarnim virom*, za katere velja, da je

$$P(X_{k+1}{=}x_{i_1}, ..., X_{k+n}{=}x_{i_n}) = P(X_1{=}x_{i_1}, ..., X_n{=}x_{i_n})$$

pri poljubni izbiri naravnih števil n in k. Verjetnostne lastnosti stacionarnih virov so torej neodvisne od časa.

5.1 Entropija diskretnega stacionarnega vira

Lastno informacijo (na znak) sporočil, ki jih oddaja diskretni stacionarni vir, zapišemo kot

$$\begin{aligned} H_n &= \frac{1}{n} H(X_1, ..., X_n) = \\ &= -\frac{1}{n} K \sum p(x_{i_1}, ..., x_{i_n}) \log_d p(x_{i_1}, ..., x_{i_n}) \ , \end{aligned}$$

kjer je $K > 0$ in je $d > 1$ [4].

Podobno za nize, ki jih oddaja diskretni stacionarni vir informacij, zapišemo tudi *pogojno entropijo* n-tega znaka (za $n \geq 2$), ko poznamo $(n-1)$ predhodnih znakov

$$\begin{aligned} H'_n &= H(X_n \mid X_{n-1}, ..., X_1) = \\ &= -K \sum p(x_{i_1}, ..., x_{i_n}) \log_d p(x_{i_n} \mid x_{i_{n-1}}, ..., x_{i_1}) \ . \end{aligned}$$

Pokažemo, da tako lastna informacija kot tudi pogojna entropija konvergirata proti isti vrednosti

$$\lim_{n\to\infty} H'_n = \lim_{n\to\infty} H_n = H < \infty,$$

ki ji pravimo *entropija stacionarnega vira* [4].

Lastna informacija sporočil H_n in pogojna entropija n-tega znaka glede na $(n-1)$ predhodnih H'_n sta povezani preko naslednje zveze

$$H_n = \frac{n-1}{n} H_{n-1} + \frac{1}{n} H'_n \ .$$

***Naloga* 5.1**

Vzemimo, da smo iz enega (dovolj dolgega) niza ocenili povprečne lastne informacije znakov v n-členih podnizih za $n = 1, ..., 5$ in pri tem dobili naslednje rezultate $H_1 = 4{,}46$ *bitov*, $H_2 = 4{,}0$ *bitov*, $H_3 = 3{,}67$ *bitov*, $H_4 = 3{,}39$ *bitov*, in $H_5 = 3{,}15$ *bitov*. Določite pogojne entropije n-tega znaka, ko poznamo $(n-1)$ predhodnih znakov!

Rešitev:

Za izračun pogojnih entropij H'_n najprej preoblikujemo enačbo

$$\begin{aligned} H_n &= \frac{n-1}{n} H_{n-1} + \frac{1}{n} H'_n \ , \\ \frac{1}{n} H'_n &= H_n - \frac{n-1}{n} H_{n-1} \mid \cdot n \ , \end{aligned}$$

da dobimo

$$H'_n = nH_n - (n-1)H_{n-1}$$

in nato izračunamo vse zahtevane pogojne entropije

$$
\begin{aligned}
H'_2 &= 2H_2 - 1H_1 = 2 \cdot 4 - 4{,}46 = 3{,}54 \ \ [bitov] \ , \\
H'_3 &= 3H_3 - 2H_2 = 3{,}01 \ \ [bitov] \ , \\
H'_4 &= 4H_4 - 3H_3 = 2{,}25 \ \ [bitov] \ , \\
H'_5 &= 5H_5 - 4H_4 = 2{,}19 \ \ [bitov] \ .
\end{aligned}
$$

5.2 Ergodičnost stacionarnih virov

V nadaljevanju bomo pozornost namenjali predvsem tako-imenovanim *ergodičnim virom*, za katere velja, da se statistična povprečja funkcij, definirana nad nizi, ki jih oddajajo takšni viri, poljubno malo razlikujejo od časovnih povprečij teh funkcij z verjetnostjo, ki je blizu vrednosti ena [4].

Povedano drugače, v primeru ergodičnih stacionarnih virov imajo vsi oddani nizi znakov enake statistične lastnosti, zato moremo porazdelitev verjetnosti vira oceniti na podlagi zgolj enega dovolj dolgega niza oddanih znakov.

5.3 Odvečnost vira

Za informacijske vire definiramo mero odvečnosti vira O, ki meri redundantnost zapisa oddanih sporočil vira glede na najučinkovitejši zapis oddanih sporočil

$$
O = 1 - \frac{dejanska\ entropija\ vira}{največja\ možna\ entropija\ vira} = 1 - \frac{H_v}{K \log_d a} \ ,
$$

kjer je H_v entropija vira, $K > 0$, $d > 1$ in je a moč abecede znakov, iz katere črpajo sporočila, ki jih vir oddaja [4].

Odvečnost vira zavzema vrednosti na intervalu [0,1] in se pogosto podaja v odstotkih. Tipični pristop k zmanjšanju odvečnosti vira je uporaba kompresije.

5.4 Vir brez spomina

Stacionarnemu viru informacij, za katerega velja, da je verjetnost nastopa n-tega simbola v nizu neodvisna od vseh $(n-1)$ predhodnih simbolov

$$p(x_{i_n} \mid x_{i_{n-1}}, ..., x_{i_2}, x_{i_1}) = p(x_{i_n}),\ n = 2, 3, ...\ ,$$

pravimo *vir brez spomina* [4].

Iz predpostavljene neodvisnosti simbolov v nizu sledi, da moremo vezano verjetnost naključnih spremenljivk niza (tj., znakov oz. simbolov niza) predstaviti kot produkt robnih verjetnosti spremenljivk

$$p(x_{i_1}, ..., x_{i_n}) = p(x_{i_1})p(x_{i_2}) \cdots p(x_{i_n}) = \prod_{j=1}^{n} p(x_{i_j})\ ,$$

kjer je $x_i \in A$, $i = 1, 2, ..., a$ [4].

Za stacionarni vir brez spomina definiramo povprečno lastno informacijo na oddani znak

$$H_n = \frac{1}{n} H(X_1, X_2, ..., X_n) = \frac{1}{n}\Big(H(X_1) + \cdots + H(X_n)\Big)\ ,$$

ki se ob predpostavki, da so spremenljivke (oz. znaki) $X_1, ..., X_n$ neodvisne in enako porazdeljene, poenostavi v

$$H_n = \frac{1}{n} n H(X_1) = H(X_1) = -K \sum_{i=1}^{a} p(x_i) \log_d p(x_i)\ .$$

Kot opazimo, lastna informacija H_n ni odvisna od n in je zato enaka entropiji vira brez spomina

$$H = \lim_{n\to\infty} H_n = H_n = -K \sum_{i=1}^{a} p(x_i) \log_d p(x_i)\ ,$$

kjer je $K > 0$, $d > 1$ in je a moč abecede vira [4].

***Naloga* 5.2**

Imamo stacionarni ergodični vir brez spomina z dvojiško abecedo $A = \{0, 1\}$. V nizih znakov, ki jih vir oddaja, je v povprečju 60% ničel in 40% enic. Izračunaj entropijo vira ter njegovo odvečnost!

Rešitev:

Vir predstavimo z naključno spremenljivko V z naslednjo verjetnostno shemo

$$V \sim \begin{pmatrix} v_1\,, & v_2 \\ p(v_1)\,, & p(v_2) \end{pmatrix} = \begin{pmatrix} 0\,, & 1 \\ 0{,}6\,, & 0{,}4 \end{pmatrix} .$$

Ker imamo opravka z dvojiško abecedo, je moč množice A enaka $a = 2$.

Entropijo vira izračunamo kot

$$H(V) = H\big(p(v_1)\,,\, p(v_2)\big) = H\big(0{,}6\,,\, 0{,}4\big) \approx 0{,}97 \ \ [bitov]\,,$$

na podlagi izračunane entropije pa še odvečnost vira O,

$$O = 1 - \frac{H(V)}{\log_2 a} = 1 - \frac{0{,}97}{\log_2 2} = 1 - \frac{0{,}97}{1} = 0{,}03\ .$$

Odvečnost vira $O = 0{,}03$ nam pove, da moremo sporočila tega vira kodirat (v povprečju) s 3% krajšimi sporočili in pri tem še zmeraj zajamemo vso informacijo.

***Naloga* 5.3**

Imamo stacionarni ergodični vir brez spomina, ki oddaja sporočila sestavljena iz treh znakov (oz. simbolov) w_1, w_2 in w_3. V nizih znakov, ki jih oddaja vir je v povprečju delež simbolov w_1 enak p, delež simbolov w_2 je prav tako enak p, preostalo so simboli w_3. Določite vrednost p, pri kateri je odvečnost vira najmanjša. Koliko takrat znaša odvečnost!

Rešitev:

Iz naloge razberemo, da imamo opravka s stacionarnim ergodičnim virom brez spomina, ki ga opišemo z naslednjo verjetnostno shemo

$$V \sim \begin{pmatrix} w_1\,, & w_2\,, & w_3 \\ p\,, & p\,, & 1-2p \end{pmatrix} .$$

Vir črpa znake iz abecede z močjo $a = 3$.

Zapišemo izraz za odvečnost vira kot funkcijo verjetnosti nastopa simbolov p, pri čemer predpostavimo, da moremo omenjeno verjetnost nastopa oceniti iz relativne frekvence (oz. deleža) nastopa simbolov v

oddanih sporočilih,

$$O = 1 - \frac{H(V)}{\log_2 a} = 1 - \frac{H(p\,,\,p\,,\,1-2p)}{\log_2 a}\ .$$

Opazimo, da bo odvečnost minimalna, ko bo entropija vira $H(V)$ maksimalna. Poiskati moramo torej vrednost p, ki bo maksimirala vrednost entropije $H(p\,,\,p\,,\,1-2p)$. Postopamo enako kot v primeru naloge 4.3: izračunamo parcialni odvod entropije po p, ga enačimo z 0 in izraz poenostavimo

$$\nabla_p\ H(V) = \frac{\partial H(V)}{\partial p_{(w_1)}} \cdot \frac{dp_{(w_1)}}{dp} + \frac{\partial H(V)}{\partial p_{(w_2)}} \cdot \frac{dp_{(w_1)}}{dp} + \frac{\partial H(V)}{\partial p_{(w_3)}} \cdot \frac{dp_{(w_1)}}{dp} = 0\ ,$$

$$0 = \left(-\log_2 p - \frac{1}{\ln 2}\right)\cdot 1 + \left(-\log_2 p - \frac{1}{\ln 2}\right)\cdot 1 + \left(-\log_2(1-2p) - \frac{1}{\ln 2}\right)\cdot(-2)\ ,$$

$$0 = -2\log_2 p + 2\log_2(1-2p)\ ,$$

$$2\log_2 p = 2\log_2(1-2p)\ ,$$

$$p = 1 - 2p\ ,$$

$$3p = 1\ ,$$

$$p = \frac{1}{3}\ .$$

Če dobljeno verjetnost vstavimo v verjetnostno shemo vira V, dobimo

$$V \sim \begin{pmatrix} w_1\,, & w_2\,, & w_3 \\ \frac{1}{3}\,, & \frac{1}{3}\,, & \frac{1}{3} \end{pmatrix}\ ,$$

kar nam da entropijo vira

$$H(V) = H\left(\frac{1}{3}\,,\,\frac{1}{3}\,,\,\frac{1}{3}\right) = \log_2 3\ \ [bitov]$$

in odvečnost vira pri izračunani verjetnosti $p = \frac{1}{3}$

$$O = 1 - \frac{H(V)}{\log_2 a} = 1 - \frac{\log_2 3}{\log_2 3} = 0\ .$$

Pri izbiri verjetnosti $p = \frac{1}{3}$ je torej odvečnost vira enaka 0.

***Naloga* 5.4**

Imamo dvojiški stacionarni ergodični vir brez spomina

$$V \sim \begin{pmatrix} v_1\,, & v_2 \\ p(v_1)\,, & p(v_2) \end{pmatrix} = \begin{pmatrix} 0\,, & 1 \\ \frac{1}{4}\,, & \frac{3}{4} \end{pmatrix} ,$$

ki oddaja nize z dolžino $n = 3$. Določite entropijo vira na oddani niz!

Rešitev:

Na podlagi podanih podatkov lahko izračunamo entropijo na oddani znak, ki je enaka

$$H(V) = H\big(p(v_1)\,,\, p(v_2)\big) = H\left(\frac{1}{4}\,,\, \frac{3}{4}\right) \approx 0{,}811 \;\; [bitov] \;\; .$$

Ker moremo vezano verjetnost niza v primeru stacionarnih virov brez spomina zapisati kot produkt robnih verjetnosti znakov niza, lahko tudi entropijo niza $H(V_1, V_2, V_3)$ določimo kot vsoto entropij posameznih znakov, torej

$$H(V_1, V_2, V_3) = H(V_1) + H(V_2) + H(V_3) = 3H(V) \approx 2{,}43 \;\; [bitov] \;\; .$$

5.5 Vir s spominom

Stacionarnemu viru informacij, za katerega velja, da je verjetnost nastopa n-tega simbola v oddanem nizu odvisna od določenega števila predhodnih simbolov, pravimo *vir s spominom*.

Posebej nas zanima vrsta vira s spominom, pri kateri je verjetnost nastopa n-tega simbola (oz. znaka) v nizu odvisna le od enega predhodnega simbola (tj., $(n-1)$-tega simbola). Takšnemu viru pravimo *Markovov vir s spominom prvega reda*, zanj pa velja

$$P(X_n{=}x_{i_n} \mid X_{n-1}{=}x_{i_{n-1}}, ..., X_2{=}x_{i_2}, X_1{=}x_{i_1}) = P(X_n{=}x_{i_n}) \;\; ,$$

kjer je $x_{i_k} \in A$, $k = 1, ..., n$ in $n \geq 2$ [4].

Pri obravnavi Markovovih virov s spominom prvega reda so pomembne takoimenovane *prehodne verjetnosti* q_{ij}, ki predstavljajo pogojne verjetnosti oddaje znaka $x_j \in A$ v diskretnem časovnem trenutku n, če je pred njim (torej

v trenutku $(n-1)$) bil oddan znak $x_i \in A$

$$q_{ij} = P(X_n{=}x_j \mid X_{n-1}{=}x_i) \geq 0 \ ,$$

kjer velja $\sum_{j=1}^{a} q_{ij} = 1$, $n \geq 2$ in $i, j = 1, 2, ..., a$.

S pomočjo prehodnih verjetnosti q_{ij} zapišemo verjetnost oddaje znaka $x_j \in A$ v trenutku n

$$\begin{aligned} P(X_n{=}x_j) &= \sum_{i=1}^{a} P(X_{n-1}{=}x_i) P(X_n = x_j{=} \mid X_{n-1}{=}x_i) = \\ &= \sum_{i=1}^{a} P(X_{n-1}{=}x_i) q_{ij}, \text{ za } j = 1, 2, ..., a \ . \end{aligned}$$

Zgornje enačbe po navadi zapišemo v matrični in homogeni obliki kot

$$\mathbf{p}_n = \mathbf{p}_{n-1} \mathbf{P}_Q \ ,$$

kjer je $\mathbf{p}_n$ porazdelitev verjetnosti n-tega znaka

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

je $\mathbf{p}_{n-1}$ porazdelitev $(n-1)$-ga znaka

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

in je $\mathbf{P}_Q$ matrika prehodnih verjetnosti

$$\mathbf{P}_Q = i \downarrow \overset{j}{\longrightarrow} [q_{ij}] = \begin{bmatrix} q_{11} & \cdots & q_{1a} \\ \vdots & \ddots & \vdots \\ q_{a1} & \cdots & q_{aa} \end{bmatrix} \ ,$$

kjer sta $i, j = 1, 2, ..., a$ [4].

5.5.1 Stacionarni Markovov vir

Markovovem viru s spominom prvega reda, pri katerem porazdelitev znakov znotraj niza ni odvisna od trenutka oddaje (indeksa n) znaka, pravimo stacionaren Markovov vir. Zanj velja

$$\mathbf{p}_n = \mathbf{p}_{n-1} = \mathbf{p} \ ,$$

iz česar sledi

$$\mathbf{p} = \mathbf{p}\mathbf{P}_Q \ .$$

Porazdelitvi $\mathbf{p} = \big(p(x_1), ..., p(x_a)\big)$, ki zadošča zgornji enačbi in naredi Markovov vir stacionaren, pravimo *stacionarna porazdelitev vira*. Stacionarna porazdelitev vira torej ustreza levemu lastnemu vektorju matrike $\mathbf{P}_Q$.

Če upoštevamo lastnosti stacionarnega Markovovega vira, pokažemo, da je pogojna entropija H'_n neodvisna od časovnega indeksa n. Entropijo Markovovega vira zato izračunamo kot

$$H = \lim_{n\to\infty} H'_n = \sum_{i=1}^{a} p(x_i)H_i = -\sum_{i=1}^{a} p(x_i)K \sum_{j=1}^{a} q_{ij} \log_d(q_{ij}) \ ,$$

kjer je $K \geq 1$, $d \geq 2$ in $H_i = -K \sum_{j=1}^{a} q_{ij} \log_d(q_{ij})$ [4].

***Naloga* 5.5**

Imamo homogen stacionarni Markovov vir s spominom prvega reda z naslednjo matriko prehodnih verjetnosti

$$\mathbf{P}_Q = [q_{ij}] = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} \ .$$

Določi stacionarno porazdelitev $\mathbf{p}$ homogenega Markovovega vira in izračunaj njegovo entropijo!

Rešitev:

Za izračun stacionarne porazdelitve $\mathbf{p}$ uporabimo zvezo $\mathbf{p} = \mathbf{p}\mathbf{P}_Q$. Če upoštevamo $p_1 = p(x_1)$ in $p_2 = p(x_2)$, zapišemo

$$(p_1 \, , \, p_2) = (p_1 \, , \, p_2) \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} \ .$$

Na podlagi zgornje matrične enačbe zapišemo dve (med seboj odvisni) enačbi

$$p_1 = \frac{1}{3}p_1 + \frac{3}{5}p_2 \ ,$$
$$p_2 = \frac{2}{3}p_1 + \frac{2}{5}p_2 \ .$$

Prvo enačbo preuredimo

$$p_1 = \frac{1}{3}p_1 + \frac{3}{5}p_2 \quad \Rightarrow \quad \frac{2}{3}p_1 = \frac{3}{5}p_2 \quad \Rightarrow \quad p_1 = \frac{9}{10}p_2$$

in upoštevamo dejstvo, da je

$$p_1 + p_2 = 1$$

ter tako dobimo rezultat

$$p_2 = \frac{10}{19} \text{ in } p_1 = \frac{9}{19} .$$

Stacionarna porazdelitev Markovovega vira je torej določena s $\mathbf{p} = \left(\frac{9}{19}, \frac{10}{19}\right)$.

Izračun entropije podanega stacionarnega Markovovega vira izvedemo na podlagi matrike prehodnih verjetnosti $\mathbf{P}_Q$

$$H_1 = \sum_{j=1}^{2} q_{1j} \log_2(q_{1j}) = H\left(\frac{1}{3}, \frac{1}{3}\right) \approx 0{,}9183 \ [bitov] ,$$

$$H_2 = \sum_{j=1}^{2} q_{2j} \log_2(q_{2j}) = H\left(\frac{3}{5}, \frac{2}{5}\right) \approx 0{,}9710 \ [bitov] ,$$

$$H = p_1 H_1 + p_2 H_2 \approx \frac{9}{19} \cdot 0{,}9183 + \frac{10}{19} \cdot 0{,}9710 \approx 1{,}3375 \ [bitov] .$$

Entropija vira je enaka 1,3375 *bitov* na oddani znak.

