Od parního stroje k teorii informace

Mezi základními pojmy jsme si mj. zhruba vysvětlili, co je to informace. Nyní se budeme věnovat otázce, jak se informace měří a proč. Budeme také pátrat po tom, jestli množství informace nějak souvisí s množstvím energie nebo hmoty – odpověď bude velice nejasná. Zvídavý čtenář se dozví víc ve výpravě do fyziky. Teprve tato výprava přinese odpovědi na některé dosud nezodpovězené otázky o základních pojmech, o informatickém pojetí světa a o souvislostech mezi informací a energií nebo hmotou.

Až dosud nás zajímalo, **co** je to informace. Na to nyní navážeme jinou otázkou: Jak informaci změříme? Jak změříme poznatky a vědomosti? Jak změříme podobnost nebo rozdílnost všeho toho, co se nám jeví? Jak změříme množství informace, které nám přináší jev?

Touto otázkou se zabývá *teorie informace* – věda, která je jedním ze základů informatiky.

V jakých jednotkách měříme informaci?

Je zřejmé, že čím víc různých věcí, hodnot nebo pojmů dokážeme rozlišit, tím víc informace máme. Tato jednoduchá představa nám sice stále ještě příliš nepomůže s měřením informace v jevech, nicméně jednu důležitou věc si můžeme ujasnit: kolik nejméně informace nám může jev přinést? Samozřejmě: jevy, které nám neumožní rozlišit nic, jsou jako zírání do tmy – zírání do tmy nepřináší žádnou informaci kromě té, že totiž právě zíráme do tmy, tj. že jsme nezískali žádnou informaci. Abychom nějakou informaci získali, musíme rozlišit alespoň jednu ze dvou možností, jednu ze dvou věcí, jednu ze dvou hodnot: tmu od světla, teplo od chladu, mokro od sucha, hlad od pocitu příjemného nasycení – to umí každé malé dítě. Pradávné mýty a legendy nás učí rozlišovat dobro od zla, pravdu od nepravdy, krásu od ošklivosti. V technice rozlišujeme např. tmu od světla, kladné napětí od záporného, jednu ze dvou magnetických orientací. Konkrétně ve výpočetní technice pak tyto fyzikální hodnoty interpretujeme např. jako pravdu a nepravdu nebo jako číslice 0 a 1.

Hodíme-li si kostkou, zjistíme jednu ze šesti možností. Tím získáme víc informace, než když si hodíme korunou, ta nám totiž nabízí jen jednu z pouhých dvou možností. Můžeme získat ještě méně než jednu ze dvou možností? Určitě ano: zjistíme-li například, že v noci nesvítí slunce, že voda v řece teče po proudu, že mrtvý pradědeček zůstává i nadále mrtvý, že volby v roce 1976 znovu vyhrála koalice Národní fronty nebo že trojúhelník má tři vrcholy, nedozvíme se žádnou informaci. Z jedné možnosti se nedá vybírat. Nemůžeme rozlišit méně než *jednu ze dvou* možností – proto je právě takové množství informace zvoleno jako jednotkové, jako jeden **bit** informace.

Jednotkové množství informace se nazývá bit.
Jeden bit nám umožňuje rozlišit jednu ze dvou možných hodnot.

Kolik informace nám přináší jev?

K základnímu nářadí elektrotechnika patří šroubovák. Každému elektrotechnikovi šroubovák někdy upadne. A upadne-li na zem, eletrotechnika to nepřekvapí, zažil to už mockrát. Nic nového se nedozví. Takový jev nepřináší žádnou informaci. Jaké je však elektrotechnikovo překvapení, když šroubovák místo dolů padá vzhůru a dopadne na strop! To je ovšem informace! (např. o tom, že se elektrotechnik právě nachází v silném magnetickém poli)

Zřejmě tedy musí být rozdíl v tom, kolik informace nám přináší jev očekávaný a neočekávaný, častý a řídký, pravděpodobný a nepravděpodobný. Čím častější nebo pravděpodobnější je jev, tím méně informace nám přináší. A naopak, čím je jev řidší nebo nepravděpodobnější, tím víc informace z něj dostaneme. Velikost informace I tedy převráceně závisí na četnosti nebo pravděpodobnosti jevu p:

$$I = f(1/p)$$

f je funkce, kterou dosud neznáme. Tato funkce zobrazí převrácenou hodnotu četnosti jevu do množství informace. Jaká funkce to asi bude?

