Základní pojmy [7,8,9,10]

SIGNÁL → fyzikální veličina, která nese informaci

ABECEDA → množina možných elementárních signálů

ZPRÁVA → vytvoříme-li z prvků abecedy nějakou sestavu symbolů, kterou použijeme k přenosu informace a bude-li tato sestava časově ohraničena, tvoří zprávu.

Každá zpráva má pak:

- ***Sémantický význam (obsah)*** – vyjadřuje významovou stránku zprávy (co říká, co je obsahem), jde o kvalitativní stránku zprávy,
- ***Pragmatický význam (obsah)*** – určuje významnost (důležitost, užitečnost, cennost) sdělení a prioritu pro příjemce
- ***Syntaktický význam (obsah)*** – určuje formu zprávy, týká vzájemného uspořádání znaků jako nositelů informace.

Příklad: Mějme zprávu „Soused umřel“. Vysvětlete jednotlivé významy zprávy.

Sémantický význam: je třeba zařídit pohřeb apod.
Pragmatický význam: jiná důležitost pro rodinu souseda než pro jiné
Syntaktický význam: jde o zprávu obsahující 12 znaků

Jednotka a zobrazení informace [3]

Jednotku informace nazýváme - BIT (**bi**nary digi**t** – dvojkové číslo)

je to informace, kterou nese zpráva o stavu systému – možné stavy pouze 0 a 1 se stejnou pravděpodobností.

BIT - má i význam pravdivostní hodnoty jednoho výroku *tzn. jeden BIT je elementární množství informace* např.: žárovka SVÍTÍ – NESVÍTÍ

Vyjádření informační hodnoty je velmi obtížné. Každý objekt může být v každém časovém okamžiku v jednom z mnoha možných stavů.

Informace o systému je tím větší, čím je pravděpodobnost výskytů jednotlivých jeho stavů menší. Informace je větší, obsahuje-li zpráva něco nového, nebo co nelze snadno uhodnout (je málo pravděpodobné).

Informační hodnota [4]

Množství informace obsažené ve zprávě X souvisí s pravděpodobností P(X) s jakou může příjemce uhodnout obsah zprávy X, neboli jaká je pravděpodobnost výskytu dané zprávy u příjemce před jejím přijetím.

Platí:

$$\mathrm{I(X)} = \mathrm{f}\left(\frac{1}{\mathrm{P(X)}}\right) = \log_2 \frac{1}{\mathrm{P(X)}} = -\log_2 \mathrm{P(X)}$$

Zpráva, která nese informaci o stavu systému, který je jistý nenese žádnou informaci; **P(X)=1 a množství informace je pak:**

$$\mathrm{I(X)} = -\log_2 1 = 0 \text{ bit}$$

Příklad:

Házíme kostkou. Jaké množství informace nese zpráva „padlo sudé číslo„?

Pravděpodobnost, že padne sudé číslo je P(X)=0,5**:** **$I(X) = -\log_2 P(X) = -\log_2 0{,}5 = 1$ bit**

Pokud se zpráva **X** skládá ze dvou nezávislých zpráv **A, B** pak celková informace, kterou získá příjemce je rovna:

$$I(X) = I(A) + I(B)$$

přičemž pravděpodobnost zprávy **X** dvou nezávislých jevů je dána:

$$P(X) = P(A)\;.\;P(B)$$

Po úpravě dostaneme:

$$I(X) = f\left(\frac{1}{P(A)}\right) + f\left(\frac{1}{P(B)}\right) = f\left(\frac{1}{P(A)\cdot P(B)}\right)$$

Příklad:

Zpráva X, která nás informuje o výsledku hodu dvěma kostkami nese informaci I(X). Zpráva o výsledku hodu první kostkou nese informaci I(A) a zpráva o výsledku hodu druhou kostkou informaci I(B). Hody jsou nezávislé, tudíž je celková informace o výsledku hodu obou kostek rovna součtu informací, které nesou zprávy o výsledku jednotlivých poloh kostek. Nechť výsledkem hodu na jedné kostce je 3 a na druhé 4. Jaké množství informace nese zpráva že padlo číslo 3 a zároveň číslo 4?