***Naloga* 5.6**

Imamo homogen stacionarni Markovov vir z abecedo $A = \{0, 1\}$. Verjetnost, da se bo znak '0' ponovil, je enaka $\frac{3}{8}$, verjetnost, da bo znaku '0' sledil znak '1' je enaka $\frac{5}{8}$. Znak '1' ima verjetnost $\frac{3}{4}$, da bo ponovljen, in $\frac{1}{4}$, da mu bo sledil znak '0'.

a) Zapišite matriko prehodnih verjetnosti.

b) Vzemimo, da je vir v trenutku $n = 0$ oddal znak '1'. Določite verjetnost, da bo vir oddal znak '1' tudi v trenutku $n = 2$!

Rešitev:

Iz podatkov v nalogi razberemo, da so prehodne verjetnosti enake

$$q_{11} = P(X_n{=}x_1 \mid X_{n-1}{=}x_1) = \frac{3}{8}\ ,$$
$$q_{12} = P(X_n{=}x_2 \mid X_{n-1}{=}x_1) = \frac{5}{8}\ ,$$
$$q_{21} = P(X_n{=}x_1 \mid X_{n-1}{=}x_2) = \frac{1}{4}\ ,$$
$$q_{22} = P(X_n{=}x_2 \mid X_{n-1}{=}x_2) = \frac{3}{4}\ ,$$

kjer je zaloga vrednosti naključnih spremenljivk X_t (oz. oddanih simbolov) za $t = 1, 2, ..., n$ definirana kot $\mathcal{Z}(X_t) = \{x_1, x_2\} \sim \{\texttt{'0'}, \texttt{'1'}\}$.

Matriko prehodnih verjetnosti zato zapišemo kot

$$\mathbf{P}_Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} = \begin{bmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}\ .$$

Vir je v trenutku $n = 0$ oddal znak 0. Ker je vir znak že oddal, je verjetnost oddaje znaka 0 v trenutku $n = 0$ enaka ena in verjetnost oddaje znaka 1 je enaka nič. Zapišemo

$$\mathbf{p}_0 = \left(p_0(x_1)\,,\, p_0(x_2)\right) = (0\,,\,1)$$

in nato izračunamo še verjetnosti nastopa spremenljivk v:

- trenutku $n = 1$:

$$\mathbf{p}_1 = \mathbf{p}_0\mathbf{P}_Q = (0\,,\,1) \begin{bmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} = \left(\frac{1}{4}\,,\,\frac{3}{4}\right) \quad \text{in}$$

- trenutku $n = 2$:

$$\mathbf{p}_2 = \mathbf{p}_1\mathbf{P}_Q = \left(\frac{1}{4}\,,\,\frac{3}{4}\right) \begin{bmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} = \left(\frac{9}{32}\,,\,\frac{23}{32}\right)\ .$$

Verjetnost oddaje znaka 1 v trenutku $n = 2$ je torej enaka $\frac{23}{32}$.

***Naloga* 5.7**

Dvojiški stacionarni homogen Markovov vir je oddal niz

$$1001000110011010010 \ .$$

Statistično ocenite s kakšno verjetnostjo ta vir enico spremeni v ničlo in ničlo v enico. Iz statističnih ocen teh pogojnih verjetnosti določite še:

a) stacionarno porazdelitev podanega vira, ter

b) entropijo podanega vira.

Rešitev:

Kadar opazovani informacijski vir odda dovolj dolgi niz, moremo pogojne verjetnosti nastopa določenega simbola ob predpostavljenem predhodnem simbolu določiti statistično - s štetjem pojavljanja kombinacij, ki nas zanimajo. Za določitev pogojnih verjetnosti

$$P(X_n{=}'0' \mid X_{n-1}{=}'1') = \frac{P(X_n{=}'0' \, , \, X_{n-1}{=}'1')}{P(X_{n-1}{=}'1')} \ ,$$
$$P(X_n{=}'1' \mid X_{n-1}{=}'0') = \frac{P(X_n{=}'1' \, , \, X_{n-1}{=}'0')}{P(X_{n-1}{=}'0')} \ ,$$

je tako potrebno določiti robne verjetnosti $P(X_{n-1}{=}'1')$, $P(X_{n-1}{=}'0')$ ter vezane verjetnosti $P(X_n{=}'0' \, , \, X_{n-1}{=}'1')$ in $P(X_n{=}'1' \, , \, X_{n-1}{=}'0')$.

Vezano verjetnost $P(X_n{=}'0' \, , \, X_{n-1}{=}'1')$ določimo kot razmerje med številom parov simbolov v oddanem nizu n_u, kjer simbolu $'0'$ sledi simbol $'1'$, ter številom vseh oddanih parov n_t. Opazimo, da je število parov n_u enako 6

$$\underset{1}{\underline{10}}\underset{2}{\underline{10}}001\underset{3}{\underline{10}}01\underset{4}{\underline{10}}\underset{5}{\underline{10}}\underset{6}{\underline{10}}$$

in je število vseh parov n_t enako 18.

Iskana vezana verjetnost $P(X_n{=}'0' \, , \, X_{n-1}{=}'1')$ je torej enaka

$$P(X_n{=}'0' \, , \, X_{n-1}{=}'1') = \frac{n_u}{n_t} = \frac{6}{18} \ .$$

Na podoben način določimo še drugo vezano verjetnost

$$P(X_n{=}'\mathtt{1}'\,,\, X_{n-1}{=}'\mathtt{0}') = \frac{5}{18}\ .$$

Pri določanju robnih verjetnosti je potrebno upoštevati zgolj simbole, ki nastopajo na $(n-1)$ mestu v nizu. Pri štetju simbolov zato ne upoštevamo zadnjega simbola na desni, saj se nikoli ne pojavi kot predhodnik drugega simbola. Ugotovimo torej, da je število vseh simbolov, ki se pojavijo na $(n-1)$ mestu enako 18, pri čemer se simbol $'\mathtt{0}'$ pojavi deset–krat

$$1\underset{1\,2}{00}1\underset{3\,4\,5}{000}11\underset{6\,7}{00}11\underset{8}{0}1\underset{9}{0}\underset{10}{0}10$$

in simbol $'\mathtt{1}'$ osem-krat

$$\underset{1}{1}00\underset{2}{1}000\underset{3\,4}{11}00\underset{5\,6}{11}0\underset{7}{1}00\underset{8}{1}0\ .$$

Poudarimo še enkrat, da zadnjega simbola pri štetju ne upoštevamo, saj se ne pojavi kot predhodnik drugega simbola. Robne verjetnosti zato zapišemo kot

$$P(X_{n-1}{=}'\mathtt{0}') = \frac{10}{18}\ ,\ \text{in}$$
$$P(X_{n-1}{=}'\mathtt{1}') = \frac{8}{18}\ .$$

Na podlagi statističnih ocen določimo iskane pogojne verjetnosti

$$P(X_n{=}'\mathtt{0}' \mid X_{n-1}{=}'\mathtt{1}') = \frac{P(X_n{=}'\mathtt{0}'\,,\, X_{n-1}{=}'\mathtt{1}')}{P(X_{n-1}{=}'\mathtt{1}')} = \frac{\frac{6}{18}}{\frac{8}{18}} = \frac{3}{4}\ ,$$
$$P(X_n{=}'\mathtt{1}' \mid X_{n-1}{=}'\mathtt{0}') = \frac{P(X_n{=}'\mathtt{1}'\,,\, X_{n-1}{=}'\mathtt{0}')}{P(X_{n-1}{=}'\mathtt{0}')} = \frac{\frac{5}{18}}{\frac{10}{18}} = \frac{1}{2}\ ,$$

kar nam omogoča zapis matrike prehodnih verjetnosti vira

$$\mathbf{P}_Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}\ .$$

S pomočjo statistične analize oddanega niza smo določili matriko prehodnih verjetnosti opazovanega Markovovega vira, na podlagi katere izračunamo stacionarno porazdelitev vira preko zveze

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

ki opisuje dve med seboj odvisni enačbi

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

Po preurediti prve enačbe dobimo

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

in z upoštevanjem zveze $p_1 + p_2 = 0$

$$1 - p_2 = \frac{3}{2}p_2 \ ,$$

kar nas pripelje do rezultata

$$p_2 = \frac{2}{5} \quad , \quad p_1 = \frac{3}{5}$$

oziroma

$$\mathbf{p} = \left(\frac{3}{5}, \frac{2}{5}\right) .$$

V zgornjih enačbah smo z p_1 in p_2 označili $p_1 = p('\mathtt{0}')$ in $p_2 = p('\mathtt{1}')$.

Pri izračunu entropije opazovanega vira prav tako izhajamo iz ocenjene matrike prehodnih verjetnosti $\mathbf{P}_Q$. Izračunamo

$$H_1 = H\left(\frac{1}{2}, \frac{1}{2}\right) = 1 \ [bitov] \ ,$$
$$H_2 = H\left(\frac{3}{4}, \frac{1}{4}\right) \approx 0{,}9710 \ [bitov] \ ,$$

ter

$$H = p_1 H_1 + p_2 H_2 \approx \frac{3}{5} \cdot 1 + \frac{2}{5} \cdot 0{,}811 \approx 0{,}924 \ [bitov]$$

in vidimo, da je entropija vira enaka 0,924 *bitov* na oddani znak.

6 Kodiranje vira informacije

Prirejanju znakov ene abecede znakom druge abecede pravimo *kodiranje*. Pri *kodiranju vira* prirejamo znakom abecede vira znake ali nize znakov abecede koda, ki ustrezajo tehničnim zahtevam za prenos po komunikacijskem kanalu.

Kod vira določa trojica (A, B, f) [4], kjer je

- A abeceda vira moči a znakov,
- B abeceda koda moči b znakov in
- f injektivna preslikava $f : A \to E$, kjer je $E \subset B^*$ množica kodnih zamenjav (posameznih znakov ali nizov znakov abecede B).

V primeru, ko množico kodnih zamenjav E tvorijo enako dolgi nizi znakov abecede B, torej $E \subset B^m$, pravimo takšnemu kodu vira *enakomerni kod*. V primeru, ko množico kodnih zamenjav E tvorijo različno dolgi nizi znakov abecede B, torej $E \subset B^1 \cup B^2 \cup B^3 \cup \cdots$, pa pravimo takšnemu kodu vira *neenakomerni kod*.

Zaradi tehničnih zahtev pri prenosu znakov po dvojiškem komunikacijskem kanalu po navadi vir kodiramo z dvojiškim kodom. V tem primeru je $B = \{0\,,\,1\}$.

***Primer* 6.1**

Primer množice dvojiških kodnih zamenjav enakomernega koda je tako

$$E = B^3 = \{000\,,\,001\,,\,010\,,\,011\,,\,100\,,\,101\,,\,110\,,\,111\}\ .$$

Primer množice dvojiških kodnih zamenjav neenakomernega koda je tako

$$E = B^1 \cup B^2 \cup B^3 =$$
$$= \{0\,,\,1\,,\,00\,,\,01\,,\,10\,,\,11\,,\,000\,,\,001\,,\,010\,,\,011\,,\,100\,,\,101\,,\,110\,,\,111\}\ .$$

Pri nalogah v nadaljevanju se bomo omejili na dvojiške kode vira.

6.1 Uporabnost, enoličnost in trenutnost

Uporabni so le tisti kodi vira, ki omogočajo *enolično dekodiranje* vsakega možnega niza kodnih zamenjav iz množice E. Enakomerne kode je možno enolično dekodirati, če je $a \leq b^m$. Neenakomerni kodi omogočajo enolično dekodiranje, če nobena kombinacija dveh ali več kodnih zamenjav iz množice E ne sestavlja druge kodne zamenjave iz te množice.

Neenakomerni kod je *trenuten kod*, če omogoča sprotno dekodiranje vseh možnih nizov kodnih zamenjav iz množice E. To velja v primeru, ko nobena kodna zamenjava iz množice E ni predpona drugi kodni zamenjavi iz te množice.

***Primer* 6.2**

Kod vira po navadi podajamo s *kodno tabelo*. Vzemimo, da vir informacije oddaja štiri znake $A = \{x_1, x_2, x_3, x_4\}$. Primeri dvojiških kodov $\mathcal{A}$, $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{C}$, $\mathcal{D}$ tega vira so podani v spodnjih kodnih tabelah.

A	kod $\mathcal{A}$	kod $\mathcal{B}_1$	kod $\mathcal{B}_2$	kod $\mathcal{C}$	kod $\mathcal{D}$
x_1	00	0	0	0	0
x_2	01	0	1	10	01
x_3	10	1	00	110	011
x_4	11	10	01	111	0111

Med navedenimi kodi sta koda $\mathcal{B}_1$ in $\mathcal{B}_2$ neuporabna, ker ne omogočata enoličnega dekodiranja nizov kodnih zamenjav. Kod $\mathcal{B}_1$ ni uporaben, ker nima injektivne preslikave, saj je $f(x_1) = f(x_2)$. Kod $\mathcal{B}_2$ pa ne omogoča enoličnega dekodiranja, ker je kombinacija dveh $f(x_1)$ enaka $f(x_3)$, kombinacija $f(x_1)$ in $f(x_2)$ pa sestavlja kodno zamenjavo $f(x_4)$. Kodiran dvojiški niz 00 tako lahko dekodiramo kot niz dveh znakov x_1x_1 ali zgolj kot en sam znak x_3.

Kod $\mathcal{A}$ je enakomeren. Poleg injektivnosti preslikave zadošča tudi pogoju za enoličnost, saj je $a = 4 \leq b^m = 2^2 = 4$.

Koda $\mathcal{C}$ in $\mathcal{D}$ sta neenakoomerna koda, ki omogočata enolično dekodiranje, saj nobena kombinacija dveh ali več kodnih zamenjav iz njune množice E ne sestavlja druge kodne zamenjave iz te iste množice. Opazimo pa, da so pri kodu $\mathcal{D}$ določene kodne zamenjave predpone drugim kodnim zamenjavam. Kodna zamenjava $f(x_1)$ je predpona vsem drugim kodnim zamenjavam, kodna zamenjava $f(x_2)$ je predpona kodnima zamenjavama $f(x_3)$ in $f(x_4)$, kodna zamenjava $f(x_3)$ pa je predpona kodni zamenjavi $f(x_4)$. Kod $\mathcal{D}$ torej ni trenuten kod.

6.2 Mera gospodarnosti in učinkovitost koda vira

Vzemimo, da diskreten vir oddaja znake iz abecede $A = \{x_1, ..., x_a\}$ v skladu z porazdelitvijo verjetnosti $\big(p(x_1), ..., p(x_i), ..., p(x_n)\big)$, kjer je $p(x_i) = P(x_i) \geq 0$ in $\sum_{i=1}^{a} p(x_i) = 1$.

Vzemimo, da je dolžina kodne zamenjave $f(x_i)$ enaka n_i, kjer n_i označuje število znakov iz abecede B v kodni zamenjavi (nizu) $f(x_i) \in E$.

Mera gospodarnosti koda je definirana kot povprečna dolžina kodnih zamenjav glede na porazdelitev verjetnosti kodiranih znakov vira [4]. Mero gospodarnosti označujemo z $\overline{n}$ in jo izračunamo kot

$$\overline{n} = \sum_{i=1}^{a} p_i n_i \ . \tag{6.1}$$

Kot opazimo, je enota mere gospodarnosti *znak*.

Uspešnost koda vira je definirana kot razmerje med entropijo vira in povprečno dolžino kodnih zamenjav. Uspešnost koda označujemo z η in jo izračunamo kot razmerje

$$\eta = H/\overline{n} \cdot 100\% \ , \tag{6.2}$$

kjer H označuje entropijo vira [4].

***Naloga* 6.1**

Vzemimo, da stacionarni vir brez spomina oddaja štiri znake iz abecede $A = \{x_1, ..., x_a\}$ v skladu s porazdelitvijo $\big(p(x_1), p(x_2), p(x_3), p(x_4)\big) =$

(0,5 , 0,25 , 0,125 , 0,125). Izračunajte mero gospodarnosti in uspešnost vseh uporabnih kodov iz primera 6.2!

Rešitev:

Uporabni kodi iz navedenega primera, ki omogočajo enolično dekodiranje nizov kodnih zamenjav, so kodi $\mathcal{A}$, $\mathcal{C}$, $\mathcal{D}$. Entropija podanega vira na oddani znak je

$$H = H(0{,}5\,, 0{,}25\,, 0{,}125\,, 0{,}125) = 1{,}75 \ \ [bitov] \ .$$

Dolžine kodnih zamenjav ugotovimo iz kodnih tabel.

A	$P(x_i)$	kod $\mathcal{A}$	n_i	kod $\mathcal{C}$	n_i	kod $\mathcal{D}$	n_i
x_1	0,5	00	2	0	1	0	1
x_2	0,25	01	2	10	2	01	2
x_3	0,125	10	2	110	3	011	3
x_4	0,125	11	2	111	3	0111	4

Mere gospodarnosti $\overline{n}$ in uspešnosti kodov η izračunamo po izrazih (6.1) in (6.2). Rezultate izračunov obeh mer podaja spodnja tabela.

kod	$\overline{n}$	η
$\mathcal{A}$	2,0	87,5%
$\mathcal{C}$	1,75	100%
$\mathcal{D}$	1,875	93,3%

Iz rezultatov je razvidno, da je najbolj uspešen kod $\mathcal{C}$ in to kar 100%, kar se zgodi le v primerih, ko so vrednosti verjetnosti znakov vira enake 2 na negativno potenco, kot v tem primeru.

6.3 Pogoj za obstoj uporabnega neenakomernega koda

Pogoj za obstoj trenutnega koda (A, B, f), ki omogoča enolično in trenutno dekodiranje vseh nizov kodnih zamenjav, pri čemer ima abeceda vira A

moč a znakov, abeceda koda B moč b znakov in so dolžine kodnih zamenjav $n_1, ..., n_i, ..., n_a$, je, da morajo dolžine kodnih zamenjav zadoščati neenačbi *Krafta* in *McMillana* [4]

$$\sum_{i=1}^{a} b^{-n_i} \leq 1 \ . \tag{6.3}$$

Velja tudi obratno, da dolžine kodnih zamenjav uporabnega koda, ki omogoča enolično in trenutno dekodiranje vseh nizov kodnih zamenjav, vedno zadoščajo omenjeni neenačbi.

***Primer* 6.3**

Koda $\mathcal{B}_1$ in $\mathcal{B}_2$ iz primera 6.2 ne zadoščata neenačbi (6.3). Pri kodu $\mathcal{B}_1$ je neenačba

$$\sum_{i=1}^{a} b^{-n_i} = 2^{-1} + 2^{-1} + 2^{-1} + 2^{-2} = \frac{7}{4} \not\leq 1 \ .$$

To pomeni, da pri dolžinah kodnih zamenjav (1,1,1,2) nikakor ni možno določiti koda, ki bi omogočal enolično in trenutno dekodiranje vseh nizov kodnih zamenjav. Enako velja za kod $\mathcal{B}_2$ in neenačbo

$$\sum_{i=1}^{a} b^{-n_i} = 2^{-1} + 2^{-1} + 2^{-2} + 2^{-2} = \frac{6}{4} \not\leq 1 \ .$$

Z neenačbo *Krafta* in *McMillana* (6.3) tako moremo vedno preveriti, ali pri izbranih dolžinah kodnih zamenjav obstaja kod, ki omogoča enolično in trenutno dekodiranje vseh nizov kodnih zamenjav.

***Naloga* 6.2**

Ali obstaja uporabni trenutni kod $(A = \{x_1, \cdots, x_6\}, B = \{0, 1\}, f)$

1. z dolžinami kodnih zamenjav {2,2,3,3,3,3}?
2. z dolžinami kodnih zamenjav {1,3,3,3,3,3}?
3. z dolžinami kodnih zamenjav {1,3,3,3,4,4}?

Rešitev:

Obstoj uporabnih trenutnih kodov pri danih dolžinah kodnih zamenjav ugotovimo s preverjanjem njihovega zadoščanja neenačbi *Krafta* in *McMillana* (6.3).