Pravděpodobnost nezávislých jevů

Představme si, že házíme kostkou. S jakou pravděpodobností padne např. trojka? Každá hodnota může padnout se stejnou pravděpodobností a kostka má šest stěn, tj. šest různých hodnot. Pravděpodobnost, že padne kterákoli zvolená hodnota (např. trojka), je tedy jedna ku šesti, tj. 1/6 = 0,1666… Jaká je pravděpodobnost, že padnou dvě trojky za sebou? Z prvních hodů by jedna šestina měla dát první trojku. Na tuto šestinu prvních hodů pak navážeme druhými hody a z nich zase jedna šestina budou trojky. Tedy šestina ze šestiny hodů nám dá druhou trojku. Pravděpodobnost, že padnou dvě trojky za sebou je tedy $1/6 \times 1/6 = 1/36$. Poznamenejme, že kostka je záměrně vymyšlena tak, aby výsledek druhého hodu nezávisel na hodu prvním – jedná se o nezávislé jevy. A pravděpodobnosti nezávislých jevů se násobí (jak jsme si právě ukázali na příkladu).

Pro dva nezávislé jevy tedy vyjádříme množství informace takto:

$$I = f(1/p_1 \times 1/p_2)$$

Poznamenejme, *že někdy potřebujeme zjišťovat množství informace ze závislých jevů (např. jev p_2 může být podmíněn jevem p_1 nebo na sobě mohou záviset i jinak). V takových případech nemůžeme počítat množství informace přesně podle vztahů, ke kterým zde dospějeme. Místo nich musíme odvodit správné vztahy na základě správného výpočtu pravděpodobnosti závislých jevů (tím se zabývá nauka o pravděpodobnosti). Takové odvození pak bude analogické dalšímu postupu, který zde uvedu pro jevy nezávislé. Zde se tedy nadále omezím jen na jevy nezávislé.*

Množství informace souvisí s pravděpodobností jevů, ale my rozhodně nechceme, aby se množství nezávisle získané informace násobilo. Když např. přijmu zprávu a v ní deset bitů informace a potom další zprávu, která mi přidá ještě pět bitů informace, budu mít celkem patnáct bitů, a ne padesát. Požadujeme tedy, aby se množství informace získané z nezávislých jevů sečítalo. Každý další nezávislý jev musí přinášet další informaci. Více informace ukládáme do přiměřeně větší paměti. Podobně také více informace přenášíme přiměřeně delší zprávou (délka dílčích částí zprávy se nenásobí, nýbrž sečítá). Celkové množství informace tedy získáváme sčítáním:

$$I = I_1 + I_2$$

A po dosazení za I_1 a I_2:

$$I = f(1/p_1) + f(1/p_2)$$

Abychom tedy z pravděpodobností nezávislých jevů dokázali spočítat množství informace, potřebujeme matematickou funkci, která nám změní násobení pravděpodobností na sčítání množství informace:

$$f(x \times y) = f(x) + f(y)$$

Logaritmus

Mnozí z nás asi vědí, jak snadné je násobit mocninami deseti: 20 × 100 = 2000, protože 1 nula + 2 nuly = 3 nuly. V tomto zvláštním případě jsme dokázali násobení převést na sčítání a tím jsme si usnadnili výpočet. 23 × 87 by se nám počítalo o poznání hůř. Kdybychom však uměli číslo 23 i číslo 87 vyjádřit jako mocninu nějakého základu (např. deseti), byl by i tento výpočet snadný:

$$10^{1,36} \times 10^{1,94} = 10^{(1,36+1,94)} = 10^{3,30} \doteq 2001$$

(rovnosti jsou přibližné, čísla jsou zaokrouhlená – v následujících výpočtech také)