Pravděpodobnost toho, že současně padne 3 a 4 je: $\mathbf{P_{3,4} = P(X) = P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}}$

Množství informace ve zprávě s výsledkem na obou kostkách je pak dáno:

$$\mathbf{I(X) = -\log_2 (1/36) = 5.17\ bitů}$$

A protože pravděpodobnost že padne 3 na první kostce resp. pravděpodobnost že padne 4 na druhé kostce je 1/6 pak dle dříve uvedeného platí:

$$\mathbf{I(X) = I(A) + I(B) = -\log_2 \frac{1}{6} - \log_2 \frac{1}{6} = 2.585 + 2.585 = 5.17\ bit}$$

Tabulka největších informačních obsahů soustav

Soustava	Bity	Vyjádření
číslice ve dvojkové soustavě	1	$I(X) = -\log_2 0.5 = 1\ \text{bit}$
číslice v desítkové soustavě	3,32	$I(X) = -\log_2 0.1 = 3.32\ \text{bit}$
číslice v šestnáctkové soustavě	4	$I(X) = -\log_2 \frac{1}{16} = 4\ \text{bit}$
symbol anglické abecedy (27)	4,75	$I(X) = -\log_2 \frac{1}{27} = 4.75\ \text{bit}$
karta v balíku mariášových karet	5	
symbol české abecedy (42)	5,39	
krátkodobá lidská paměť	10^4	
průměrná kniha	10^7	
paměť RAM (PC)	10^8	
genetická informace	10^10	
lidská paměť	10^20	

Hartleyho formule [1,7,10]

definuje množství informace, které je nutné k charakterizaci jednoho prvku množiny E o N prvcích.

$$I\,(E_i) = \log_2 N$$

- Tento vzorec je možné zobecnit.
- Jestliže množina *M* obsahuje *m* symbolů o stejné pravděpodobnosti výskytu (tj. *p=1/m*), vyslání každého symbolu nese kvantum informace podle vzorce:

$$I = K \cdot \log_b m = -K \cdot \log_b \left(\frac{1}{m}\right) = -K \cdot \log_b p$$

kde ***K*** je konstanta odpovídající zvolenému základu logaritmů:

b=2 binární jednotka – ***bit***

b=10 desítková jednotka – ***dit (decit)***

b=e jednotka natural - ***nat***.

Příklad: Házíme mincí. Dva možné stavy: hlava, orel. Jaké množství informace nese zpráva o výsledku náhodného pokusu?

$$I(E_i) = \log_2 2 = 1\,\text{bit}$$

Shannonova formule [6,8,9]

pokud všechny stavy náhodného jevu nemají stejnou pravděpodobnost pak platí:

$$I = \sum_{k=1}^{n} p_k \cdot \log_2 \frac{1}{p_k}$$

pro $p_1=p_2=\ldots p_n=1/n$ dostáváme **Hartleyho formuli**

Příklad:

Česká abeceda obsahuje 42 znaků (včetně diakritiky). Zamyslete se nad množstvím informace jednoho znaku!!!

Za předpokladu rovnoměrného výskytu všech písmen je pravděpodobnost výskytu u všech znaků stejná: **P(X) = 1/42 = 0,0238095238**

Pro výpočet množství informace, kterou nese jeden znak abecedy lze použít Hartleho formuli popřípadě Shannonovu formuli: **$I(X) = \log_2 42$** popř. **$I(X)=-\log_2 0{,}238095238 = 5.39$ bitů**

Ve skutečnosti **neexistuje** zcela **rovnoměrný výskyt** všech znaků v české abecedě (ani v ostatních).

Příklad:

Zkoumejte množství informace jednoho znaku s uvažováním pravděpodobnostního výskytu znaků v daném jazyce.

Bez dlouhého zkoumání je nám zřejmé, že písmeno **e** se v českých textech vyskytuje častěji než **w**. Použijeme-li výsledky analýzy frekvence jednotlivých písmen české abecedy v dostatečně dlouhém textu, zjistíme, že informační hodnotu jednoho písmene musíme vypočítat podle následujícího schématu:

mezera	a	á …… z	z
0,166	0,054	0,022…0,019	0,009

Výsledná informační hodnota jednoho znaku je pak:

$$\mathbf{I = 0{,}166.\log_2 1/0{,}166 + 0.054.\log_2 1/0{,}054 + \;.... = 4{,}67\ bitů}$$

Míru informace však můžeme odvodil ještě jinou úvahou. Vyjdeme z ***Nyquistova-Küpfmullerova časového zákona*** vyjadřujícího vztah mezi šířkou pásma a dobou přenosu, který je jedním z prvních pokusů o kvantitativní vyjádření informace

$$B \,.\, T = \text{konst} \;\; [c]$$

Kde **B** - šířka pásma [Hz]
T - doba přenosu [s]