1. Dolžine kodnih zamenjav {2,2,3,3,3,3} neenačbi

$$\sum_{i=1}^{a} b^{-n_i} = 2^{-2} + 2^{-2} + 2^{-3} + 2^{-3} + 2^{-3} + 2^{-3} = \frac{8}{8} \leq 1$$

zadoščajo, torej, uporabni trenutni kod pri teh dolžinah kodnih zamenjav obstaja.

2. Dolžine kodnih zamenjav {1,3,3,3,3,3} neenačbi

$$\sum_{i=1}^{a} b^{-n_i} = 2^{-1} + 2^{-3} + 2^{-3} + 2^{-3} + 2^{-3} + 2^{-3} = \frac{9}{8} \not\leq 1$$

ne zadoščajo, torej, uporabni trenutni kod pri teh dolžinah kodnih zamenjav ne obstaja.

3. Dolžine kodnih zamenjav {1,3,3,3,4,4} neenačbi

$$\sum_{i=1}^{a} b^{-n_i} = 2^{-1} + 2^{-3} + 2^{-3} + 2^{-3} + 2^{-4} + 2^{-4} = \frac{16}{16} \leq 1$$

zadoščajo, torej, uporabni trenutni kod pri teh dolžinah kodnih zamenjav obstaja.

6.4 Gospodarni kodi

Denimo, da imamo opravka z diskretnim stacionarnim virom brez spomina, ki oddaja znake $x_i \in A$ v skladu z verjetnostno porazdelitvijo $p(x_1), ..., p(x_a)$, kjer je $\sum_{i=1}^{a} p(x_i) = 1$. Pokažemo, da obstaja neenakomerni kod (A, B, f), katerega povprečna dolžina kodnih zamenjav se podreja naslednji neenačbi

$$\frac{H}{K \log_d b} \leq \overline{n} \leq \frac{H}{K \log_d b} + 1 \ , \tag{6.4}$$

kjer je b moč abeceda koda vira in je $H = -K \sum_{i=1}^{a} p(x_i) \log_d p(x_i)$ entropija stacionarnega vira brez spomina. Neenakomernemu kodu, ki ustreza

neenačbi, pravimo *gospodarni kod*. Za kod, ki je gospodaren in ima hkrati najmanjšo povprečno dolžino kodnih zamenjav $\overline{n}$ med vsemi gospodarnimi kodi danega vira, pravimo, da je *(globalno) optimalen* za dani vir [4].

Kot smo videli v primeru naloge 6.1, lahko spodnjo mejo neenačbe (6.4) in s tem 100% uspešnost koda dosežemo zgolj v primeru, ko so verjetnosti oddaje posameznih simbolov celoštevilske potence osnove b, kar je v praksi precej redko. Se pa moremo spodnji meji poljubno približati, če namesto posameznih znakov kodiramo bloke (oz. sporočila) dolžine r znakov, ki jih obravnavamo kot znake iz hiperabecede A^r. V tem primeru velja

$$\frac{H}{K \log_d b} \leq \frac{\overline{n}}{r} \leq \frac{H}{K \log_d b} + \frac{1}{r} \ . \tag{6.5}$$

Kot vidimo, se z daljšanjem dolžine sporočil r poljubno približamo spodnji meji neenačbe in s tem povečamo uspešnost koda [4].

6.5 Huffmanov kod

Algoritem za izgradnjo Huffmanovega koda omogoča sestavljanje gospodarnega trenutnega koda, če so podani abeceda vira $A = \{x_1, x_2, ..., x_a\}$, abeceda koda $B = \{0, 1\}$ in verjetnostna porazdelitev oddaje znakov iz abeceda vira $p(x_1), p(x_2), ..., p(x_a)$, kjer velja $\sum_{i=1}^{a} p(x_i) = 1$. Huffmanov kod je v praksi precej razširjen in se, na primer, uporablja za zmanjševanje razsežnosti digitalnih slik pri JPEG kompresiji.

Za Huffmanov kod je značilno, da imajo znaki z večjo verjetnostjo nastopa krajše kodne zamenjave, medtem ko imajo znaki z manjšo verjetnostjo oddaje krajše kodne zamenjave. Znaka z najmanjšima verjetnostma oddaje imata enako dolgi kodni zamenjavi in se razlikujeta le v zadnjem kodnem znaku.

Algoritem za izgradnjo Huffmanovega koda obsega naslednje tri korake, ki se jih izvede v skladu s spodnjimi navodili [4]:

1. korak:

- inicializiraj začetno abecedo $A_0 \leftarrow A$,
- znake v začetni abecedi A_0 uredi v skladu z verjetnostmi njihove oddaje $p(x_1) \geq p(x_2) \geq \cdots p(x_a)$,
- predzadnjemu znaku x_{a-1} v urejeni abecedi A_0 priredi kodni znak 0, zadnjemu znaku x_a kodni znak 1.

2. korak: Iteriraj od $j = 1$ do $j = a - 2$

- sestavi novo abecedo A_j z združitvijo zadnjih dveh znakov iz abecede A_{j-1}, pri tem novemu znaku pripiši vsoto verjetnosti združenih znakov,
- znake v abecedi A_j uredi v skladu z verjetnostmi njihove oddaje $p(x_{1_j}) \geq p(x_{2_j}) \geq \cdots p(x_{a_j})$,
- predzadnjemu znaku x_{a_j-1} v urejeni abecedi A_j priredi kodni znak 0, zadnjemu znaku x_{a_j} kodni znak 1.

3. korak:

- kodno zamenjavo za x_i sestavi tako, da vse znake iz B, ki so se pojavili z indeksom i, vzameš po vrsti od konca proti začetku.

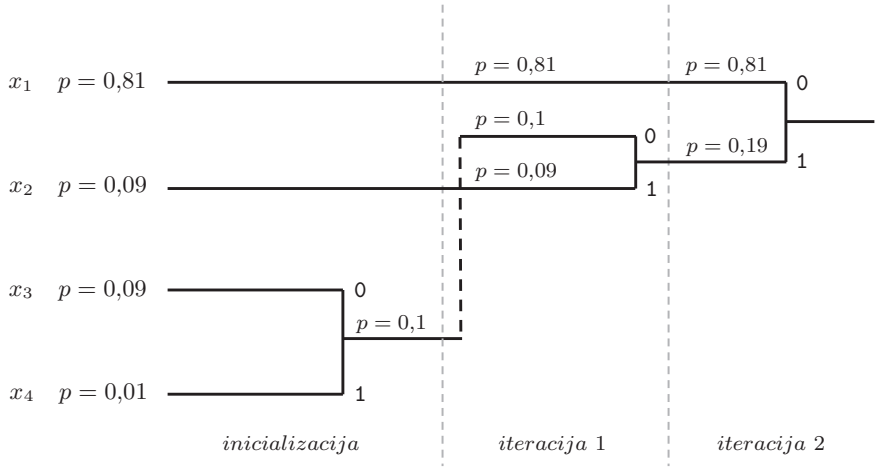
***Naloga* 6.3**

Vzemimo, da je podan vir brez spomina V, ki oddaja znake iz abecede $A = \{x_1, x_2, x_3, x_4\}$ in je moč abecede torej enaka $a = 4$. Vzemimo, da so verjetnosti oddaje znakov iz abecede A enake $p(x_1) = 0{,}81$, $p(x_2) = 0{,}09$, $p(x_3) = 0{,}09$ in $p(x_4) = 0{,}01$. Določite Huffmanov kod podanega vira, izračunajte povprečno dolžino kodnih zamenjav $\overline{n}$, entropijo vira $H(V)$ in uspešnost koda η!

Rešitev:

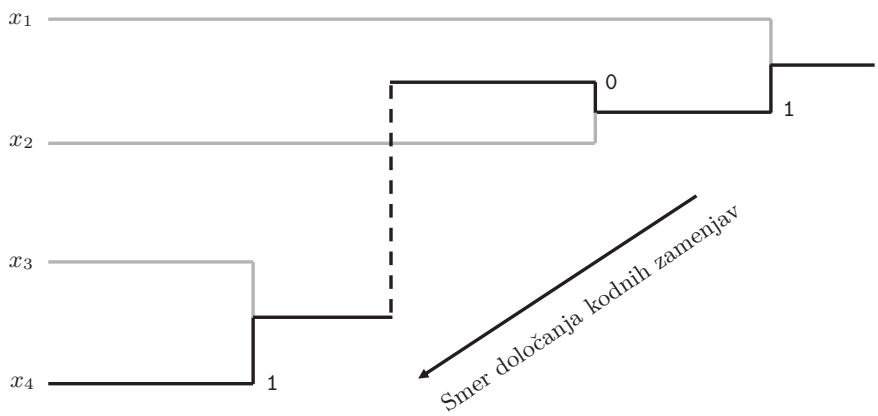
Iz naloge lahko razberemo, da so znaki iz abecede že urejeni v skladu s njihovimi verjetnostmi. Začetna abeceda $A_0 \leftarrow A$ je že ustrezno urejena, zato moramo v okviru prvega koraka (oz. inicializacije) zgolj še pripisati znakoma, ki imata najnižji verjetnosti oddaje, kodna znaka '0' in '1'. Kot vidimo na sliki 6.1, sta to v našem primeru znaka x_3 in x_4. Zato ju združimo in jima pripišemo skupno verjetnost nastopa $p = 0{,}1$.

V okviru prve iteracije (korak 2) tvorimo novo abecedo A_1, ki je tokrat sestavljena iz znakov x_1, x_2 ter združenega znaka (x_3, x_4), ki jim ustrezajo verjetnosti $p(x_1) = 0{,}81$, $p(x_2) = 0{,}09$ in $p((x_3, x_4)) = 0{,}1$. Znake uredimo v skladu z velikostmi njihovih verjetnosti. Znakoma, ki imata najnižji verjetnosti, priredimo kodna znaka '0' in '1'. Kot vidimo na sliki 6.1, je potrebno pri urejanju abecede A_1 v prvi iteraciji zamenjati vrstni red znaka x_2 ter združenega znaka (x_3, x_4), saj ima slednji večjo



Slika 6.1: Sestavljanje Huffmanovega koda.

verjetnost oddaje. V drugi iteraciji (korak 2) postopek še enkrat ponovimo in s tem zaključimo iteracije drugega koraka, saj smo združili že vse simbole iz abecede vira A.



Slika 6.2: Določanje kodnih zamenjav Huffmanovega koda - primer za x_4.

V zadnjem, tretjem koraku algoritma za vsak znak iz abecede A poiščemo pot od začetka drevesa do konca in zapišemo vse kodne znake,

ki jih prečkamo na poti v vzvratnem vrstnem redu. Zaporedje kodnih znakov, ki jih prečkamo od izbranega simbola do korena drevesa nam v vzvratnem vrstnem redu določa kodno zamenjavo Huffmanovega koda za izbrani simbol. Primer določitve kodne zamenjave za znak x_4 je prikazan na sliki 6.2, kjer lahko vidimo, da znaku x_4 ustreza kodna zamenjava 101.

Ostale kodne zamenjave so skupaj z njihovimi verjetnostmi oddaje ter dolžinami kodnih zamenjav podane v spodnji tabeli.

A	Huffmanov kod	n_i	p
x_1	0	n_1=1	$p(x_1)$=0,81
x_2	11	n_2=2	$p(x_2)$=0,09
x_3	100	n_3=3	$p(x_3)$=0,09
x_4	101	n_4=3	$p(x_4)$=0,01

Vidimo, da so bolj pogosti (oz. bolj verjetni) znaki iz A dejansko kodirani s krajšimi kodnimi zamenjavami.

Na podlagi podatkov v zgornji tabeli izračunamo še ostale iskane veličine

$$\begin{aligned}\overline{n} &= \sum_{i=1}^{a} p(x_i)n_i = \\ &= 0{,}81 \cdot 1 + 0{,}09 \cdot 2 + 0{,}09 \cdot 3 + 0{,}02 \cdot 3 = \\ &= 1{,}29 \ \ [znaka] \ ,\end{aligned}$$

$$\begin{aligned}H(V) &= H\big(p(x_1)\,,\, p(x_2)\,,\, p(x_3)\,,\, p(x_4)\big) = \\ &= H\big(0{,}81\,,\, 0{,}09\,,\, 0{,}09\,,\, 0{,}01\big) \approx \\ &\approx 0{,}938 \ \ [bitov] \ ,\end{aligned}$$

$$\eta = \frac{H(V)}{\overline{n}} \cdot 100\% = 72{,}7\% \ .$$

***Naloga* 6.4**

Vir brez spomina oddaja znake iz abecede $A = \{x_1, x_2\} = \{0, 1\}$ v skladu z naslednjo porazdelitveno shemo $\big(p(x_1), p(x_2)\big) = (0{,}9\,,\, 0{,}1)$.

Poiščite Huffmanov kod sporočil dolžine 3 znakov. Določite entropijo na znak in entropijo na sporočilo ter uspešnost dobljenega koda!

Rešitev:

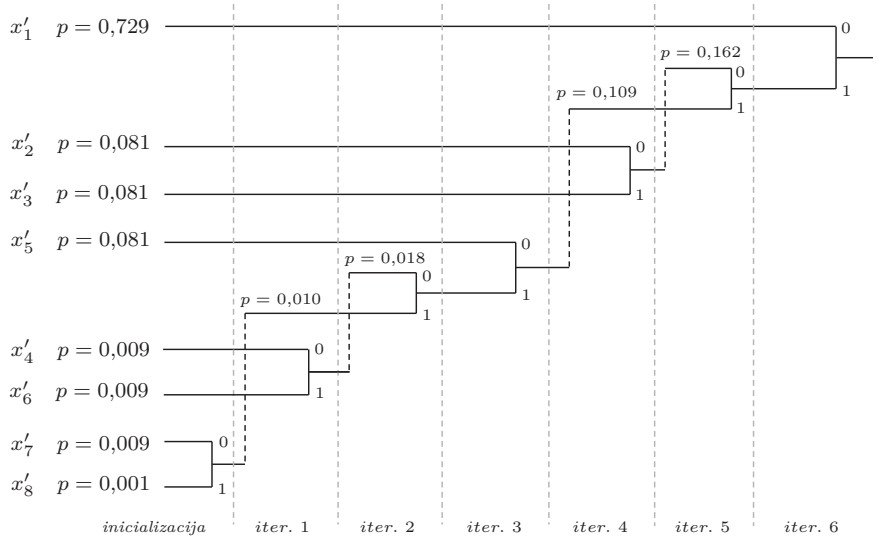
Iz naloge razberemo, da imamo opravka z dvojiškim virom, ki črpa iz abecede znakov A, katere moč je enaka $a = 2$. Ker želimo sestaviti Huffmanov kod za sporočila dolžine treh znakov, moramo najprej določiti novo hiperabecedo A^3, ki vsebuje vsa možna sporočila, ki jih je mogoče tvoriti na podlagi treh znakov iz A. Iz kombinatoričnega računa vemo, da je takšnih kombinacij $a^3 = 8$. Za vsako od osmih možnih sporočil moramo določiti še verjetnost oddaje, ki jo izračunamo na osnovi verjetnosti oddaje osnovnih znakov iz A.

Zapišemo torej novo abecedo A^3 in pripadajoče verjetnosti, kot so podane v spodnji tabeli.

A^3	p
$x'_1 = 000$	$p(x'_1)$=0,729
$x'_2 = 001$	$p(x'_2)$=0,081
$x'_3 = 010$	$p(x'_3)$=0,081
$x'_4 = 011$	$p(x'_4)$=0,009
$x'_5 = 100$	$p(x'_5)$=0,081
$x'_6 = 101$	$p(x'_6)$=0,009
$x'_7 = 110$	$p(x'_7)$=0,009
$x'_8 = 111$	$p(x'_8)$=0,001

Ker imamo opravka z virom brez spomina, so verjetnosti oddaje posameznega znaka v sporočilu neodvisne od verjetnosti oddaje drugih dveh znakov. Verjetnosti sporočil v zgornji tabeli smo zato določili kot produkt verjetnosti oddaje znakov, ki sporočilo sestavljajo. Za sporočilo $x'_2 = 001$, na primer, smo $p(x'_2)$ določili kot $p(x'_2) = p(x_1)\cdot p(x_1)\cdot p(x_2) = 0{,}9 \cdot 0{,}9 \cdot 0{,}1 = 0{,}081$. Podobno smo izračunali tudi verjetnosti za preostala sporočila.

Opazimo, da sporočila iz nove abecede A^3 tokrat niso urejena v skladu z velikostmi pripadajočih verjetnosti, zato jih je potrebno pred pričetkom algoritma za izgradnjo Huffmanovega koda še preurediti. Ko so sporočila urejena, pričnemo z izvajanjem algoritma in generiramo kodne



Slika 6.3: Izgradnja Huffmanovega koda.

zamenjave Huffmanovega koda za vseh osem sporočil, kot je ilustrirano na sliki 6.3. Rezultat postopka je osem kodnih zamenjav, ki so poleg verjetnosti oddaje ter dolžine kodnih zamenjav navedene v spodnji tabeli.

A^3	p	n'_i	Huffmanov kod
$x'_1 = 000$	$p(x'_1)$=0,729	$n'_1 = 1$	0
$x'_2 = 001$	$p(x'_2)$=0,081	$n'_2 = 3$	100
$x'_3 = 010$	$p(x'_3)$=0,081	$n'_3 = 3$	101
$x'_5 = 100$	$p(x'_5)$=0,081	$n'_5 = 3$	110
$x'_4 = 011$	$p(x'_4)$=0,009	$n'_4 = 5$	11100
$x'_6 = 101$	$p(x'_6)$=0,009	$n'_6 = 5$	11101
$x'_7 = 110$	$p(x'_7)$=0,009	$n'_7 = 5$	11110
$x'_8 = 111$	$p(x'_8)$=0,001	$n'_8 = 5$	11111

Izračunamo entropijo na oddani znak (oz. simbol iz abecede A), ki je enaka

$$H_1 = H(V) = H\big(p(x_1), p(x_2)\big) = H\big(0{,}9\,,\, 0{,}1\big) \approx 0{,}469 \ \ [bitov] \ .$$

Na podlagi te entropije lahko določimo še entropijo na sporočilo (oz. simbol iz hiperabecede A^3)

$$H_3 = H(V_1, V_2, V_3) = 3 \cdot H(V) \approx 1{,}4 \ \ [bitov] \ .$$

Če primerjamo uspešnost koda za primer, ko smo oddajali zgolj po en znak iz abecede A,

$$\overline{n} = 1\,[znak] \quad \Rightarrow \quad \eta = \frac{H_1}{\overline{n}} 100\% \approx \frac{0{,}469}{1} 100\% = 46{,}9\%$$

in uspešnost koda za primer, ko smo kodirali sporočila, sestavljena iz treh znakov,

$$\overline{n}_3 = \sum_{i=1}^{a^3} n_i' p(x_i') = 0{,}729 \cdot 1 + 0{,}081 \cdot 3 + \cdots + 0{,}001 \cdot 5 \approx 1{,}598\,[znakov] \quad \Rightarrow$$

$$\eta = \frac{H_3}{\overline{n}_3} 100\% \approx 88\% \ ,$$

ugotovimo, da smo s kodiranjem daljših sporočil bistveno izboljšali uspešnost koda - iz 46,9% na 88%.

6.5.1 Dekodiranje Huffmanovega koda

Za dekodiranje Huffmanovega koda nujno potrebujemo kodno tabelo, ki je bila uporabljena za kodiranje. Kodno tabelo zato pripnemo nizu kodnih zamenjav, ki jih prenašamo preko kanala. V praksi se kodna tabela zapiše v glavo kodirane datoteke in se pri dekodiranju prebere in uporabi za dekodiranje vsebine datoteke.

***Primer* 6.4**

Predpostavimo, da smo nize znakov kodirali s Huffmanovim kodom v skladu s kodno tabelo podano v nalogi 6.3, to je

A	Huffmanov kod
x_1	0
x_2	11
x_3	100
x_4	101

in smo k sprejemniku poslali naslednjo zaporedje znakov

$$011100100111010100 \ .$$

S pomočjo kodne tabele lahko enolično dekodiramo poslani niz, in sicer kot

$$\underset{x_1}{\underline{0}}\ \underset{x_2}{\underline{11}}\ \underset{x_3}{\underline{100}}\ \underset{x_3}{\underline{100}}\ \underset{x_2}{\underline{11}}\ \underset{x_4}{\underline{101}}\ \underset{x_1}{\underline{0}}\ \underset{x_3}{\underline{100}} \ .$$

Oddani niz bi torej dekodirali kot $x_1x_2x_3x_3x_2x_4x_1x_3$.

7 Tajno kodiranje

S tajnim kodiranjem sporočilo M, ki je navadno niz znakov iz abecede $\Sigma = \{x_1, \cdots, x_\sigma\}$, prikrijemo s šifrirnim postopkom E, za katerega obstaja dešifrirni postopek D, s katerim razkrijemo prvotno prikrito sporočilo - *tajnopis* ali *kriptogram*

$$D(E(M) = M \ .$$

Ključ K odpravlja zahtevo po tajnosti E in D. Ključ je parameter šifrirne in dešifrirne funkcije. Tajnopis potem zapišemo kot

$$C = E(M, K)$$

in

$$D(C, K) = D(E(M, K), K) = M$$

Poskus razkrivanja tajnopisa s strani oseb, ki jim sporočilo ni namenjeno, imenujemo *napad* na kriptografski sistem ali *kriptoanaliza* [4].

7.1 Vigenerjev kriptografski sistem

Vigenerjev kriptografski sistem je posplošitev osnovne zamisli zamenjave črk. Temelji na seštevanju znakov sporočila in znakov ključa po modulu σ, pri čemer ključ sestavlja d znakov in je σ moč abecede Σ. S takšnim načinom kriptiranja dosežemo, da so vsi znaki v sporočilu, ključu in tajnopisu iz iste abecede.

Cezarjev kriptografski sistem je Vigenerjev sistem z dolžino ključa $d = 1$, Vernamov pa z dolžino ključa, ki je enaka dolžini sporočila.

Šifrirni in dešifrirni postopek Vigenerjevega kriptografskega sistema zapišemo kot

$$E : c_i \ \equiv \ (m_i + k_i) \mod \sigma$$
$$D : m_i \ \equiv \ (c_i + k_i) \mod \sigma$$

kjer je: m_i številski kod i-tega znaka sporočila, k_i številski kod i-tega znaka ključa in c_i številski kod i-tega znaka tajnopisa [4]

Številski kodi znakov m_i, k_i in c_i so cela števila med 0 in $\sigma - 1$ ter so navadno kar zaporedni indeksi znakov v abecedi Σ. To pomeni, da Vigenerjev šifrirni postopek izvajamo tako, da vsoti $(m_i + k_i)$ po potrebi prištevamo ali odštevamo večkratnik σ, dokler ni rezultat v razponu med 0 in $\sigma - 1$. Podobno velja za dešifrirni postopek.

Izraz "$u \equiv v \mod m$" pomeni, da sta števili u in v *kongruentni* po *modulu* m, kar poenostavljeno pomeni, da je njuna razlika $u - v$ deljiva z m, oziroma da imata števili u in v enak ostanek pri deljenju z m. Pri danem številu v lahko kongruentna števila u določimo tudi po izrazu $u = k\,m + v$, kjer je k poljubno celo število.

***Primer* 7.1**

Izraz

$$11 \equiv 5 \mod 3$$

beremo: *števili* 11 *in* 5 *kongruentni po modulu* 3. To pomeni, da je njuna razlika, $11 - 5 = 6$, deljiva s 3, oziroma, da imata obe enak ostanek po celoštevilskem deljenju s 3, ta je 2.

Določimo nekaj števil, ki so kongruentna s številom 37 po modulu 10, torej

$$u \equiv 37 \mod 10\ .$$

Takšna števila so vsa, za katere velja

$$u = k\,10 + 37\ .$$

Nekaj teh števil je $17, 27, 47, 57,$ itd.

***Naloga* 7.1**

Denimo, da smo prejeli tajnopis $C =$ (PPMFČAPBA), za katerega vemo, da je bil ustvarjen z Vigenerjevim kriptografskim sistemom, ki uporablja abecedo Σ s samimi velikimi slovenskimi črkami in presledkom, kodirano po naslednji kodni tabeli.

A	B	C	Č	D	E	F	G	H	I	J	K	L
0	1	2	3	4	5	6	7	8	9	10	11	12
M	N	O	P	R	S	Š	T	U	V	Z	Ž	
13	14	15	16	17	18	19	20	21	22	23	24	25

Vemo tudi, da je bil pri šifriranju uporabljen ključ $K = (\texttt{LB})$. Dešifrirajte poslano sporočilo!

Rešitev:

Število znakov je $\sigma = 26$. Dešifriranje izvedemo tako, da kodam znakov tajnopisa c_i zaporedoma podpisujemo kode znakov ključa k_i. Nato izračunamo njihove razlike $(c_i - k_i)$, iz katerih določimo kode znakov prvotnega sporočila, ki so kongruentne tej razliki. To opravimo tako, da razliki $(c_i - k_i)$ po potrebi prištejemo ali odštejemo 26, tako da je rezultat ponovno v razponu med 0 in 26.

C	P	P	M	F	Č	A	P	B	A
c_i	16	16	13	6	4	0	16	1	0
K	L	B	L	B	L	B	L	B	L
k_i	12	1	12	1	12	1	12	1	12
$(c_i - k_i)$	4	15	1	5	-9	-1	4	1	-12
$m_i \equiv (c_i - k_i) \mod 26$	4	15	1	5	17	25	4	1	14
$M = D(C, K)$	D	O	B	E	R		D	A	N

Dešifrirano sporočilo je torej $M = (\texttt{DOBER DAN})$.

V nadaljevanju se bomo omejili le na kriptografske sisteme z javnim ključem, saj so sodobni kritpografski sistemi s tajnim ključem kot so IBM Lucifer, DES, 3-DES, ali AES računsko prezahtevni za naloge.

7.2 Kriptografski sistemi z javnim ključem

Kriptografski sistemi z javnim ključem temeljijo na navidezno enosmernih funkcijah, ki namesto enega samega uporabljajo dva med seboj odvisna, vendar različna ključa. Tak pristop odpravi potrebo po izmenjavi tajnih ključev pred prenosom sporočila.

Vsak uporabnik tovrstnega sistema mora razpolagati s/z

1. šifrirnim algoritmom E,
2. dešifrirnim algoritmom D, ter
3. parom ključev (K_e, K_d).

Če želi uporabnik U_i poslati tajnopis uporabniku U_j, šifrira sporočilo M tako, da vstavi v šifrirni algoritem E javni ključ K_e^j uporabnika U_j. Tako

dobi tajnopis

$$C = E(M, K_e^j) \ .$$

Uporabnik U_j dešifrira tajnopis C z dešifrirnim algoritmom D in svojim tajnim ključem K_d^j. Tako razkrije sporočilo

$$M = D(E(M, K_e^j), K_d^j) = D(C, K_d^j) \ .$$

Varnost kriptografskega sistema z javnim ključem temelji na predpostavki, da je računsko praktično nemogoče določiti tajni ključ K_d^j, tudi če poznamo E, D, K_e^j, C in M [4].

Primer takšnega kriptografskega sistema je sistem, ki uporablja šifrirni in dešifrirni postopek RSA (avtorjev Rivest, Shamir in Adleman).

Pred spoznavanjem postopka RSA na kratko povzemimo najnujnejše iz teorije števil.

7.2.1 Najnujnejše iz teorije števil

Deljivost celih števil

Število u je deljivo s številom v, če lahko u zapišemo kot produkt

$$u = k\,v \quad ,$$

kjer je k poljubno celo število. Številu v pravimo, da je *delitelj* števila u, oziroma tudi, da število v *deli* število u

Če je v naravno število, lahko vsako celo število u zapišemo

$$u = k\,v + r, \quad 0 \leq r < v \quad ,$$

kjer je r ostanek po deljenu u z v.

Če število v deli vsa števil $u_1, ..., u_n$, pravimo, da je v *skupni delitelj* teh števil. Največjemu skupnemu delitelju števil $u_1, ..., u_n$ pravimo *največja skupna mera* teh števil. Označimo jo z $NSM(u_1, ..., u_n)$. Števila $u_1, ..., u_n$ so si paroma *tuja*, če je $NSM(u_1, ..., u_n) = 1$ [4].

***Primer* 7.2**

Ostanek deljenja števila 39 s 7 je 4, ker velja $39 = 5 \cdot 7 + 4$. S kalkulatorjem ostanek izračunamo tako, da najprej delimo $39/7 \doteq 5{,}571$. Pri rezultatu nato upoštevamo samo celi del, to je 5, in izračunamo ostanek $r = 39 - 5 \cdot 7 = 4$.

Največja skupna mera števil $NSM(18, 27, 36) = 9$, ker je največji skupni delitelj teh treh števil 9.

Praštevila in sestavljena števila

Praštevilo je od 1 različno naravno število q, ki nima drugih deliteljev kot 1 in q. Naravno število, ki ni 1 in ni praštevilo, je *sestavljeno*. Vsako tako število lahko zapišemo kot produkt praštevilskih deliteljev - *prafaktorjev* [4]. Na primer

$$49896 = 2^3 \cdot 3^4 \cdot 7 \cdot 11 \ .$$

Eulerjeva funkcija

Naravnemu številu u je vsako izmed števil $0, 1, ..., u-1$ ali tuje ali pa ima kakšnega skupnega delitelja. S $\varphi(u)$ označimo število tistih števil med $0, 1, ..., u-1$, ki so številu u tuja. Funkcijo $\varphi(u)$ imenujemo *Eulerjeva funkcija*.

Za vsako praštevilo q velja

$$\varphi(q) = q - 1 \ ,$$

saj je med števili $0, 1, ..., q-1$ le število 0, ki ni tuje številu q [4].

Prav tako bi lahko ugotovili, da za tuji števili u in v velja

$$\varphi(u \cdot v) = \varphi(u) \cdot \varphi(v) \ .$$

***Primer* 7.3**

S številom $u = 8$ imajo med števili $0, 1, 2, 3, 4, 5, 6, 7$ števila $0, 2, 4, 6$ skupnega delitelja, števila $1, 3, 5, 7$ pa so tuja številu 8, zato velja $\varphi(8) = 4$.

Vrednost funkcije $\varphi(11) = 10$, ker je 11 praštevilo.

Vrednost funkcije $\varphi(35) = 24$, ker je $\varphi(35) = \varphi(5{\cdot}7) = \varphi(5){\cdot}\varphi(7) = 6{\cdot}4$.

7.2.2 Kriptografski sistem RSA

Varnost kriptografskega sistema z javnim ključem RSA temelji na zahtevnosti faktorizacije števila n, ki je produkt dveh velikih praštevil q in q'.

Sistem RSA določa naslednji šifrirni in dešifrirni postopek [4]:

1. Uporabnik U_i naključno izbere dve veliki praštevili q in q', izračuna njun produkt $n = q\,q'$ in Eulerjevo funkcijo $\varphi(n) = (q-1)(q'-1)$.

2. Naključno izbere število d, ki je tuje številu $\varphi(n)$ in torej zanj velja $NSM(d, \varphi(n)) = 1$. Pogoj $NSM(d, \varphi(n)) = 1$ izpolnjuje vsako praštevilo večje od $\max\{q, q'\}$. Števili d in n tvorita njegov tajni ključ $K_d = (d, n)$.

3. Nato izračuna še naravno število e, kjer je $1 < e < \varphi(n)$ in velja

$$e \cdot d \equiv 1 \mod \varphi(n) \ ,$$

   kar lahko zapišemo tudi kot

$$e \cdot d = k\,\varphi(n) + 1 \ ,$$

   kjer je k poljubno celo število. Števili e in n tvorita njegov javni ključ $K_e = (e, n)$, ki ga posreduje uporabniku U_j.

4. Uporabnik U_j informacijske bloke sporočila zakodira na običajen način v niz naravnih števil $M = (m_1, ..., m_i, ...)$, kjer je vsak $m_i < n$. Niz nato šifrira tako, da izračuna

$$c_i = E(m_i, K_e) = m_i^e \mod n \ .$$

5. Uporabnik U_i nato dešifrira prejeti tajnopis tako, da izračuna števila

$$m_i = D(c_i, K_d) = c_i^d \mod n \ ,$$

   iz katerih lahko na običajni način dekodira izvirne informacijske bloke poslanega sporočila.

Za vse uporabljene računske operacije so razviti računalniški algoritmi, ki hitro izvedejo potrebne izračune tudi za zelo velika praštevila q in q' (nekaj sto mestna decimalna števila oziroma nekaj tisoč mestna dvojiška števila). Iz opisanega šifrirnega in dešifrirnega postopka je razvidno, da je ob poznavanju javnega ključa $K_e = (e, n)$ možno priti do tajnega ključa $K_d = (d, n)$ le z izračunom Eulerjeve funkcije $\varphi(n) = (q-1)(q'-1)$, ki zahteva faktorizacijo števila n v produkt dveh neznanih praštevil q in q'. Zahtevnost te faktorizacije se šteje za NP-polni problem, ki je pri velikih številih praktično nerešljiv.

Omenimo še, da je od izbire q in q' odvisna tudi velikost informacijskih blokov m_i, ki jih še moremo kodirati. Velikost informacijskih blokov (oz. naravnih števil, s katerimi predstavljamo informacijo), ki jih lahko kodiramo s postopkom RSA, je omejena na $0 \leq m_i < n = qq'$. Pri tem sporočilo m_i tudi ne sme biti deljivo z q ali q', kar glede na velikost praštevil q in q' po navadi ni problematično.

***Naloga* 7.2**

S kriptografskim sistemom RSA želimo poslati osebi U_i tajno kodirano sporočilo. Oseba U_i ima javni ključ $K_e = (e = 5\,,\, n = 119)$ in tajni ključ $K_d = (d = 77\,,\, n = 119)$. Določite tajnopis, če je sporočilo $M = (66)$. Iz sprejetega tajnopisa nato dešifrirajte odposlano sporočilo.

Rešitev:

Sporočilo očitno vsebuje le en informacijski blok, ki je kodirani z decimalnim številom $m_1 = 66$. Uporabnik U_j določi tajnopis z izračunom

$$\begin{aligned} c_1 &= m_1^e \bmod n = \\ &= 66^5 \bmod 119 = \\ &= 1252332576 \bmod 119 = \\ &= 1252332576 - 119 \cdot 10523803 = \\ &= 19\,, \end{aligned}$$

kjer smo uporabili celi del ulomka $1252332576/119 \doteq 10523803{,}159$.

Uporabnik U_i izvede dešifriranje prejetega tajnopisa $C = (19)$ z izra-

čunom

$$\begin{aligned} m_1 &= c_1^d \bmod n = \\ &= 19^{77} \bmod 119 = \\ &= 291089...31139 \bmod 119 = \\ &= 291089...31139 - 119 \cdot 244613...3764967 = \\ &= 66 , \end{aligned}$$

kjer smo zopet uporabili celi del ulomka $19^{77}/119$. Sporočilo je tako pravilno dešifrirano kot $M = (66)$.

Naloga 7.3

Prestregli smo sporočilo v kriptografskem sistemu RSA. Prestreženi tajnopis je $C = (4)$. Vemo, da je tajnopis namenjen uporabniku U_i, katerega javni ključ je $K_e = (e = 5\,, n = 35)$. Ugotovite prvotno sporočilo $M = (m_1)$.

Rešitev:

Za uspešno dešifriranje prvotnega sporočila je potrebno najprej ugotoviti tajni ključ $K_d = (d = ?\,, n = 35)$. Za število d vemo, da mora veljati kongruenca

$$e \cdot d \equiv 1 \bmod \varphi(n) ,$$

pri čemer je $\varphi(n) = (q-1)(q'-1)$ in $n = q\,q'$. Znano število $n = 35$ moramo torej faktorizirati na produkt dveh praštevil q in q'. Ker je število n po vrednosti majhno, ugotovimo, da je

$$n = 35 = q\,q' = 5 \cdot 7 .$$

in

$$\varphi(35) = (5-1)(7-1) = 4 \cdot 6 = 24 .$$

Glede na to, da je znano število $e = 5$, velja kongruenca

$$5 \cdot d \equiv 1 \bmod \varphi(35) \quad \text{oziroma} \quad 5 \cdot d \equiv 1 \bmod 24 ,$$

kar zapišemo tudi kot

$$5 \cdot d = k\,24 + 1 ,$$

kjer je k poljubno celo število. Število d izpostavimo

$$d = \frac{k\,24 + 1}{5}$$

in s poskušanjem ugotovimo, pri katerem celem številu k dobimo celoštevilski rezultat. Pri $k = 1$ je rezultat celo število $d = 5$. S tem smo prišli do tajnega ključa, ki je v tem primeru kar enak $K_d = (d = 5\,,\, n = 35)$. Dešifriranje prejetega tajnopisa $C = (4)$ sedaj izvedemo z izračunom

$$\begin{aligned} m_1 &= c_1^d \bmod n = \\ &= 4^5 \bmod 35 = \\ &= 9\,. \end{aligned}$$

Sporočilo je tako dešifrirano kot $M = (9)$.

Poudarimo, da rešitev $d = 5$ ni enolična, saj dobimo celoštevilske rezultate za d tudi pri $k = 6$, $k = 11$, $k = 16$, itd. Pri teh vrednostih bi dobili $d = 29$, $d = 53$, $d = 77$ in tako naprej, ki bi vsi pravilno dekodirali prejeti tajnopis. Kot vidimo, so veljavne vrednosti za parameter d med seboj razmaknjene za $\varphi(35)$. Čeprav je teh vrednosti več, pa jih brez faktorizacije n ali $\varphi(35)$ tako rekoč ni mogoče uganiti oz. določiti.

***Naloga* 7.4**

Vzemimo, da smo prestregli RSA tajnopis v obliki zaporedja desetiških kod $C =$ (1,3,8,1,1,2,1,1,0,0,8,5,1,4,0,0,8,8,0,5,8,7,0,0) in da smo uspeli izvedeti, da je bil tajnopis ustvarjen z javnim ključem $K_e = (e = 3\,,\, n = 10)$. Iz javnega ključa določite tajni ključ in dekodirajte tajnopis v zaporedje desetiških kod prvotnega sporočila . Glede na to, da je $n = 10$, se bodo desetiške kode tajnopisa prekodirale v desetiške kode med 0 in 9. Vzemimo, da zaporedni pari desetiških kod predstavljajo kodne zamenjave velikih črk slovenske abecede (s presledkom) po spodnji kodni tabeli.

A	B	C	Č	D	E	F	G	H	I	J	K	L
00	01	02	03	04	05	06	07	08	09	10	11	12
M	N	O	P	R	S	Š	T	U	V	Z	Ž	
13	14	15	16	17	18	19	20	21	22	23	14	25

Kakšno je bilo torej izvirno sporočilo?