Vyjádříme-li tuto rovnici graficky, dostaneme soustavu obdélníků, jejichž vrcholy leží na hyperbole. Plochy obdélníků charakterizují do jisté míry obsah informace. Dosadíme-li do předchozí rovnice příslušné rozměry, je obsah informace vyjádřen v kmitech

$$B\,[c / s] \,.\, T\,[s] = \text{konst} \;\; [c]$$

Informační obsah zůstane zachován, zmenšíme-li např. m-krát dobu **T** a naopak m-krát zvětšíme šířku pásma **B**.
Časový zákon doplnil Hartley o další veličinu (třetí rozměr) a to o **amplitudu** resp. **počet amplitudových stupňů s** .

$$\mathbf{s = \log_2 Z = lb\ Z}$$

kde Z určuje počet stavů jednoho místa

Míra informace je pak dána rovnicí:

$$\mathbf{I = 2 \cdot B \cdot T \cdot s = 2 \cdot B \cdot T \cdot lb\ Z \quad [bit]}$$

Výraz násobíme dvěma proto, abychom mohli respektovat každou půl vlnu zvlášť a abychom dostali výsledek v **bitech** a nikoli v kmitech. Informace dostává takto prostorové vyjádření v souřadnicích - čas, šířka pásma, amplituda. Změnou kteréhokoli z těchto parametrů měníme informační obsah.

Příklad:

Určete informační obsah dálnopisné značky kde B = 25 Hz a T = 100 ms (na jedno písmeno).

$$\mathbf{I = 2 \cdot B \cdot T \cdot lb\ Z = 2 \cdot 25 \cdot 100 \cdot 10^{-3} \cdot lb\ 2 \ = 5\ bitů}$$

Množství informace prošlé za jednotku času se nazývá **informační tok**.

Platí:

$$\mathbf{T_{inf} = \frac{I}{t} = 2 \cdot B \cdot lb\ Z \quad [bit/s]}$$

Příklad:

Přenos řeči v běžném telefonním styku vyžaduje šířku pásma B = 3000 Hz. Nahradíme-li 32 písmen abecedy odpovídajícím počtem amplitudových stupňů, je teoreticky informační tok dán:

$$\mathbf{T_{inf} = 2 \cdot B \cdot lb\ Z = 2 \cdot 3000 \cdot lb\ 32 \ = 30\,000\ bit/s}$$

ENTROPIE (míra neurčitosti) [3,5,10]

děje v přírodě dělíme na :

- **deterministické** – neurčitost rovna nule
- **náhodné** – jistá míra neurčitosti – velikost neurčitosti závisí na celkovém počtu možných výsledků a na pravděpodobnosti výskytu jednotlivých výsledků

Míra neurčitosti *i*-tého výsledku $\boldsymbol{H_i}$ pak bude funkcí pravděpodobnosti $\boldsymbol{p_i}$

$$H_i = f(p_i)$$

platí: $f(1) = 0$

$f(p_i)$ je monotónní klesající funkcí pravděpodobnosti $\mathbf{p_i}$, tj. $\mathbf{f(p_i) < f(p_j)}$, **pro** $\mathbf{p_i > p_j}$

Jediná funkce, která těmto podmínkám vyhovuje, je funkce logaritmického tvaru:

$$H_i = c \, . \log p_i$$

kde **c** je libovolná konstanta (když c = -1 pak H kladné)

Entropie je tedy míra neurčitosti v nějaké zprávě *X* o daném systému.

- Tato míra neurčitosti se po příjmu zprávy odstraňuje a tím se vyjádří míra získané informace.
- Při růstu informace klesá entropie a naopak.

Entropii zprávy X je pak možné vyjádřit podle Shannonova vzorce:

$$\mathbf{H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i}$$

Entropii náhodné veličiny můžeme chápat i jako průměrné množství informace, které získáme, vykonáme-li několik nezávislých opakování náhodné veličiny.

Je to vlastně střední hodnota náhodné veličiny.

Příklad:

Oblíbeným demonstračním příkladem je házení hrací kostkou se šesti stěnami. Může nastat jeden ze šesti stavů, tj. hodíme jedničku, nebo dvojku, nebo ... nebo šestku. Všechny možnosti mají stejnou pravděpodobnost výskytu = 1/6. Jaká je entropie výsledku hodu kostkou?