Rešitev:

Za uspešno dešifriranje prvotnega sporočila moramo najprej ugotoviti tajni ključ $K_d = (d = ?, n = 10)$. Za število d vemo, da mora veljati

kongruenca

$$e \cdot d \equiv 1 \mod \varphi(n) \ ,$$

pri čemer je $\varphi(n) = (q-1)(q'-1)$ in $n = q\,q'$. Znano število $n = 10$ moramo torej faktorizirati na produkt dveh praštevil q in q'. Ker je število n po vrednosti majhno, ugotovimo, da je

$$n = 10 = q\,q' = 2 \cdot 5 \quad \text{in} \quad \varphi(10) = (2-1)(5-1) = 1 \cdot 4 = 4 \ .$$

Glede na to, da je znano število $e = 3$, torej velja kongruenca

$$3 \cdot d \equiv 1 \mod \varphi(10) \quad \text{oziroma} \quad 3 \cdot d \equiv 1 \mod 4 \ ,$$

kar zapišemo tudi kot

$$3 \cdot d = k\,4 + 1 \ ,$$

kjer je k poljubno celo število. Število d izpostavimo

$$d = \frac{k\,4+1}{3}$$

in s poskušanjem ugotovimo, pri katerem celem številu k dobimo celoštevilski rezultat. Pri $k = 2$ je rezultat celo število $d = 3$. S tem smo prišli do tajnega ključa, ki je v tem primeru kar enak $K_d = (d = 3\,,\ n = 10)$.

Dešifriranje desetiških kod c_i prestreženega tajnopisa sedaj izvedemo z izračunom

$$m_i \ = \ c_i^3 \mod 10 \ .$$

Desetiške kode tajnopisa torej potenciramo na 3 in ugotavljamo ostanek po deljenju z 10. Ti izračuni so

$$\begin{aligned} m_1 &= c_1^3 \bmod 10 = 1^3 \bmod 10 = 1 \bmod 10 = 1 \ , \\ m_2 &= c_2^3 \bmod 10 = 3^3 \bmod 10 = 27 \bmod 10 = 7 \ , \\ m_3 &= c_3^3 \bmod 10 = 8^3 \bmod 10 = 512 \bmod 10 = 2 \ , \\ &\vdots \end{aligned}$$

Po dešifriranju desetiških kod tajnopisa pridemo do niza desetiških kod sporočila, ki je

$$M = (\texttt{1,7,2,1,1,8,1,1,0,0,2,5,1,4,0,0,2,2,0,5,2,3,0,0}) \ .$$

Zaporedne pare desetiških kod nato dekodiramo po dani kodni tabeli in pridemo do sporočila.

17	21	18	11	00	25	14	00	22	05	23	00
R	U	S	K	A		N	A	V	E	Z	A

Prvotno sporočilo je torej (`RUSKA NAVEZA`).

8 Komunikacijski kanali

Komunikacijski kanal poleg vira informacij in sprejemnika informacij predstavlja enega od treh ključnih elementov informacijsko-komunikacijskih sistemov. Kanal na vhodu sprejema znake, ki jih generira vir informacij in jih preslikuje v znake na izhodni, sprejemnikovi strani. Podobno kot ostale elemente komunikacijskih sistemov tudi komunikacijski kanal opišemo s pomočjo verjetnostne teorije, saj se pri prenosu informacije preko kanala pojavljajo naključne motnje.

8.1 Diskretni komunikacijski kanali

Teoretični model diskretnega komunikacijskega kanala predpostavlja, da poznamo množico znakov na vhodu v kanal, $U = \{x_1, ..., x_u\}$, množico znakov na izhodu iz kanala, $V = \{y_1, ..., y_v\}$, ter pogojne verjetnosti $\{p(\mathbf{y} \mid \mathbf{x})\}$, to so verjetnosti, da dobimo na izhodu kanala niz n znakov $\mathbf{y} = (y_1, ..., y_n) \in V^n$, če je na vhodu niz n znakov $\mathbf{x} = (x_1, ..., x_n) \in U^n$, pri $n \in \mathbb{N}$ [4].

Diskretni kanal *brez spomina* je kanala, za katerega za vsak $n \in \mathbb{N}$ velja

$$\begin{aligned} p(\mathbf{y} \mid \mathbf{x}) &= p((y_1, ..., y_k, ..., y_n) \mid (x_1, ..., x_k, ..., x_n)) = \\ &= p(y_1 \mid x_1) \cdots p(y_k \mid x_k) \cdots p(y_n \mid x_n) \ , \end{aligned}$$

kjer $p(y_k \mid x_k)$ označuje pogojno verjetnost, da se v k-tem koraku pri oddanem znaku $x_k \in U$ na izhodu kanala pojavi znak $y_k \in V$. Pri kanalu brez spomina je preslikava vhodnega znaka v izhodni znak naključni dogodek, ki je neodvisen od prejšnjih preslikav znakov.

Diskretni komunikacijski kanal brez spomina tako v celoti opisuje trojica $(U, \mathbf{P}_K, V)$, kjer so [4]:

- množica vhodnih znakov, $U = \{x_1, ..., x_i, ..., x_u\}$,
- množica izhodnih znakov, $V = \{y_1, ..., y_j, ..., y_v\}$, in
- verjetnostna matrika kanala, $\mathbf{P}_K$ z elementi

$$\mathbf{P}_K = i \downarrow \overset{j}{\underset{\longrightarrow}{[a_{ij}]}} , \quad a_{ij} = p(y_j \mid x_i) , \quad i = 1, ..., u, \quad j = 1, ..., v ,$$

za katere velja

$$\sum_{j=0}^{v} p(y_j \mid x_i) = 1 , \quad i = 1, ..., u .$$

Verjetnost $a_{ij} = p(y_j \mid x_i)$ označuje verjetnost, da se na izhodu pojavi znak $y_j \in V$, če se je na vhodu oddal znak $x_i \in U$. Posamezna vrstica matrike kanala $\mathbf{P}_K$ tako vsebuje porazdelitev pogojnih verjetnosti vseh možnih izhodnih znakov pri danem vhodnem znaku.

Nadaljnjo obravnavo diskretnih komunikacijskih kanalov bomo v celoti omejili na diskretne komunikacijske kanale brez spomina.

***Naloga* 8.1**

Imamo primer diskretnega komunikacijskega kanala brez spomina z dvema vhodnima znakoma in tremi izhodnimi znaki, ki ga opisuje trojka:

$$U = \{\alpha , \beta\} , \quad V = \{\alpha , \beta , \gamma\} , \quad \mathbf{P}_k = \begin{bmatrix} 0{,}9 & 0 & 0{,}1 \\ 0 & 0{,}9 & 0{,}1 \end{bmatrix} .$$

Izračunajte, kolikšna je verjetnost, da na izhodu podanega kanala prejmeš niz $\mathbf{y} = (\alpha , \gamma , \alpha , \beta)$, če je na vhodu v kanal oddan niz $\mathbf{x} = (\alpha , \alpha , \alpha , \beta)$.

Rešitev:

Ker je podan diskretni komunikacijski kanal *brez spomina*, je verjetnost $p(\mathbf{y} \mid \mathbf{x}) = p\big((\alpha , \gamma , \alpha , \beta) \mid (\alpha , \alpha , \alpha , \beta)\big)$ enaka produktu pogojnih verjetnosti za posamezne pare znakov. Te verjetnosti razberemo iz matrike $\mathbf{P}_k$ in rešitev je

$$\begin{aligned} p\big((\alpha , \gamma , \alpha , \beta) \mid (\alpha , \alpha , \alpha , \beta)\big) &= p(\alpha \mid \alpha)p(\gamma \mid \alpha)p(\alpha \mid \alpha)p(\beta \mid \beta) = \\ &= 0{,}9 \cdot 0{,}1 \cdot 0{,}9 \cdot 0{,}9 = \\ &= 0{,}0729 . \end{aligned}$$

Pri diskretnem komunikacijskem kanalu predpostavljamo, da se na vhodu kanala naključno pojavljajo znaki po neki porazdelitvi verjetnosti

$$p(x_i) \geq 0, \quad i = 1, ..., u, \quad \sum_{i=1}^{u} p(x_i) = 1 \ ,$$

kjer $p(x_i)$ označuje verjetnost, da se na vhodu kanala pojavi znak $x_i \in U$. Z upoštevanjem pogojnih verjetnosti $p(y_j \mid x_i)$ lahko določimo tudi (robno) porazdelitev verjetnosti pojavljanja znakov na izhodu kanala

$$p(y_j) = \sum_{i=1}^{u} p(x_i)p(y_j|x_i) \quad j = 1, ..., v, \quad \sum_{j=1}^{v} p(y_j) = 1 \ ,$$

kjer $p(y_j)$ označuje verjetnost, da se na izhodu kanala pojavi znak $y_j \in V$.

Vhod kanala lahko zato obravnavamo kot naključno spremenljivko X z zalogo vrednosti $\mathcal{Z}(X)$, ki jo priredimo abecedi U in s porazdelitvijo verjetnosti $P_X = \big(p(x_1), ..., p(x_u)\big)$, izhod kanala pa kot naključno spremenljivko Y z zalogo vrednosti $\mathcal{Z}(Y)$, ki jo priredimo abecedi V, in s porazdelitvijo verjetnosti $P_Y = \big(p(y_1), ..., p(y_v)\big)$.

Za par naključnih spremenljivk (X, Y) iz podanih porazdelitev verjetnosti izpeljemo vseh pet obravnavanih entropij:

$$H(X) = -K \sum_{i=1}^{u} p(x_i) \log_d p(x_i) \ ,$$

$$H(Y) = -K \sum_{i=1}^{u} p(y_j) \log_d p(y_j) \ ,$$

$$H(X, Y) = -K \sum_{i=1}^{u} \sum_{j=1}^{v} p(x_i)p(y_j \mid x_i) \log_d \big(p(x_i)p(y_j \mid x_i)\big) \ ,$$

$$H(Y \mid X) = -K \sum_{i=1}^{u} \sum_{j=1}^{v} p(x_i)p(y_j \mid x_i) \log_d p(y_j \mid x_i) \ ,$$

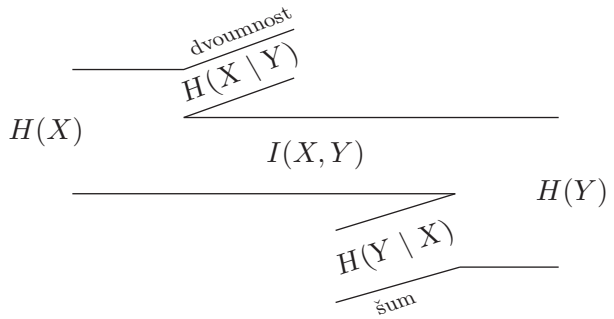
$$H(X \mid Y) = -K \sum_{i=1}^{u} \sum_{j=1}^{v} p(x_i)p(y_j \mid x_i) \log_d \left(\frac{p(x_i)p(y_j \mid x_i)}{p(y_j)} \right) \ ,$$

kjer je $K > 0$ in $d > 1$.

Matrika kanala $\mathbf{P}_K$ opisuje motnje pri prenosu znakov v diskretnem komunikacijskem kanalu. Prevajana informacija,

$$I(X,Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X) \ ,$$

je odvisna tako od kanala kot tudi od porazdelitve verjetnosti vhodnih znakov. Zgornjo enačbo in odnose med prevajano informacijo in entropijami lahko ponazorimo s sliko 8.1. *Entropija šuma* ali *irelevanca* $H(Y \mid X)$ je povprečna informacija, ki jo potrebujemo za določitve izhodnega znaka, če poznamo vhodni znak. *Dvoumnost* ali *ekvivokacija* $H(X \mid Y)$ je povprečna informacija, ki se izgubi v kanalu pri prenosu enega znaka [4].



Slika 8.1: Ponazoritev odnosa med prevajano informacijo, dvoumnostjo in entropijo šuma v komunikacijskem kanalu.

Prevajano informacijo lahko v celoti izpeljemo iz vhodne porazdelitve in matrike kanala:

$$I(X,Y) = -K \sum_{i=1}^{u} \sum_{j=1}^{v} p(x_i) p(y_j \mid x_i) \log_d \left(\frac{p(y_j \mid x_i)}{\sum\limits_{k=1}^{u} p(x_k) p(y_j \mid x_k)} \right) \ .$$

Prevajana informacija je torej funkcija vhodne porazdelitve P_X in matrike kanala $\mathbf{P}_K$, torej

$$I(X,Y) = f(P_X, \mathbf{P}_K) \ .$$

Pri danem kanalu lahko ugotavljamo, pri kakšni vhodni porazdelitvi bo prevajana informacija največja. Tej informaciji pravimo *kapaciteta komunikacijskega kanala* [4].

Kapaciteta diskretnega komunikacijskega kanala brez spomina je tako definirana kot:

$$C = \max_{P_X} \{ I(X,Y) \} \ .$$

Pri osnovi logaritma $d = 2$ je kapaciteta kanala podana v bitih na znak. Če poznamo čas τ, potreben za prenos enega znaka, potem lahko kapaciteto C' podamo tudi v bitih na sekundo

$$C' = \frac{C}{\tau} \ .$$

Čas, potreben za prenos enega znaka, določa hitrost kanala

$$\nu = \frac{1}{\tau}$$

v znakih na sekundo. Zvezo med kapacitetama kanal C in C' je potem tudi

$$C' = C\nu \ .$$

Po matriki kanala $\mathbf{P}_K$ že na pogled prepoznamo vrsto kanala. Prepoznamo lahko kanal brez izgub, kanal brez šuma, kanal brez motenj, neuporaben kanal in simetričen kanal [4].

8.1.1 Kanal brez izgub

Kanal brez izgub je kanal brez dvoumnosti. Prepoznamo ga po matriki kanala $\mathbf{P}_K$, ki ima po stolpcih le eno vrednost različno od nič [4]. Primera takšnih matrik sta

$$\mathbf{P}_k = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad \mathbf{P}_k = \begin{bmatrix} 0{,}5 & 0{,}5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \ .$$

Za kanale s takšnimi matrikami $\mathbf{P}_K$ ugotovimo, da je pogojna entropija (dvoumnost ali ekvivokacija)

$$H(X \mid Y) = 0 \ ,$$

ne glede na to kakšna je vhodna porazdelitev. V tem primeru je prevajana informacija

$$I(X,Y) = H(X) = H(Y) - H(Y \mid X) \ .$$

Največja prevajana informacija in s tem kapaciteta kanala brez izgub je potem enaka največji možni vhodni entropiji $H(X)$, torej

$$C = \max_{P_X}\{H(X)\} = H\left(\frac{1}{u}, ..., \frac{1}{u}\right) = \log_d u \ .$$

8.1.2 Kanal brez šuma

Kanal brez šuma je kanal, ki prenešeni informaciji ne doda dodatne entropije šuma, ki sicer poveča entropijo izhoda. Prepoznamo ga po matriki kanala $\mathbf{P}_K$, ki ima po vrsticah le eno vrednost različno od nič [4]. Primera takšnih matrik sta

$$\mathbf{P}_k = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad \mathbf{P}_k = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad .$$

Za kanale s takšnimi matrikami $\mathbf{P}_K$ ugotovimo, da je pogojna entropija

$$H(X \mid Y) = 0 \ ,$$

ne glede na to kakšna je vhodna porazdelitev. V tem primeru je prevajana informacija enaka izhodni entropiji, torej

$$I(X,Y) = H(Y) = H(X) - H(X \mid Y) \ .$$

Največja prevajana informacija in s tem kapaciteta kanala brez motenj je potem enaka največji možni izhodni entropiji $H(Y)$, torej

$$C = \max_{P_X}\{H(Y)\} = H\left(\frac{1}{v}, ..., \frac{1}{v}\right) = \log_d v \ .$$

8.1.3 Kanal brez motenj

Kanal brez motenj je kanal brez izgub in brez šuma. Prepoznamo ga po matriki kanala $\mathbf{P}_K$, ki ima po stolpcih in tudi vrsticah le eno vrednost, ki je enaka 1, vse preostale vrednosti so enake 0 [4]. To je mogoče le v primeru, ko je število vrstic in stolpcev enako, torej je tudi število vhodnih in izhodnih znakov enako, $u = v$. Primera takšnih matrik sta

$$\mathbf{P}_k = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad \mathbf{P}_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad .$$

Za kanale s takšnimi matrikami $\mathbf{P}_K$ ugotovimo, da sta obe pogojni entropiji

$$H(X|Y) = H(Y \mid X) = 0 \ ,$$

in

$$I(X,Y) = H(X) = H(Y) \ .$$

Entropiji $H(X)$ in $H(Y)$ sta enaki, ker je izhodna porazdelitev le druga razporedba vhodne porazdelitve (leva zgornje matrike) ali ista kot vhodna porazdelitev (desna zgornje matrike). Največja prevajana informacija in s tem kapaciteta kanala brez motenj je potem enaka največji možni izhodni entropiji $H(X) = H(Y)$, torej

$$C = \log_d u = \log_d v \ ; \ u = v \ .$$

8.1.4 Neuporaben kanal

Neuporaben kanal je kanal, ki ne prevaja informacije. Prepoznamo ga po matriki kanala $\mathbf{P}_K$, ki ima po stolpcih enake vrednosti (ima povsem enake vrstice) [4]. Primera takšnih matrik sta

$$\mathbf{P}_k = \begin{bmatrix} 0{,}5 & 0{,}5 \\ 0{,}5 & 0{,}5 \end{bmatrix} \quad , \quad \mathbf{P}_k = \begin{bmatrix} 0{,}2 & 0{,}3 & 0{,}5 \\ 0{,}2 & 0{,}3 & 0{,}5 \end{bmatrix} \ .$$

V tem primeru sta naključni spremenljivki X in Y neodvisni, ker zaradi enakosti vrednosti po stolpcih matrike $\mathbf{P}_K$ velja

$$p(y_j) = p(y_j \mid x_i) \ \ \forall i \ .$$

Posledično je $H(Y) = H(Y \mid X)$ in $I(X,Y) = H(Y) - H(Y \mid X) = 0$, torej je tudi kapaciteta kanala $C = 0$.

8.1.5 Simetričen kanal

Simetričen kanal prepoznamo po matriki kanala $\mathbf{P}_K$, ki ima po vrsticah in stolpcih le različne razporedbe istih vrednosti pogojnih verjetnosti [4]. Primera takšnih matrika sta

$$\mathbf{P}_k = \begin{bmatrix} 0{,}4 & 0{,}1 & 0{,}5 \\ 0{,}5 & 0{,}4 & 0{,}1 \\ 0{,}1 & 0{,}5 & 0{,}4 \end{bmatrix} \quad , \quad \mathbf{P}_k = \begin{bmatrix} 0{,}1 & 0{,}1 & 0{,}4 & 0{,}4 \\ 0{,}4 & 0{,}4 & 0{,}1 & 0{,}1 \end{bmatrix} \ .$$

Ker je entropijska funkcija $H(Y \mid X)$ neodvisna od razporedb porazdelitvenega zakona, ugotovimo, da je

$$H(Y \mid X) = H(Y \mid X = x_r)$$

enaka za vsak $r = 1, ..., u$. Kapaciteto simetričnega komunikacijskega kanala brez spomina potem izpeljemo kot

$$C = H\left(\frac{1}{v}, ..., \frac{1}{v}\right) - H\left(p(y_1 \mid x_i), ..., p(y_v \mid x_i)\right)$$

oziroma

$$C = \log_d v + \sum_{j=1}^{v} p(y_j \mid x_i) \log_d p(y_j \mid x_i) \ ,$$

kjer i predstavlja indeks poljubno izbrane vrstice matrike $\mathbf{P}_K$.

8.1.6 Dvojiški simetričen kanal

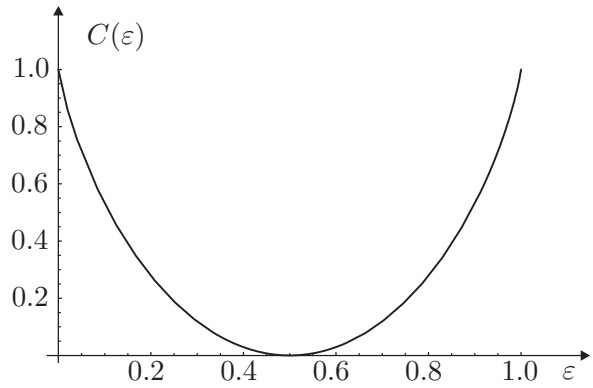
Dvojiški simetričen kanal določa matrika kanala [4]

$$\mathbf{P}_k = \begin{bmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix} \ ,$$

kjer ε predstavlja verjetnost napake na dvojiškem znaku. Njegova kapaciteta je

$$C = H\left(\frac{1}{2}, \frac{1}{2}\right) - H(1-\varepsilon\,,\,\varepsilon) = 1 + (1-\varepsilon)\log_2(1-\varepsilon) + \varepsilon \log_2 \varepsilon \ .$$

Kapaciteta je očitno funkcija napake ε. Iz lastnosti entropijske funkcije ugotovimo odvisnost kapacitete od napake ε. Ta odvisnost je ponazorjena na spodnji sliki.



***Naloga* 8.2**

Imamo izgubni dvojiški simetrični komunikacijski kanal brez spomina z verjetnostjo napake na dvojiškem simbolu $\varepsilon = 0{,}1$. Določite matriko $\mathbf{P}_K$ tega kanala in njegovo kapaciteto v bitih na dvojiški znak.

Rešitev:

Verjetnost napake na dvojiškem simbolu predstavlja obe pogojni verjetnosti $P(y_2 \mid x_1) = P(y_1 \mid x_2) = \varepsilon$. Preostali pogojni verjetnosti sta potem $P(y_1 \mid x_1) = P(y_2 \mid x_2) = 1 - \varepsilon$. Matriko kanala tako zapišemo kot

$$\mathbf{P}_k = \begin{bmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix} .$$

Kanal je simetričen, zato je njegova kapaciteta

$$\begin{aligned} C &= \log_2 2 - H(1-\varepsilon\,,\,\varepsilon) = 1 + (1-\varepsilon)\log_2(1-\varepsilon) + \varepsilon \log_2 \varepsilon = \\ &= 1 + (0{,}9)\log_2(0{,}9) + (0{,}1)\log_2(0{,}1) \approx \\ &\approx 0{,}531\,[bitov] \end{aligned}$$

na dvojiški znak.

***Naloga* 8.3**

Izračunajte kapaciteto diskretnega komunikacijskega kanala brez spomina (v bitih na sekundo) z naslednjo matriko kanala:

$$\mathbf{P}_k = \begin{bmatrix} \frac{1-p}{2} & \frac{1-p}{2} & \frac{p}{2} & \frac{p}{2} \\ & & & \\ \frac{p}{2} & \frac{p}{2} & \frac{1-p}{2} & \frac{1-p}{2} \end{bmatrix} ,$$

pri $p = 0{,}2$. Čas prenosa znaka je $\tau = 0{,}001$ sekunde. Skicirajte še odvisnost kapacitete od verjetnosti p kot funkcijo $C(p)$.

Rešitev:

V matriki so po stolpcih in vrsticah le različne razporedbe istih vrednosti, torej imamo opravka s simetričnim kanalom z $u = 2$ in $v = 4$

ter kapaciteto

$$\begin{aligned} C &= H\Big(\frac{1}{v},...,\frac{1}{v}\Big) - H(r_1,...,r_v) = \\ &= \log_2(v) + \sum_{i=1}^{v} r_i \log_2(r_i) \ = \\ &= \log_2(4) - H\Big(\frac{1-p}{2}\,,\ \frac{1-p}{2}\,,\ \frac{p}{2}\,,\ \frac{p}{2}\Big) = \\ &= 2 - H(0{,}4\,,\ 0{,}4\,,\ 0{,}2\,,\ 0{,}2) \approx \\ &\approx 0{,}278\,[bitov] \end{aligned}$$

na znak.

Izraz za kapaciteto lahko tudi razvijemo:

$$\begin{aligned} C &= \log_2(4) - H\Big(\frac{1-p}{2}\,,\ \frac{1-p}{2}\,,\ \frac{p}{2}\,,\ \frac{p}{2}\Big) = \\ &= 2 - \Big(-2\left(\frac{1-p}{2}\right)\log_2\left(\frac{1-p}{2}\right) - 2\,\frac{p}{2}\,\log_2\left(\frac{p}{2}\right)\Big) = \\ &= 2 - \Big(-(1-p)\,\log_2(1-p) + (1-p) - p\,\log_2(p) + p\Big) = \\ &= 2 - \Big(1 - p\,\log_2(p) - (1-p)\,\log_2(1-p)\Big) = \\ &= 1 - \Big(-p\,\log_2(p) - (1-p)\,\log_2(1-p)\Big) = \\ &= 1 - H(1-p\,,\,p) \ = \ 1 - H(0{,}8\,,\,0{,}2) \approx \\ &\approx 0.278\,[bitov] \end{aligned}$$

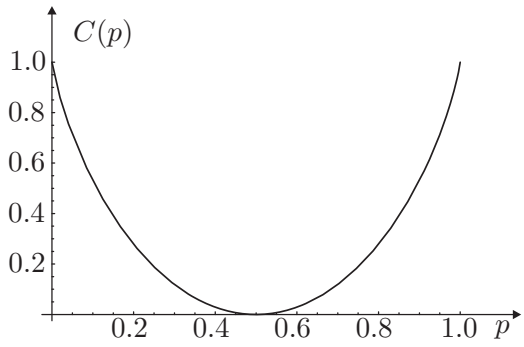
na znak. Kapaciteta in tudi $C(p)$ je torej enaka kot pri dvojiškem simetričnem kanalu:

$$C = 1 - H(1-p\,,\,p) = C(p)\ .$$

Odvisnost kapacitete od verjetnosti p je ponazorjena na spodnji sliki

Iskana kapaciteta v bitih na sekundo pa je:

$$C' = \frac{C}{\tau} \approx 278\left[\frac{bitov}{sekundo}\right]\ .$$



***Naloga* 8.4**

Diskretni stacionarni vir V oddaja štiri enako verjetne znake. Vir priključimo na kanal s kapaciteto $C' = 10$ bitov na sekundo. S kolikšno najvišjo hitrostjo ν v znakih na sekundo sme vir oddajati, da bi še bilo možno brez izgub prenašati oddane znake skozi kanal.

Rešitev:

Entropija vira V je dva bita na znak, ker je $H(V) = H(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}) = 2$. Oddane znake bi lahko brez izgub prenesli skozi kanal kvečjemu, če je entropija vira manjša ali enaka kapaciteti kanala C v bitih na znak. Ker je $C = C'/\nu$, ta pogoj zapišemo kot

$$H(V) \leq \frac{C'}{\nu} \quad \Longrightarrow \quad \nu \leq \frac{C'}{H(V)} \quad \Longrightarrow \quad \nu \leq 5 \ \left[\frac{znakov}{sekundo}\right] .$$

Najvišja hitrost oddajanja je torej 5 znakov na sekundo.

8.2 Zvezni komunikacijski kanali

Kapaciteto zveznega komunikacijskega kanala (v bitih na sekundo) določa izraz [4]:

$$C' = F \ \log_2 \left(1 + \frac{S}{N}\right) ,$$

kjer je F frekvenčna pasovna širina kanala v Hz, S je največja moč signala in N je največja moč šuma. S in N pogosto podajamo v obliki razmerja signal-šum, ki se meri v dB. Razmerje določa izraz [4]:

$$SNR = 10\ \log_{10}\left(\frac{S}{N}\right)\ [dB]\ .$$

V primeru, ko je dano razmerje signal šum v dB, določa kapaciteto izraz [4]:

$$C' = F\ \log_2\left(1 + 10^{\left(\frac{1}{10}\ SNR\right)}\right)\ .$$

***Naloga* 8.5**

Slikovni vzorčevalnik iz kamere daje sive slike S razsežnosti 625×625 slikovnih točk. Predpostavimo, da so svetlosti slikovnih točk T med sabo neodvisne in enakomerno kvantizirane na 256 sivih nivojev od črne do bele barve. Zakodirane slike želimo preko modema prenašati po starem analognem telefonskem kanalu s frekvenčno pasovno širino od 300 do $3400Hz$ in z razmerjem signal šum $20dB$. Izračunajte, koliko časa je najmanj potrebno za prenos dvojiške kodirane sive slike, če predpostavimo, da smo slikovne točke kodirali z enakomernim dvojiškim kodom.

Rešitev:

Lastna informacija slikovne točke je

$$I(T) = H(T) = H\left(\frac{1}{256}, \ldots, \frac{1}{256}\right) = \log_2(256) = 8\ \ [bitov]\ .$$

Lastna informacija celotne slike je

$$I(S) = H(S) = 625 \cdot 625 \cdot H(T)\ =\ 625 \cdot 625 \cdot 8\ =\ 3.125.000\,[bitov]\ .$$

Kapaciteta kanala (v bitih na sekundo) je

$$\begin{aligned} C' &= F\ \log_2\left(1 + 10^{\left(\frac{1}{10}\ SNR\right)}\right) = \\ &= (3400 - 300)\ \log_2\left(1 + 10^{\left(\frac{20}{10}\right)}\right) = \\ &= 3100\ \log_2(101) = \\ &= 20.640\left[\frac{bitov}{sekundo}\right]\ . \end{aligned}$$

Najkrajši potreben čas za prenos celotne slike je

$$\begin{aligned} C' &= C/\tau \;\Rightarrow\; C = C' \cdot \tau \;, \\ H(S) &= I(S) \;\leq\; C \;=\; C' \cdot \tau \;\Rightarrow \\ \tau &\geq H(S)/C' \;\Rightarrow \\ \tau &\geq 152 \;\; [sekund] \;\;, \end{aligned}$$

torej, najmanj 152 sekund za celo sliko.

9 Kod kanala

Znake oz. nize znakov, ki jih prenašamo preko komunikacijskega kanala kodiramo s kodom kanala. Namen takšnega kodiranja je odkrivanje ter morebitno popravljanje napak, ki se pojavljajo med prenosom.

Kod $\mathcal{K}(n,k)$ kanala $(U, \mathbf{P}_K, V)$ sestavljajo [4]:

- množica M informacijskih blokov D, kjer je vsak blok sestavljen kot niz k znakov iz abecede B,
- bijektivna kodirna funkcija h, ki preslika vsak informacijski blok $\mathbf{z}$ iz množice D v kodno zamenjavo $\mathbf{x}$, sestavljeno kot niz n znakov iz abecede U, torej

$$h : \mathbf{z} \in D \subseteq B^k \quad \rightarrow \quad \mathbf{x} \in K \subseteq U^n \ ,$$

- funkcija odločanja, ki preslika vsak sprejeti niz (vektor) $\mathbf{z}$, sestavljen iz n-znakov iz abecede V v niz veljavno kodno zamenjavo $\hat{\mathbf{x}} \in K$, torej

$$g : \mathbf{y} \in V^n \quad \rightarrow \quad \hat{\mathbf{x}} \in K \ .$$

Hitrost koda $\mathcal{K}(n,k)$ je določena z razmerjem [4]

$$R = \frac{\log_2 M}{n} \ .$$

V primeru, da je $M = 2^k$, je $R = \frac{k}{n}$. Hitrosti koda R pravimo tudi koeficient prenosa koda.

9.1 Dekodiranje koda kanala

Ker se pri naših obravnavah omejimo le na dvojiške kode, imamo tudi opravka le z dvojiškim dekodiranjem.

Iz zgornjih izrazov je razvidno, da pri $n > k$ vektorji $\mathbf{x} \in K \subseteq \{0,1\}^n$ zasedajo $M \leq 2^k$ točk prostora $\{0,1\}^n$ (t.j. prostor n-členih dvojiških nizov). Vektorji $\mathbf{y}$ pa zasedajo vse možne točke tega prostora.

Vektor napake je razlika med prejetim in oddanim dvojiškim nizom, torej $\mathbf{e} = \mathbf{y} - \mathbf{x}_i$. Vektor $\mathbf{e}$ nam pove, na katerem mestu v nizu je prišlo do napake.

Pri $n > k$ lahko pri dekodiranju napake odkrivamo in tudi popravljamo. Prostor $\{0,1\}^n$ razdelimo na podpodročja $\Omega_1, ..., \Omega_M$, kjer je

$$\Omega_i = \{\mathbf{y}_j : g(\mathbf{y}_j) = \mathbf{x}_i\} \ .$$

Prostor Ω_i torej obsega vse tiste nize $\mathbf{y}_j$, za katere se odločimo, da so lahko posledica prenosa niza $\mathbf{x}_i$ po danem kanalu [4].

Verjetnost pravilnega dekodiranja je potem enaka

$$p_{PD}(\mathbf{x}_i) = \sum_{\mathbf{y}_j \in \Omega_i} p(\mathbf{y}_j \mid \mathbf{x}_i) \ ,$$

verjetnost napačnega dekodiranja pa je $p_{ND}(\mathbf{x}_i) = 1 - p_{PD}(\mathbf{x}_i)$ [4].

Optimalno dekodiranje določa funkcija $\hat{\mathbf{x}} = g(\mathbf{y})$, pri kateri velja

$$p(\hat{\mathbf{x}} \mid \mathbf{y}) = \max_{1 \leq i \leq M} \{p(\mathbf{x}_i \mid \mathbf{y})\} \ .$$

Tako definirano preslikavo g imenujemo funkcija odločanja z najmanjšo verjetnostjo napake ali *idealni opazovalec* [4].

Verjetnost $p(\mathbf{x}_i \mid \mathbf{y})$ lahko zapišemo tudi kot:

$$p(\mathbf{x}_i \mid \mathbf{y}) = \frac{p(\mathbf{x}_i) p(\mathbf{y} \mid \mathbf{x}_i)}{\sum_{j=1}^{M} p(\mathbf{x}_j) p(\mathbf{y} \mid \mathbf{x}_j)} \ .$$

Očitno je, da iskani maksimum ni odvisen od imenovalca v zgornjem izrazu, ker ta ni odvisen od i-ja. Verjetnost $p(\mathbf{x}_i)$ je a priorna verjetnost celotnega informacijskega bloka. Če predpostavimo, da so vsi informacijski bloki enako verjetni, torej $p(\mathbf{x}_i) = \frac{1}{M}$, potem je maksimum odvisen le še od verjetnosti $p(\mathbf{y} \mid \mathbf{x}_i)$.

Pri teh predpostavkah funkcija odločanja $\hat{\mathbf{x}} = g(\mathbf{y})$, ki je določena kot

$$p(\mathbf{y} \mid \hat{\mathbf{x}}) = \max_{1 \leq i \leq M} \{p(\mathbf{y} \mid \mathbf{x}_i)\} \ ,$$

tudi omogoča optimalno dekodiranje in jo imenujemo *idealna funkcija odločanja* [4].

Pri diskretnih kanalih brez spomina je verjetnost $p(\mathbf{y} \mid \mathbf{x}_i)$ določena kot

$$p(\mathbf{y} \mid \mathbf{x}_i) = p(y_1 \mid x_{i_1}) \cdot p(y_2 \mid x_{i_2}) \cdots p(y_n \mid x_{i_n}) \ ,$$

torej kot produkt pogojnih verjetnost prenosa posameznega dvojiškega znaka. Pri idealni funkciji odločanja je v tem primeru iskani maksimum odvisen od maksimumov teh posameznih pogojnih verjetnosti. Sprejeti vektor dekodiramo tako, da dekodiramo vsak znak prejetega niza posebej. Idealno funkcijo odločanja potem določimo iz matrike pogojnih verjetnost kanala $\mathbf{P}_K$ tako, da poiščemo največje vrednosti verjetnosti po stolpcih te matrike [4].

***Naloga* 9.1**

Določite idealno funkcijo odločanja za diskretni kanal brez spomina, ki ga določa naslednja matrika pogojnih verjetnosti:

$$\mathbf{P}_K = i \downarrow \overset{j}{\underset{\longrightarrow}{[a_{ij}]}} = \begin{bmatrix} 0{,}1 & 0{,}8 & 0{,}1 \\ 0{,}0 & 0{,}0 & 1{,}0 \\ 0{,}9 & 0{,}0 & 0{,}1 \end{bmatrix} \ ,$$

kjer je $a_{ij} = p(y_j \mid x_i)$ in $i, j = 1, 2, 3$.

Rešitev:

Z iskanjem največjih vrednosti po stolpcih ugotovimo, da idealno funkcijo odločanja določajo sledeče preslikave posameznega znaka sprejetega vektorja: $y_1 \to x_3$, $y_2 \to x_1$ in $y_3 \to x_2$.

9.2 Idealna funkcija odločanja za dvojiški simetrični kanal

Omejimo se le na dvojiški kod in dvojiške simetrične kanale brez spomina. To pomeni, da so imamo opravka le z dvojiškimi abecedami ($B = U = V = \{0, 1\}$) in dvojiškimi nizi dolžin k in n, torej

$$\begin{aligned}\mathbf{z} &\in D \subseteq B^k = \{0,1\}^k \ , \\ \mathbf{x} &\in K \subseteq U^n = \{0,1\}^n \ , \\ \mathbf{y} &\in V^n = \{0,1\}^n \ .\end{aligned}$$

Idealna funkcija odločanja se pri dvojiškem simetričnem kanalu brez spomina poenostavi na določanje najmanjše Hammingove razdalje med prejetim in oddanim nizom dvojiških znakov [4].

Hammingova razdalja $d_h(\mathbf{x}, \mathbf{y})$ je definirana kot število mest, na katerih se niza $\mathbf{x}$ in $\mathbf{y}$ razlikujeta, torej

$$d_h(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \mid x_i - y_i \mid \ .$$

Razdalja $d_h\big((00110), (10101)\big)$ je denimo enaka 3.

Pri dvojiških simetričnih kanalih brez spomina lahko idealno funkcijo odločanja enostavno uresničimo tudi kot

$$d_H(\hat{\mathbf{x}}, \mathbf{y}) = \min_{1 \le i \le M} \{d_H(\mathbf{x}_i, \mathbf{y})\} \ .$$

Število napak, ki jih je kod dvojiškega simetričnega kanala brez spomina še sposoben popravljati pri idealni odločitveni funkciji, je odvisno od najmanjše razdalje med dvema veljavnima kodnima zamenjavama. Kod je z idealno funkcijo odločanja sposoben popravljati še vse e-kratne napake, če velja

$$\min_{i \ne j} \{d_h(\mathbf{x}_i, \mathbf{x}_j)\} = 2e + 1 \ ,$$

in je sposoben odkrivati še vse e-kratne napake, če velja

$$\min_{i \ne j} \{d_h(\mathbf{x}_i, \mathbf{x}_j)\} = 2e \ .$$

Hammingov pogoj, da bi lahko z idealno funkcijo odločanja pravilno dekodirali še vse odposlane nize po danem diskretnem dvojiškem simetričnem kanalu brez spomina, na katerih je prišlo do e ali manj napak, je določen z neenačbo [4]

$$\frac{2^n}{\sum\limits_{i=0}^{e} \binom{n}{i}} \ge M \ .$$

***Naloga* 9.2**

Ali lahko sestavimo kod s kodnimi zamenjavami $K \subset \{0,1\}^5$ za šest informacijskih blokov pri čemer želimo popravljati še vse enkratne napake?

Rešitev:

Potreben pogoj za obstoj takšnega koda je Hammingov pogoj

$$\frac{2^n}{\sum\limits_{i=0}^{e} \binom{n}{i}} \geq M \ .$$

Preverimo veljavnost pogoja pri $n = 5$, $M = 6$ in $e = 1$

$$\frac{2^n}{\sum\limits_{i=0}^{1} \binom{n}{i}} = \frac{2^5}{\binom{5}{0} + \binom{5}{1}} = \frac{32}{1+5} = \frac{32}{6} \not\geq 6 \ .$$

Pogoj torej ni izpolnjen in takšnega koda ne moremo sestaviti.