Ze zadání vyplývá konkrétní konečné schéma:
$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix}$$

pak ENTROPIE:

$$\mathbf{H} = -\left(1/6 \cdot \log_2 1/6 + 1/6 \cdot \log_2 1/6 + \ldots + 1/6 \cdot \log_2 1/6\right) =$$
$$= -\log_2 1/6 = +\log_2 6 = 2.5849$$

Vlastnosti entropie [7]

Entropie náhodné veličiny **X**, která nabývá **n** hodnot $\mathbf{x_i}$ s příslušnými pravděpodobnostmi $\mathbf{p_i}$, má tyto základní vlastnosti:

- Entropie je spojitá a nezáporná funkce – všechny náhodné procesy mají nezápornou neurčitost. $\mathbf{H(X) \geq 0}$
- Entropie je rovna nule tehdy, když pravděpodobnost výskytu některého znaku x_i je :

  $$p(x_i) = 1 \Rightarrow p(x_j) = 0 \ \text{ pro všechny } i \neq j$$

- Není jednoznačnou funkcí svých argumentů.
- Jsou-li pravděpodobnosti výskytu jednotlivých stavů stejné má zpráva maximální entropii.

  $$H(X) \leq H(1/n, 1/n, \ldots .1/n) = \log_2 n = \max H(X)$$

Často se vyjadřuje ***průměrná entropie*** systému jako míra průměrné neurčitosti jednoho stavu systému.

Může-li systém nabývat **<u>n</u>** možných stavů s pravděpodobnostmi $\mathbf{p_1, p_2, \ldots, p_n}$ pak průměrná entropie je rovna:

$$\overline{H(X)} = H_{stř} = \frac{p_1 H_1 + p_2 H_2 + \ldots + p_n H_n}{p_1 + p_2 + \ldots + p_n} = \frac{\sum_{i=1}^{n} p_i H_i}{\sum_{i=1}^{n} p_i}$$

$$= \sum_{i=1}^{n} p_i H_i = -\sum_{i=1}^{n} p_i \log_2 p_i$$

pro $\mathbf{p_1 = p_2 = \ldots = p_n = 1/n}$ pak platí: $\mathbf{H(X) = -\log_2 1/n = \log_2 n = \max H(X)}$

Je-li skutečná entropie menší než maximální, znamená to, že zdroj se plně nevyužívá
– redundance (nadbytečnost) zdroje

$$R = 1 - \frac{H}{H_{max}} = 1 - \mu \qquad \mu = \frac{H}{\log_2 n}$$

μ - relativní entropie, nebo-li účinnost zdroje

Mějme nějaký systém **X**. Než dostaneme zprávu o jeho stavu je jeho entropie **H_X**. Po obdržení přesné zprávy o stavu systému **X** je jeho entropie **$H'_x = 0$**.

Informace **I_X** o systému **X**, kterou jsme získali, je:

$$I_X = H_X - H'_X = H_X - 0 = H_X = -\sum_{i=1}^{n} p_i \log_2 p_i$$

Množství informace, které dostaneme ve zprávě sdělující nám přesný stav nějakého systému ***je rovno entropii tohoto systému***.

Příklad:

Znázorněte graficky závislost entropie náhodné veličiny, která nabývá dvě možné hodnoty $A=\{a_1, a_2\}$ na pravděpodobnosti p. (přičemž $P(a_1)=p$, $P(a_2)=1-p$). Vypočtěte hodnotu entropie pro p=0.3.