***Naloga* 9.3**

Dani so štirje dvojiški infomacijski bloki

$$D = \{\mathbf{z}_1 \, , \, \mathbf{z}_2 \, , \, \mathbf{z}_3 \, , \, \mathbf{z}_4\} = \{(00) \, , \, (01) \, , \, (10) \, , \, (11)\} \ .$$

Število informacijskih blokov je $M = 4 = 2^2$, torej je $k = 2$. Dana je bijektivna kodirna funkcija, ki preslika dvojiške informacijske bloke v sledeče kodne zamenjave

$$\begin{aligned} \mathbf{z}_1 = (00) &\ \rightarrow \ \mathbf{x}_1 = (00000) \ , \\ \mathbf{z}_2 = (01) &\ \rightarrow \ \mathbf{x}_2 = (01011) \ , \\ \mathbf{z}_3 = (10) &\ \rightarrow \ \mathbf{x}_3 = (10110) \ , \\ \mathbf{z}_4 = (01) &\ \rightarrow \ \mathbf{x}_4 = (11101) \ . \end{aligned}$$

Kod $\mathcal{K}$(5,2) dvojiškega simetričnega kanala brez spomina je tako določen z naslednjo množico kodnih zamenjav:

$$K = \{\mathbf{x}_1 \, , \, \mathbf{x}_2 \, , \, \mathbf{x}_3 \, , \, \mathbf{x}_4\} = \{(00000) \, , \, 01011) \, , \, (10110) \, , \, (11101)\} \ .$$

Na izhodni strani kanala smo prejeli niz $\mathbf{y} = (10100)$. Katera kodna zamenjava $\mathbf{x}_i$ (pri $i = 1, 2, 3, 4$) bi bila izbrana z idealno funkcijo odločanja? Vse kolikokratne napake je še sposoben popravljati podani kod z idealno funkcijo odločanja? Kolikšna je hitrost oziroma koeficient prenosa R danega koda?

Rešitev:

Idealno funkcijo odločanja, $\hat{\mathbf{x}} = g(\mathbf{y})$, uresničimo z iskanjem najmanjše Hammingove razdalje, torej

$$d_H(\hat{\mathbf{x}}, \mathbf{y}) = \min\{d_H(\mathbf{x}_1, \mathbf{y}) \,,\, d_H(\mathbf{x}_2, \mathbf{y}) \,,\, d_H(\mathbf{x}_3, \mathbf{y}) \,,\, d_H(\mathbf{x}_4, \mathbf{y})\} = \\ = d_H(\mathbf{x}_3, \mathbf{y}) \ ,$$

kjer je

$$d_H(\hat{\mathbf{x}}, \mathbf{y}) = d_H(\mathbf{x}_3, \mathbf{y}) = d_H\big((10110) \,,\, (10100)\big) = 1 \ .$$

Odločitev je torej $\hat{\mathbf{x}} = \mathbf{x}_3 = (10110)$.

Najmanjša kodna Hammingova razdalja med dvema različnima veljavnima kodnima zamenjavama je

$$\min_{i \neq j}\{d_h(\mathbf{x}_i, \mathbf{x}_j)\} = d_h(\mathbf{x}_1, \mathbf{x}_2) = 3 = 2e + 1 \quad \Rightarrow e = 1 \ .$$

Kod je torej z idealno funkcijo odločanja sposoben odkrivati še vse enkratne napake.

Glede na to, da je $M = 4 = 2^k = 2^2$ in da je $n = 5$, je hitrost oziroma koeficient prenosa danega koda $R = \frac{k}{n} = \frac{2}{5}$.

***Naloga* 9.4**

V dvojiški simetrični kanal z verjetnostjo napake na dvojiškem simbolu $p_n = \varepsilon = p(y_2 \mid x_1) = p(y_1 \mid x_2) = 0{,}45$ pošiljamo spodnje kodne zamenjave

$$K = \{\mathbf{x}_1 \,,\, \mathbf{x}_2 \,,\, \mathbf{x}_3\} \ = \ \{(000000) \,,\, (010010) \,,\, (111111)\} \ .$$

Verjetnosti posameznih informacijskih blokov in s tem kodnih zamenjav so

$$p(\mathbf{z}_1) = p(\mathbf{x}_1) = 0{,}025 \ ,$$
$$p(\mathbf{z}_2) = p(\mathbf{x}_2) = 0{,}175 \ ,$$
$$p(\mathbf{z}_3) = p(\mathbf{x}_3) = 0{,}8 \ .$$

Na izhodu kanala smo prejeli niz $\mathbf{y} = (100000)$. Kako bi se izvedlo dekodiranje, če bi enkrat uporabljali *idealnega opazovalca* in drugič *idealno funkcijo odločanja*?

Rešitev:

Pri idealni funkciji odločanja bi sprejeti $\mathbf{y}$ dekodirali kot $\mathbf{x}_1$, ker je Hammingova razdalja $d_H(\mathbf{x}_1, \mathbf{y}) = d_H\big((000000)\,,\,(100000)\big) = 1$ najmanjša med tremi razdaljami

$$\begin{aligned} d_H(\mathbf{x}_1\,,\,\mathbf{y}) &= d_H\big((000000)\,,\,(100000)\big) = 1\ , \\ d_H(\mathbf{x}_2\,,\,\mathbf{y}) &= d_H\big((001010)\,,\,(100000)\big) = 3\ , \\ d_H(\mathbf{x}_3\,,\,\mathbf{y}) &= d_H\big((111111)\,,\,(100000)\big) = 5\ . \end{aligned}$$

Pri dekodiranju z idealnim opazovalcem pa moramo poiskati

$$p(\hat{\mathbf{x}} \mid \mathbf{y}) = \max_{1\le i\le M}\{p(\mathbf{x}_i \mid \mathbf{y})\}\ .$$

Pogojne verjetnosti $p(\mathbf{x}_i \mid \mathbf{y})$ izračunamo kot

$$p(\mathbf{x}_i \mid \mathbf{y}) = \frac{p(\mathbf{x}_i)p(\mathbf{y} \mid \mathbf{x}_i)}{\sum\limits_{j=1}^{M} p(\mathbf{x}_j)p(\mathbf{y} \mid \mathbf{x}_j)}\ .$$

Verjetnosti $p(\mathbf{x}_i)$ so dane. Verjetnosti $p(\mathbf{y} \mid \mathbf{x}_j)$ izračunamo iz verjetnosti napak na posameznem dvojiškem simbolu, torej

$$\begin{aligned} p(\mathbf{y} \mid \mathbf{x}_1) &= p\big((100000) \mid (000000)\big) = p_n^1(1-p_n)^5 \approx 0.0226\ , \\ p(\mathbf{y} \mid \mathbf{x}_2) &= p\big((100000) \mid (001010)\big) = p_n^3(1-p_n)^3 \approx 0.0152\ , \\ p(\mathbf{y} \mid \mathbf{x}_3) &= p\big((100000) \mid (111111)\big) = p_n^5(1-p_n)^1 \approx 0.0101\ . \end{aligned}$$

Upoštevajmo, da je

$$p(\mathbf{y}) \;=\; \sum_{j=1}^{M} p(\mathbf{x}_j)p(\mathbf{y} \mid \mathbf{x}_j) \approx 0{,}013$$

in dobimo iskane verjetnosti

$$\begin{aligned} p(\mathbf{x}_1 \mid \mathbf{y}) &= \frac{p(\mathbf{x}_i)p(\mathbf{y} \mid \mathbf{x}_1)}{p(\mathbf{y})} \approx 0{,}049\ , \\ p(\mathbf{x}_2 \mid \mathbf{y}) &= \frac{p(\mathbf{x}_i)p(\mathbf{y} \mid \mathbf{x}_2)}{p(\mathbf{y})} \approx 0{,}234\ , \\ p(\mathbf{x}_3 \mid \mathbf{y}) &= \frac{p(\mathbf{x}_i)p(\mathbf{y} \mid \mathbf{x}_3)}{p(\mathbf{y})} \approx 0{,}716\ . \end{aligned}$$

Ker ima $p(\mathbf{x}_3 \mid \mathbf{y})$ največjo vrednost, bi sprejeti $\mathbf{y}$ dekodirali kot $\mathbf{x}_3$, torej drugače kot z idealno funkcijo določanja.

10 Varno kodiranje

V tem poglavju predstavimo izbrane primere postopkov izgradnje kodov, ki omogočajo na motnje odporno prevajanje informacije preko kanala. Postopki temeljijo na zamisli preverjanje sodosti števila enic pri določanju in preverjanju dvojiških kodnih zamenjav.

10.1 Linearni bločni kodi

Omejimo se zopet le na dvojiški kod. To pomeni, da imamo opravka le z dvojiškimi abecedami ($B = U = V = \{0, 1\}$), torej

$$\begin{aligned} \mathbf{z} &\in D \subseteq B^k = \{0,1\}^k \ , \\ \mathbf{x} &\in K \subseteq U^n = \{0,1\}^n \ , \\ \mathbf{y} &\in V^n = \{0,1\}^n \ , \end{aligned}$$

kjer D označuje množico M informacijskih blokov (vektorjev) $\mathbf{z}$ dolžine k, K označuje množico M kodnih zamenjav $\mathbf{x}$ z dolžino $n > k$, ki jo pošljemo v dvojiški kanal z motnjami, in $\mathbf{y}$ označuje kodno zamenjavo z isto dolžino n, ki jo sprejmemo na izhodni strani tega kanala (vseh možnih realizacij kodnih zamenjav $\mathbf{y}$ je 2^n). Razlika $m = n-k$, pri $n > k$, je število kontrolnih dvojiških znakov, ki jih dodajamo informacijskim blokom [4].

Pri določanju kodnih zamenjav varnega dvojiškega koda vedno najprej preverimo potreben Hammingov pogoj:

$$\frac{2^n}{\sum_{i=0}^{e} \binom{n}{i}} \geq M \ ,$$

kjer je M število kodnih zamenjav (po navadi 2^k), $n > k$ dolžina kodnih zamenjav in e najmanjše število napak na dvojiških znakih, ki jih s kodom še lahko popravimo.

Za preverjanje sodosti zaenkrat uporabljamo le vsoto po modulu 2, ki jo določa spodnja tabela.

$+$	0	1
0	0	1
1	1	0

Varno kodiranje s preverjanjem sodosti je določeno kot bijektivna preslikava:

$$h: \quad \mathbf{z} = (z_1, ..., z_{n-1}) \in \{0,1\}^{n-1} \quad \longmapsto \quad \mathbf{x} = (z_1, ..., z_{n-1}, x_n) \in \{0,1\}^n \ ,$$

kjer je

$$x_n \equiv \sum_{i=1}^{n-1} z_i \bmod 2 \ .$$

Za vsako kodno zamenjavo $\mathbf{x} = (x_1, ..., x_{n-1}, x_n)$ torej velja:

$$0 \equiv \sum_{i=1}^{n} x_i \bmod 2 \ ,$$

kjer je $(x_1, ..., x_{n-1}) = (z_1, ..., z_{n-1})$.

Povedano preprosteje, izvorno kodno zamenjavo $\mathbf{z}$ podaljšamo za en kontrolni dvojiški znak, torej $n = k + 1$ in $m = n - k = 1$. Kontrolni dvojiški znak določimo tako, da postane število enic v kodni zamenjavi $\mathbf{x}$ sodo $(0, 2, 4, ...)$.

Varno kodiranje s preverjanjem sodosti omogoča samo odkrivanje še vseh enkratnih napak, ne omogoča pa popravljanje napak, ker je minimalna Hammingova razdalja med kodnimi zamenjavami vedno enaka 2, torej [4]

$$\min\{d_H(\mathbf{x}_i, \mathbf{x}_j)\} = 2 \geq 2e + 1, \quad \text{pri} \quad e = 0 \ , \quad \text{oziroma}$$
$$\min\{d_H(\mathbf{x}_i, \mathbf{x}_j)\} = 2 \geq 2e, \quad \text{pri} \quad e = 1 \ .$$

***Naloga* 10.1**

Vzemimo, da želimo po kanalu z motnjami prenašati informacijske bloke iz naslednje množice:

$$D = \{(00)\,,\ (01)\,,\ (10)\,,\ (11)\} \ .$$

Določite kodne zamenjave varnega koda s preverjanjem sodosti. Vse kolikokratne napake je ta kod sposoben popravljati ali vsaj odkrivati?

Rešitev:

Število informacijskih blokov je $M = 4$ in njihova dolžina je $k = 2$. Dolžina kodnih zamenjav je $n = k+1 = 3$ in število kontrolnih dvojiških znakov $m = n - k = 1$. Preverimo najprej Hammingov pogoj

$$e_{max} = \arg\max_{e} \left\{ \frac{2^3}{\sum_{i=0}^{e} \binom{3}{i}} \geq 4 \right\} = 0 \quad .$$

Kod torej ne more biti sposoben popravljati napak. Kodne zamenjave, pri katerih smo dodali kontrolni dvojiški znak (ponazorjeno z ;) in poskrbeli za sodo število enic, je naslednji:

$$K = \{(00;0)\,,\ (01;1)\,,\ (10;1)\,,\ (11;0)\} \ .$$

Ker kod ni sposoben popravljati napak, preverimo, koliko jih je sposoben odkrivati. Varno kodiranje s preverjanjem sodosti omogoča samo odkrivanje še vseh enkratnih napak, ne omogoča pa popravljanje napak, ker je minimalna Hammingova razdalja med kodnimi zamenjavami vedno enaka 2,

$$\min\{d_H(\mathbf{x}_i, \mathbf{x}_j)\} = 2 \geq 2e \ , \quad \text{pri} \quad e = 1 \ .$$

Sposoben je torej odkrivati še vse enkratne napake, $e = 1$.

10.2 Določanje linearnih bločnih kodov

Osnovna zamisel preverjanja sodosti se je najprej razširila na vodoravno in navpično preverjanje sodosti. Nato pa se je razširila še v določanje in reševanje sistema linearno neodvisnih enačb nad Galoisovim obsegom GO(2) [4].

Galoisov obseg GO(2) je določen z množico elementov $\{0, 1\}$ ter operacijama seštevanja in množenja, predstavljenima v spodnjih tabelah.

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Koeficiente sistema m linearno neodvisnih enačb označujemo s $h_{ij} \in \{0,1\}$, pri $i = 1,...,m\ ;\ j = 1,...,n\ ;\ n \geq m$, torej:

$$\begin{array}{ccccccccccc} h_{11}\cdot x_1 & + & \cdots & + & h_{1j}\cdot x_j & + & \cdots & + & h_{1n}\cdot x_n & = & 0\ , \\ & & \vdots & & & & \vdots & & & & \\ h_{i1}\cdot x_1 & + & \cdots & + & h_{ij}\cdot x_j & + & \cdots & + & h_{in}\cdot x_n & = & 0\ , \\ & & \vdots & & & & \vdots & & & & \\ h_{m1}\cdot x_1 & + & \cdots & + & h_{mj}\cdot x_j & + & \cdots & + & h_{mn}\cdot x_n & = & 0\ . \end{array}$$

Sistem enačb lahko zapišemo v matrični obliki:

$$\mathbf{H}\cdot\mathbf{x}^T = \mathbf{0}^T\ ,$$

kjer je

$$\mathbf{H} = \begin{bmatrix} h_{11} & \cdots & h_{1j} & \cdots & h_{1n} \\ & \vdots & & \vdots & \\ h_{i1} & \cdots & h_{ij} & \cdots & h_{in} \\ & \vdots & & \vdots & \\ h_{m1} & \cdots & h_{mj} & \cdots & h_{mn} \end{bmatrix}$$

matrika za preverjanje sodosti,

$$\mathbf{x} = (x_1,...,x_n)$$

kodna zamenjava in

$$\mathbf{0}^T = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

m-razsežni ničelni vektor.

Iz linearne algebre je znano, da v primeru, ko je rang matrike $\mathbf{H}$ enak m, ima ta matrika poleg m linearno neodvisnih vrstic tudi m linearno neodvisnih stolpcev. V sistemu enačb izberemo $k = n - m$ x-ov poljubne vrednosti in določimo preostalih $m = n - k$ vrednosti x-ov z reševanjem sistema enačb. Z reševanjem dobimo natanko 2^k vseh možnih kodnih zamenjav, ki jih uporabimo za varen prenos po kanalu z motnjami.

Matriko za preverjanje sodosti sestavimo tako, da najprej določimo tri linearno neodvisne stolpce (na primer prve po vrsti), ostali pa so lahko linearno odvisni. Paziti moramo le, da v matriki ni enakih stolpcev ali stolpca samih

ničel. Sistem enačb je najlažje reševati, če so na začetku matrike različni stolpci s po eno samo enico.

Kod kanala na splošno označujemo s $\mathcal{K}(n,k)$. Linearni bločni kod $\mathcal{L}(n,k)$ je takšen kod kanala, pri katerem je vsaka kodna zamenjava linearna kombinacija drugih kodnih zamenjav iz množice K. Z matriko za preverjanje sodosti torej določamo linearne bločne kode $\mathcal{L}(n,k)$ [4].

***Naloga* 10.2**

Vzemimo, da želimo preko kanala prenašati 8 informacijskih blokov iz množice:

$$D = \{(000)\,,\,(001)\,,\,(010)\,,\,(011)\,,\,(100)\,,\,(101)\,,\,(110)\,,\,(111)\}\ .$$

Z idealnim pravilom odločanja želimo pravilno dekodirati še vse enkratne napake, torej velja $e = 1$. Vsaj koliko kontrolnih dvojiških znakov moramo še dodati informacijskim, da bi izpolnili to zahtevo. Določite še primerno matriko za preverjanje sodosti in vse kodne zamenjave.

Rešitev:

S preverjanjem Hammingovega pogoja najprej določimo minimalno potrebno število kontrolnih dvojiških znakov. Upoštevamo, da je $M = 8$ in $e = 1$. Hammingov pogoj je potem določen kot:

$$\frac{2^n}{\sum\limits_{i=0}^{e}\binom{n}{i}} \geq M \quad \Rightarrow \quad \frac{2^n}{\binom{n}{0}+\binom{n}{1}} \geq 8 \quad \Rightarrow \quad \frac{2^n}{1+n} \geq 8\ .$$

Preverimo izpolnjevanje pogoja za $m = 1, 2, ...$

m	$n = k + m$	Hammingov pogoj
1	4	$\frac{2^4}{1+4} = \frac{16}{5} = 3{,}2 \ngeq 8$
2	5	$\frac{2^5}{1+5} = \frac{32}{6} \approx 5{,}3 \ngeq 8$
3	6	$\frac{2^6}{1+6} = \frac{64}{7} \approx 9{,}14 \geq 8$

Glede na to, da mora veljati vsaj (pogoj je potreben, ni pa zadosten) $m = 3$ in $n = 6$ sestavimo matriko po osnovnih navodilih v

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}\ .$$

Sistem linearnih enačb lahko potem zapišemo kot:

$$\mathbf{H}\cdot\mathbf{x}^T=\mathbf{0}^T \;\Rightarrow\; \begin{array}{ccccccccccccc} x_1 & + & & & & & x_4 & + & x_5 & + & x_6 & = & 0 \\ & & x_2 & + & & & x_4 & + & x_5 & & & = & 0 \\ & & & & x_3 & + & x_4 & + & & & x_6 & = & 0 \end{array} .$$

Glede na to, da prve tri stolpce obravnavamo kot linearno neodvisne, bomo za x_4, x_5 in x_6 izbirali vse možne poljubne dvojiške vrednosti, x_1, x_2 in x_3 pa določili z enostavnim reševanjem zgornjih enačb nad GO(2), in sicer:

$$\begin{array}{ccccccccc} x_1 & = & x_4 & + & x_5 & + & x_6 & & \\ x_2 & = & x_4 & + & x_5 & & & & . \\ x_3 & = & x_4 & + & & & x_6 & & \end{array}$$

Množica kodnih zamenjav je torej:

				z_1	z_2	z_3
$\mathbf{x}_i$	x_1	x_2	x_3	x_4	x_5	x_6
$\mathbf{x}_1$	0	0	0	0	0	0
$\mathbf{x}_2$	1	0	1	0	0	1
$\mathbf{x}_3$	1	1	0	0	1	0
$\mathbf{x}_4$	0	1	1	0	1	1
$\mathbf{x}_5$	1	1	1	1	0	0
$\mathbf{x}_6$	0	1	0	1	0	1
$\mathbf{x}_7$	0	0	1	1	1	0
$\mathbf{x}_8$	1	0	0	1	1	1

Ugotovimo, da je minimalna razdalja med dvema kodnima zamenjavama enaka 3, torej je tako določen kod dejansko sposoben odkrivati in popravljati še vse enkratne napake.

10.3 Dekodiranje linearnih bločnih kodov

Dekodiranje linearnih bločnih kodov izvajamo z računanjem sindroma po spodnji matrični enačbi:

$$\mathbf{H}\cdot\mathbf{y}^T=\mathbf{s}^T ,$$

kjer je $\mathbf{s}$ sindrom prejetega vektorja $\mathbf{y}$. Sindrom je enak $\mathbf{0}^T$, če je $\mathbf{y}$ enak enemu od $\mathbf{x}$ množice veljavnih kodnih zamenjav K. Napake odkrivamo tako, da ugotavljamo, ali je sindrom $\mathbf{s}^T$ različen od $\mathbf{0}^T$ [4].

Napake pa lahko tudi popravljamo, ker je sindrom odvisen le od vektorja napake $\mathbf{e} = \mathbf{y} - \mathbf{x}$ po spodnji enačbi:

$$\mathbf{H} \cdot \mathbf{y}^T \;=\; \mathbf{H} \cdot (\mathbf{x} + \mathbf{e})^T \;=\; \mathbf{H} \cdot \mathbf{x}^T + \mathbf{H} \cdot \mathbf{e}^T \;=\; \mathbf{H} \cdot \mathbf{e}^T \;=\; \mathbf{s}^T \;,$$

kjer smo upoštevali, da velja $\mathbf{H} \cdot \mathbf{x}^T = \mathbf{0}^T$.

Dekodiranje izvajamo tako, da za vsak možen neničelni sindrom $\mathbf{s}_i^T$ ugotovimo pripadajoči vektor napake $\mathbf{e}_i$.

Pri odkrivanju še vseh enkratnih napak je pri enkratni napaki sindrom eden od stolpcev matrike za preverjanje sodosti.

***Naloga* 10.3**

Linearni bločni kod ima naslednjo matriko za preverjanje sodosti:

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \;.$$

a) Koliko različnih informacijskih blokov lahko kodiramo s tem kodom?
b) Vse kolikokratne napake bi kod lahko bil sposoben popravljati?
c) Ali je (1111111) kodna zamenjava tega koda?
d) Na katerem znaku je prišlo do napake pri sprejetem nizu (1011100)?

Rešitev:

Iz dimenzij matrike je očitno, da je $m = 3$ in $n = 7$. Torej je $k = n - m = 4$. Število različnih informacijskih blokov M je torej lahko $M = 2^k = 16$.

Preverimo pri katerem e_{max} je še izpolnjen potreben Hammingov pogoj

$$e_{max} = \arg\max_{e} \left\{ \frac{2^n}{\sum_{i=0}^{e} \binom{n}{i}} \geq M \right\} \;.$$

Preverimo torej pogoj za $e = 0, 1, 2, ...$

e	Hammingov pogoj
0	$\frac{2^7}{\binom{7}{0}} = \frac{128}{1} = 128 \geq 16$
1	$\frac{2^7}{\binom{7}{0}+\binom{7}{1}} = \frac{128}{1+7} = 16 \geq 16$
2	$\frac{2^7}{\binom{7}{0}+\binom{7}{1}+\binom{7}{2}} = \frac{128}{1+7+21} \approx 4.41 \not\geq 16$

Potreben pogoj za to, da bi kod popravljal še vse enkratne napake je izpolnjen, torej je možno, da kod odkriva in popravlja še vse enkratne napake. Iz matrike je tudi jasno, da je njen rang enak m in da je torej mogoče določiti kodne zamenjave, ki bodo omogočale tudi popravljanje vseh enkratnih napak.

Preverimo, če je (1111111) veljavna kodna zamenjava.

$$\mathbf{H} \cdot \mathbf{x}^T = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Vidimo, da je vektor (1111111) veljavna kodna zamenjava, saj je njegov sindrom enak **0**.

Preverimo še kakšen sindrom ima niz (1011100)

$$\mathbf{H} \cdot \mathbf{y}^T = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} .$$

Sindrom je očitno različen od **0**. Ker predpostavljamo odkrivanje enkratnih napak, je potrebno poiskati sindrom v stolpcih matrike **H**. Napaka je na tretjem dvojiškem znaku, torej $\mathbf{e} = (0010000)$, kar je

razvidno tudi iz tega, da je

$$\mathbf{H} \cdot \mathbf{e}^T = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \mathbf{s}^T \ .$$

10.4 Lastnosti linearnih bločnih kodov

Vsaka matrika za preverjanje sodosti $\mathbf{H}'$, ki jo dobimo iz matrike $\mathbf{H}$ s prerazporeditvijo stolpcev ali vrstic je matrika koda, ki je enakovreden kodu, ki ga določa izvirna matrika $\mathbf{H}$.

S prerazporeditvijo stolpcev ali vrstic dobimo standardno matriko oblike [4]:

$$\mathbf{H}' = [\mathbf{I}_m \mid \mathbf{B}_{mk}] \ .$$

Generatorska matrika $\mathbf{G}$ je matrika razsežnosti $k \times n$ z rangom k. Njene vrstice tvori k linearno neodvisnih temeljnih n-razsežnih kodnih zamenjav. Vse kodne zamenjave koda $\mathcal{L}(n, k)$ dobimo, če vse k-razsežne informacijske bloke $\mathbf{z} \in D$ množimo na desni z $\mathbf{G}$, torej [4]

$$\mathbf{x}_i = \mathbf{z}_i \cdot \mathbf{G} \quad \text{za vsak} \quad \mathbf{z}_i \in D \ .$$

V primeru, ko je matrika $\mathbf{H}$ zapisana v standardni obliki $\mathbf{H} = [\mathbf{I}_m \mid \mathbf{B}_{mk}]$, velja, da je standardna oblika generatorske matrike $\mathbf{G}$ enaka

$$\mathbf{G} = \left[\mathbf{B}_{mk}^T \mid \mathbf{I}_k\right] \ .$$

Na ta način določimo generatorsko matriko, s katero tvorimo vse kodne zamenjave.

***Naloga* 10.4**

Vzemimo, da želimo popravljati še vse enkratne napake pri prenašanju štirih različnih dvojiških informacijskih blokov. Določite standardno

matriko za preverjanje sodosti ustreznega linearnega bločnega koda. Določite še generatorsko matriko in z njo tvorite vse veljavne kodne zamenjave.

Rešitev:

Glede na to, da je $M = 4$, je najmanjši možni $k = 2$. S preverjanjem Hammingovega pogoja najprej določimo minimalno potrebno število kontrolnih dvojiških znakov in upoštevamo, da je $M = 4$ in $e = 1$. Hammingov pogoj je potem določen kot:

$$\frac{2^n}{\sum_{i=0}^{e}\binom{n}{i}} \geq M \quad \Rightarrow \quad \frac{2^n}{\binom{n}{0}+\binom{n}{1}} \geq 4 \quad \Rightarrow \quad \frac{2^n}{1+n} \geq 4$$

Preverimo izpolnjevanje pogoja pri $k = 2$ za $m = 1, 2, ...$

m	$n = k + m$	Hammingov pogoj
1	3	$\frac{2^3}{1+3} = \frac{8}{4} = 2 \ngeq 4$
2	4	$\frac{2^4}{1+4} = \frac{16}{5} = 3{,}2 \ngeq 4$
3	5	$\frac{2^5}{1+5} = \frac{32}{6} \approx 5{,}3 \geq 4$

Iz izračunov je razvidno, da mora biti $m = 3$ in $n = 5$. Sestavimo standardno matriko za preverjanje sodosti

$$\mathbf{H} = [\mathbf{I}_m \mid \mathbf{B}_{mk}] = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right].$$

Stolpci v delu, ki pripada matriki $\mathbf{B}_{mk}$, bi lahko bili tudi drugačni. Pomembno je le, da vsebujejo več kot eno enico in da so med seboj različni.

Standardna generatorska matrika je potem določena kot:

$$\mathbf{G} = \left[\mathbf{B}_{mk}^T \mid \mathbf{I}_k\right] = \left[\begin{array}{ccc|cc} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{array}\right].$$

Za informacijske bloke $\mathbf{z}_i$ iz množice:

$$D = \{(00), (01), (10), (11)\}$$

tvorimo kodne zamenjave po izrazu

$$\mathbf{x}_i = \mathbf{z}_i \cdot \mathbf{G} \quad \text{za vsak} \quad \mathbf{z}_i \in D \ .$$

Za informacijski blok (01) je, denimo, kodna zamenjava določena kot

$$(01) \cdot \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \ = \ (11001) \ ,$$

kar je kar spodnja vrstica matrike $\mathbf{G}$. S podobnimi izračuni dobimo še vse preostale kodne zamenjave

$$K = \{(00000)\,,\ (11001)\,,\ (01110)\,,\ (10111)\} \ .$$

10.5 Hammingov kod

Hammingov kod $\mathcal{H}(n,k)$ je linearni bločni kod z dolžino kodnih zamenjav $n = 2^m - 1$, pri $m \geq 2$. Matrika za preverjanje sodosti je sestavljena tako, da so po stolpcih dvojiško zapisana števila $1, 2, ..., 2^m - 1$. Hammingov kod je vedno sposoben odpravljati še vse enkratne napake, zato je dekodiranje zelo preprosto. Morebitni neničelni sindrom poiščemo med stolpci matrike in pri odločitvi predpostavimo, da se je na mestu tega stolpca tudi zgodila napaka [4].

Dimenzija matrike $\mathbf{H}$ je torej vnaprej določena z zvezo $n = 2^m - 1$. Spodaj je podanih nekaj možnih dimenzij, ki navajajo tudi predvidene dolžine informacijskih blokov k in s tem tudi število možnih različnih informacijskih blokov M.

m	n	k	M
1	1	0	1
2	3	1	2
3	7	4	16
4	15	11	2048

***Naloga* 10.5**

Določite matriko za preverjanje sodosti Hammingovega koda, s katerim lahko kodiramo 10 različnih informacijskih blokov! Določite še vsaj eno veljavno kodno zamenjavo tega koda!

Rešitev:

Upoštevamo zvezi $n = 2^m - 1$ in $k = n - m$, od koder sledi

$$n = 2^{n-k} - 1 \ .$$

Ker velja $M \leq 2^k$, mora biti pri $M = 10$ vrednost k vsaj 4, ker je $10 \leq 2^4$. Preverimo, pri katerih (najmanjših) vrednostih n in k velja zveza $n = 2^{n-k} - 1$, pri čemer zahtevamo, da je $k \geq 4$.

Uporabimo lahko kar izračune v zgornji tabeli, iz katerih ugotovimo, da je najmanjši še veljavni k prav 4, torej je rešitev pri $n = 7, m = 3$ in $k = 4$.

Matrika za preverjanje sodosti je sestavljena tako, da so po stolpcih le dvojiško zapisana števila od 1 do 7, torej

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \ .$$

Sistem linearnih enačb lahko potem zapišemo v matrični obliki kot:

$$\mathbf{H} \cdot \mathbf{x}^T = \mathbf{0}$$

oziroma eksplicitno v obliki sistema enačb

$$\begin{array}{ccccccccccccc} & & & & & x_4 & + & x_5 & + & x_6 & + & x_7 & = & 0 \\ & x_2 & + & x_3 & & & + & & & x_6 & + & x_7 & = & 0 \\ x_1 & & & x_3 & + & & + & x_5 & & & + & x_7 & = & 0 \end{array} \ .$$

Za tri linearno neodvisne stolpce vzamemo tiste pri x_1, x_2 in x_4. Za x_3, x_5, x_6 in x_7 pa vzamemo poljubne dvojiške vrednosti in nato x_1, x_2 in x_4 določimo z enostavnim reševanjem zgornjih enačb nad GO(2), in sicer:

$$\begin{array}{ccccccccc} x_4 & = & x_5 & + & x_6 & + & x_7 & & \\ x_2 & = & x_3 & + & x_6 & + & x_7 & . & \\ x_1 & = & x_3 & + & x_5 & + & x_7 & & \end{array}$$

pri $\mathbf{z} = (0000)$ bi vzeli $x_3 = 0$, $x_5 = 0$, $x_6 = 0$ in $x_7 = 0$ in dobili $x_1 = 0$, $x_2 = 0$ in $x_4 = 0$, torej je veljavna kodna zamenjava kar $\mathbf{x} = (000000)$.

A Zgledi pisnega izpita

A.1 Zgled 1

[64207] Informacija in kodi - pisni izpit Nekdaj

Največ točk za izpit: 30
Najmanj točk za uspešno opravljen izpit: 15
Čas pisanja: 60 minut

Naloga 1 (7 točk)

Vzemimo, da je statistična analiza krvnih skupin pri zahodnih in vzhodnih Evropejcih podala naslednje rezultate:

krvna skupina	zahodni Evropejci	vzhodni Evropejci
0	60%	50%
A	30%	30%
B	5%	15%
AB	5%	5%

Vzemimo, da je v Evropski uniji po grobi oceni 65 % prebivalcev po poreklu zahodnih Evropejcev in 35 % vzhodnih Evropejcev. Izračunajte, koliko informacije (v bitih) lahko v povprečju pridobimo o poreklu Evropejca zgolj iz poznavanja njegove krvne skupine.

Naloga 2 (8 točk)

Diskretni stacionarni informacijski vir brez spomina oddaja štiri različne znake. Negospodarni enakomerni dvojiški kod vira je podan v spodnji tabeli skupaj s porazdelitvenim zakonom.

$$V = \begin{pmatrix} x_1 \to 00 & x_2 \to 01 & x_3 \to 10 & x_3 \to 11 \\ p_1 = 0{,}3 & p_2 = 0{,}5 & p_3 = 0{,}1 & p_4 = 0{,}1 \end{pmatrix}.$$

Določite Huffmanov gospodarni dvojiški kod vira. Primerjajte uspešnost, η, Huffmanovega koda in enakomernega koda, ki je dan v zgornji tabeli.

Naloga 3 (7 točk)

Določite kapaciteto diskretnih komunikacijskih kanalov, ki jih določajo spodnje štiri matrike pogojnih verjetnosti:

$$\mathbf{P}_{k_1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad \mathbf{P}_{k_2} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad ,$$

$$\mathbf{P}_{k_3} = \begin{bmatrix} 0 & 0{,}2 & 0{,}8 \\ 1 & 0 & 0 \end{bmatrix} \quad , \quad \mathbf{P}_{k_4} = \begin{bmatrix} 0{,}4 & 0{,}1 & 0{,}5 \\ 0{,}5 & 0{,}4 & 0{,}1 \\ 0{,}1 & 0{,}5 & 0{,}4 \end{bmatrix} \quad .$$

Naloga 4 (8 točk)

Naj bo linearni bločni kod $\mathcal{L}(5{,}2)$ definiran z naslednji matriko za preverjanje sodosti:

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad .$$

a) Zapišite matriko $\mathbf{H}$ v standardni obliki.
b) Določite standardno obliko generatorske matrike.
c) Določite množico informacijskih vektorjev in kodnih zamenjav.
d) Demonstrirajte odkrivanje napak z danim kodom.

A.2 Zgled 2

[64207] Informacija in kodi - pisni izpit Nekoč

Največ točk za izpit: 30
Najmanj točk za uspešno opravljen izpit: 15
Čas pisanja: 60 minut

Naloga 1 (8 točk)

V diskontni prodajalni so na polici zapisljivi diski DVD-R slabše kakovosti v ovitkih, ki jih je izdelal preprodajalec. Kljub enakim ovitkom so v njih diski z dvema različnima barvama nanosa (srebrna in zelena barva) dveh različnih izdelovalcev. Od 200 diskov na polici jih je 50 od prvega in 150 od drugega izdelovalca. Vsi diski prvega izdelovalca imajo srebrni nanos. Od drugega izdelovalca je diskov s srebrnim nanosom 50, vsi preostali diski drugega izdelovalca pa imajo zeleni nanos.

V prodajalno pride kupec, ki naključno izbere en disk. Predpostavimo, da kupec navedena dejstva pozna.

a) Izračunajte, kolikšna je verjetnost, da kupec izbere disk s srebrnim nanosom.

b) Predpostavite, da kupec izbere disk s srebrnim nanosom. Izračunajte, kolikšna je verjetnost, da je disk izdelal prvi izdelovalec.

c) Izračunajte, koliko informacije v bitih bi v povprečju kupec pridobil o tem, kateri od obeh izdelovalcev je izdelal disk, po tem, ko ugotovi barvo nanosa (in pozna zgoraj navedena dejstva).

Naloga 2 (7 točk)

Dvojiški stacionaren homogen Markovov vir je oddal dvojiški niz

$$110011100000000101111000.$$

Statistično ocenite, s kakšno verjetnostjo ta vir enko spremeni v ničlo in ničlo v enko. Iz statističnih ocen teh pogojnih verjetnosti ocenite še:

a) Stacionarno porazdelitev podanega vira.

b) Entropijo podanega vira.

Naloga 3 (7 točk)

Vir z abecedo štirih znakov, ki smo jih enakomerno dvojiško kodirali, oddaja znake z verjetnostmi po spodnjem porazdelitvenem zakonu.

$$K \sim \begin{pmatrix} \mathbf{x}_1 = \texttt{(00000)} & \mathbf{x}_2 = \texttt{(11001)} & \mathbf{x}_3 = \texttt{(11110)} & \mathbf{x}_4 = \texttt{(00111)} \\ P(\mathbf{x}_1) = 0{,}4 & P(\mathbf{x}_2) = 0{,}1 & P(\mathbf{x}_3) = 0{,}3 & P(\mathbf{x}_4) = 0{,}2 \end{pmatrix}.$$

Kodne zamenjave prenašamo po dvojiškem simetričnem kanalu z motnjami z verjetnostjo napake na dvojiškem simbolu $p_n = 0{,}25$: Denimo, da dekodirnik na izhodu kanala sprejme dvojiški niz $\mathbf{y} = \texttt{(11000)}$.

a) Katero kodno zamenjavo $\hat{\mathbf{x}}_i \in K$ bi izbral dekodirnik z *idealno funkcijo*?

b) Katero kodno zamenjavo $\hat{\mathbf{x}}_i \in K$ bi izbral dekodirnik z *idealnim opazovalcem*?

Naloga 4 (8 točk)

Vir z abecedo štirih znakov oddaja tri znake na sekundo. Znake dvojiško kodiramo za prenos po dvojiškem kanalu z motnjami, ki ima kapaciteto 15 bitov na sekundo. Preverite, če je mogoče pri tem viru in kanalu določiti linearni bločni dvojiški kod, ki bo sposoben odpravljati še vse enkratne napake. Če je mogoče, določite primer takšnega koda, torej, določite primerno matriko za preverjanje sodosti in vse štiri kodne zamenjave.

Literatura

[1] D. P. Bertsekas, J. N. Tsitsiklis: *Introduction to Probability*, Athena Scientific, ISBN-13: 978-1886529403, 2002, 430 strani.

[2] R. M. Gray: *Entropy and Information Theory*, Springer, ISBN-13: 978-1441979698, 2011, 2. izdaja, 409 strani.

[3] C. M. Grinstead, J. L. Snell: *Introduction to Probability*, American Mathematical Society, ISBN-13: 978-0821807491, 1997, 2 popravljena izdaja, 510 strani.

[4] N. Pavešić: *Informacija in kodi - druga, spremenjena in dopolnjena izdaja*, Založba FE in FRI, ISBN 978-961-243-145-7, 2010, 309 strani.