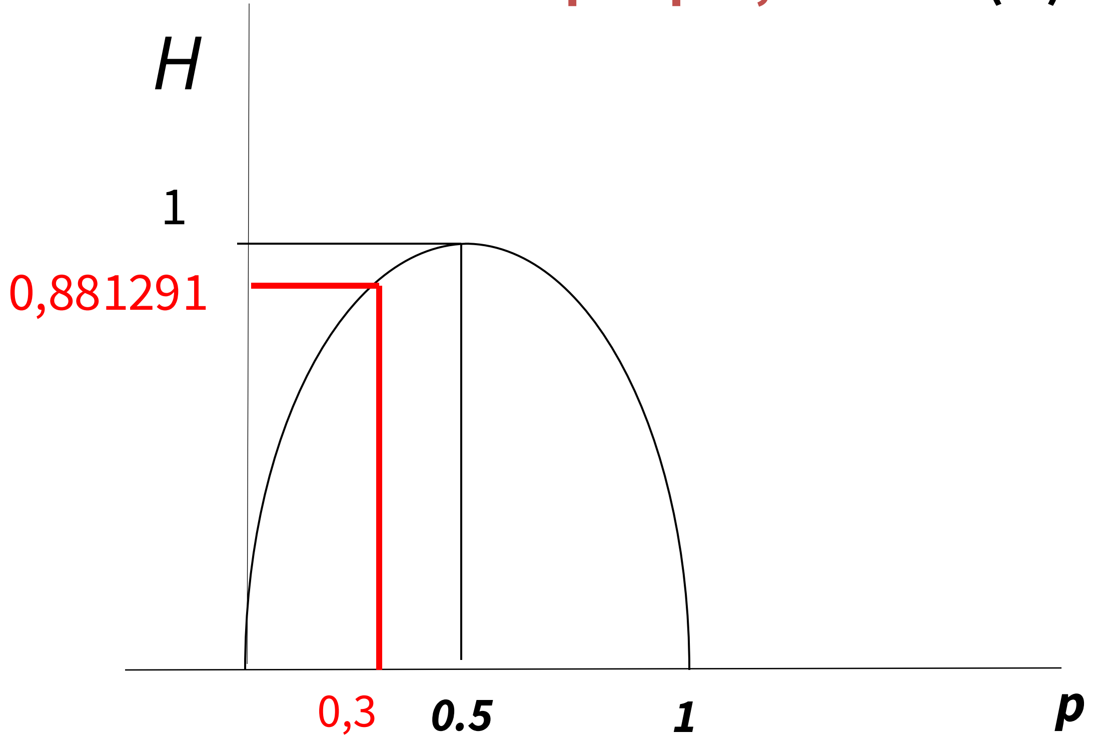
Pro entropii takového systému platí:

pro p=0,3

$$\mathbf{H(A) = -p \log_2 p - (1-p) \log_2 (1-p)}$$

$$\mathbf{H(A)} = -0{,}3 \log_2 0{,}3 - (1-0{,}3) \log_2 (1-0{,}3) =$$

$$0{,}52109 + 0{,}360201 = \mathbf{0{,}881291}$$



Grafické znázornění závislosti entropie dané náhodné veličiny na pravděpodobnosti $p \in \langle 0, 1 \rangle$.

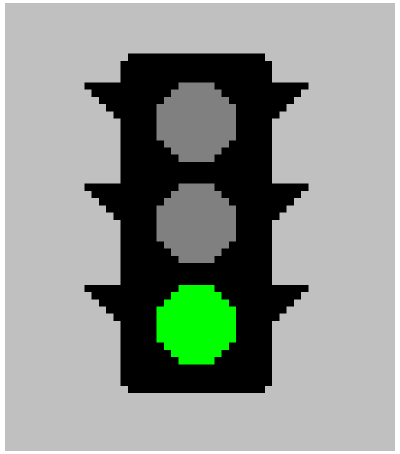
Červeně je vynesen výsledek příkladu:

Příklad:

Světelný semafor vysílá červený, zelený a žlutý signál, přičemž červený a zelený signál svítí vždy 40s, kdežto žlutý signál, který vždy odděluje zbývající 2 signály , svítí jen 20s. Určete entropii H náhodné veličiny X výskytu barvy signálu.

Pravděpodobnost uvedených světelných signálů, uvádí následující tabulka:

X	Č	Ž	Z	$\sum$
P_k	0,4	0,2	0,4	1



Řešení:

$$H(X) = -\,0{,}4 \text{ lb } 0{,}4 - 0{,}2 \text{ lb } 0{,}2 - 0{,}4 \text{ lb } 0{,}4 =$$

$$= 0{,}5288 + 0{,}4644 + 0{,}5288 = 1{,}522 \text{(bit)}$$

Příklad:

Máme dvě urny, z nichž každá obsahuje 20 koulí. V prvním osudí je 10 bílých, 5 červených a 5 žlutých koulí. V druhém osudí je 8 bílých, 8 červených a 4 žluté koule. Z každé urny vytáhneme jednu kouli. O kterém z těchto dvou pokusů lze tvrdit. Že jeho výsledek je méně neurčitý?

	Barva vytažené koule			
	bílá	červená	žlutá	Σ
Pravděpodobnost vytažení dané koule při 1. pokusu	0,5	0,25	0,25	1
Pravděpodobnost vytažení dané koule při 2. pokusu	0,4	0,4	0,2	1

Pro entropie jednotlivých pokusů dostaneme:

$$H_1 = -0{,}5 \text{ lb } 0{,}5 - 0{,}25 \text{ lb } 0{,}25 - 0{,}25 \text{ lb } 0{,}25 = 0{,}50 + 0{,}5 + 0{,}5 = \underline{\mathit{1{,}50\ bit}}$$

$$H_2 = -0{,}4 \text{ lb } 0{,}4 - 0{,}4 \text{ lb } 0{,}4 - 0{,}2 \text{ lb } 0{,}2 = 0{,}5288 + 0{,}5288 + 0{,}4644 = \underline{\mathit{1{,}522\ bit}}$$

Výsledek 1.pokusu je tedy méně neurčitý.

Příklad:

Podle dlouholetých zkušeností bylo v místě N zjištěno, že pravděpodobnost výskytu deště 15. června je p_1=0,4; kdežto pravděpodobnost, že nebude uvedeného dne pršet, je p_2=0,6. Podobně v témže místě N pravděpodobnost výskytu srážek (tj. deště nebo sněžení 15. listopadu je q_1=0,80, kdežto pravděpodobnost, že 15. listopad bude beze srážek, je q_2 = 0,20. Zajímá-li nás při počasí jen výskyt srážek, máme určit, který z obou dnů má určitější počasí.

Řešení:

Pro 15. červen je entropie:

$$H_1 = -0{,}4 \text{ lb } 0{,}4 - 0{,}6 \text{ lb } 0{,}6 = \underline{0{,}971\ \textit{bit}}$$

Pro 15. listopad je entropie:

$$H_2 = -0{,}8 \text{ lb } 0{,}8 - 0{,}2 \text{ lb } 0{,}2 = \underline{0{,}7219\ \textit{bit}}$$

Počasí 15. listopadu je v daném případě méně neurčité než počasí 15. června neboť $H_2 < H_1$.

Příklad:

Mějme abecedu A, B, C, D. Text sestavený z těchto abeced má 800 písmen. Předpokládejme, že A se vyskytuje 400 krát, B – 200 krát, C a D – 100 krát. Jaká je entropie takové abecedy? Jaká je redundance?

Pravděpodobnost výskytu jednotlivých písmen je dána takto:

P(A)=1/2, P(B)=1/4, P(C)=P(D)=1/8.

Entropie abecedy je pak dána: $$\mathbf{H = -\frac{1}{2}\,lb\,\frac{1}{2} - \frac{1}{4}\,lb\,\frac{1}{4} - \frac{1}{8}\,lb\,\frac{1}{8} - \frac{1}{8}\,lb\,\frac{1}{8} = \frac{7}{4}}$$

Pro výpočet redundance je zapotřebí určit maximální entropii dané abecedy. Abeceda bude mít maximální entropii tehdy, budou-li pravděpodobnosti výskytu jednotlivých stavů stejné. Tzn. **P(A)=P(B)=P(C)=P(D)=1/4**

Maximální entropie abecedy je pak dána: $$\mathbf{H_{max} = -\frac{1}{4}\,lb\,\frac{1}{4} - \frac{1}{4}\,lb\,\frac{1}{4} - \frac{1}{4}\,lb\,\frac{1}{4} - \frac{1}{4}\,lb\,\frac{1}{4} = 2}$$

Redundance je pak dána:

$$\mathbf{R = 1 - \frac{H}{H_{MAX}} = 1 - \frac{\frac{7}{4}}{2} = 1 - \frac{7}{8} = \frac{1}{8}}$$

V praxi často potřebujeme určit entropii systému, který je tvořen z několika podsystémů. Např. dva podsystémy **X** a **Y**, které mohou nabývat stavy: ***x_1, ... ,x_n*** resp. ***y_1, ...y_m***, sloučené do jednoho systému **XY**, který teď může nabývat stavy **(x_i , y_j)**. Počet stavů tohoto systému je ***n x m.***

Pravděpodobnost daného stavu je $$\mathbf{p_{ij}} = \mathbf{P}\left[\left(\mathbf{X} = \mathbf{x_i}\right)\left(\mathbf{Y} = \mathbf{y_j}\right)\right]$$

Entropie složeného systému XY je pak dána:

$$\mathbf{H_{XY}} = -\sum_{i=1}^{n}\sum_{j=1}^{m}\mathbf{p_{ij}} \cdot \log_2 \mathbf{p_{ij}}$$

Jsou-li podsystémy X, Y vzájemně nezávislé, platí: $\mathbf{H_{XY} = H_X + H_Y}$

Jsou-li podsystémy X, Y vzájemně závislé, platí: $\mathbf{H_{XY} = H_X + H_{Y|X} = H_Y + H_{X|Y}}$