Zvolíme-li jiný základ, bude násobení stejně snadné:

$$e^{3,14} \times e^{4,46} = e^{(3,14+4,46)} = e^{7,60} \doteq 1998$$

kde $e \doteq 2,72$ (Eulerovo číslo)

$$2^{4,52} \times 2^{6,45} = 2^{(4,52+6,45)} = 2^{10,97} \doteq 2006$$

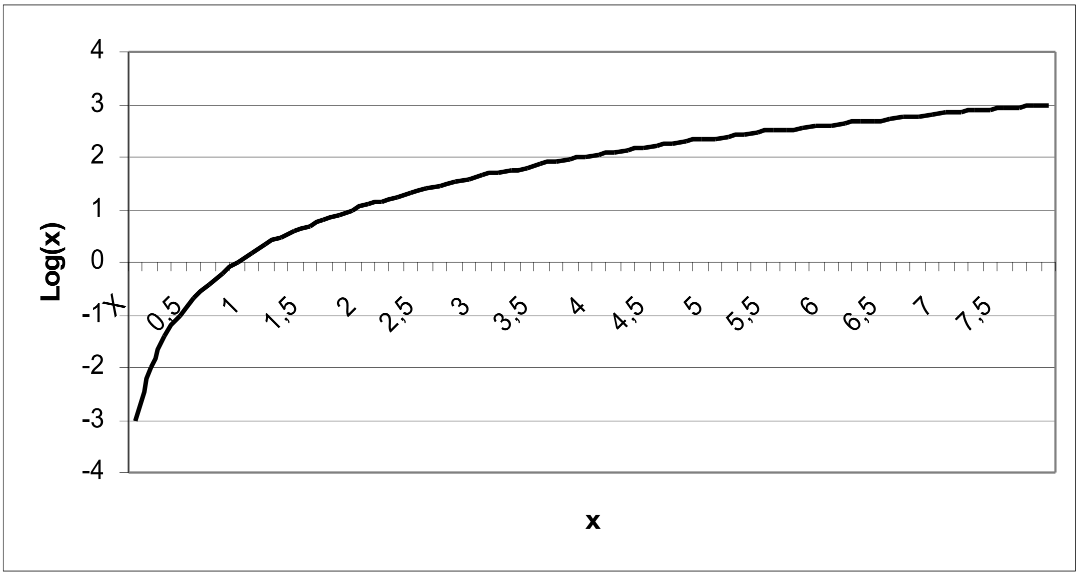
Podobný vztah platí i pro dělení:

$$2001 / 23 = 10^{3,30} / 10^{1,36} = 10^{(3,30 - 1,36)} = 10^{1,94} \doteq 87$$

Samozřejmě by bylo daleko snazší počítat přímo s exponenty. Proto byla definována právě ta funkce *f,* kterou my potřebujeme k výpočtu množství informace. Je to inverzní funkce vůči umocňování, nazývá se *logaritmus*. Jestliže tedy *x, y, z* jsou reálná čísla a jestliže *z* je navíc ještě kladné, tak

$$y = \log_z(x)$$

znamená právě tolik jako $x = z^y$. Číslo *z* je základem mocniny a proto se nazývá i *základem* logaritmu (čteme: logaritmus x při základu z). V matematice se běžně používají logaritmy desítkové (se základem 10) nebo přirozené (se základem e), ve fyzice často přirozené logaritmy a v informatice dvojkové logaritmy (protože množství informace pak vychází v bitech – bit je informace získaná výběrem jedné ze dvou možností, jedna ze dvou je jedna polovina a dvojkový logaritmus poloviny je –1, na znaménku přitom moc nezáleží, protože jenom odlišuje informaci získanou od informace ztracené). Funkci logaritmus znázorníme jejím grafem:

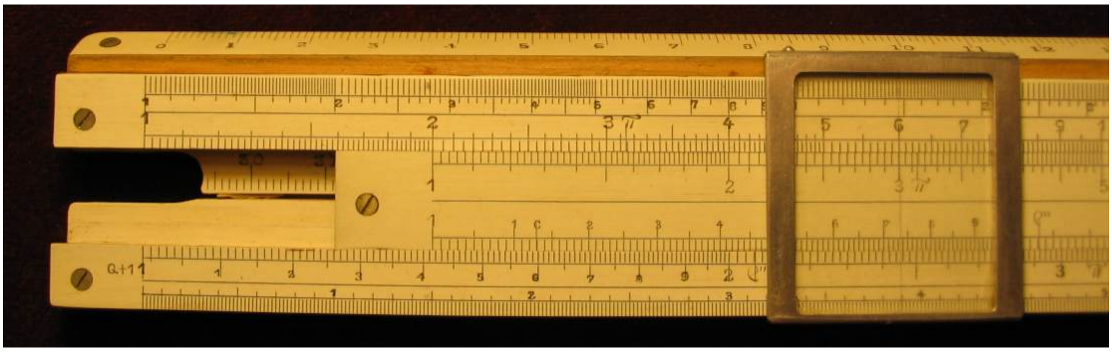


Obr.: graf funkce logaritmus se základem 2

Všimněme si např., že dvojkový logaritmus 8 jsou 3, protože $2^3=8$. Podobně logaritmus 1 je 0, a to při každém základu, protože každé číslo umocněné na nultou je jedna. Logaritmy čísel mezi nulou a jedničkou jsou záporné, protože mocniny se záporným exponentem mají hodnoty zlomků mezi nulou a jedničkou. A konečně nula logaritmus nemá, protože čím dál do záporných čísel jdeme s exponentem, tím menší bude i mocnina, ale nikdy nedosáhne nulové hodnoty. Podobně ani záporná čísla nemají logaritmus, protože umocňováním kladného základu se nemůžeme dostat mimo oblast kladných čísel.

Logaritmy se začaly používat v době renesance nejprve v astronomii a astrologii, později při výpočtech v námořní navigaci. Logaritmus objevil již v první polovině 16. století německý mnich Michael Stifel.[1] Počítání s logaritmy se však rozšířilo až na přelomu 16. a 17. století, kdy skotský matematik John Napier[2] vydal první logaritmické tabulky. Napiera znal i náš Tycho Brahe a s logaritmy počítal. Největší praktický užitek logaritmy přinesly nejprve námořníkům při výpočtech polohy lodí a dost možná tak pomohly evropským mocnostem ovládnout světové oceány a vybudovat celosvětové koloniální říše. Později se logaritmy staly neodmyslitelnou výbavou inženýrů, kteří na logaritmických pravítkách hbitě počítali složité výpočty ocelových konstrukcí mostů a hal, parních strojů i výbušných motorů, lokomotiv, lodí, aut a letadel, rozvodů elektrické energie, rozhlasových vysílačů i spotřební elektroniky – to vše až zhruba do sedmdesátých let dvacátého století, kdy kapesní elektronické kalkulačky věku informačního vypudily logaritmická pravítka doby industriální na smetiště dějin.

[1] viz Michael Stifel. Wikipedia. http://en.wikipedia.org/wiki/Michael_Stifel (navštíveno 15.5.2010) [**!!! odkaz !!!**]
[2] viz John Napier. Wikipedia. http://en.wikipedia.org/wiki/John_Napier (navštíveno 15.5.2010) [**!!! odkaz !!!**]



Obr.: Logaritmické pravítko Matěje Lisa, profesora matematiky na písecké reálce (vyrobil Theodor Zlocha ve Vídni přibližně v devadesátých letech 19. století)

Logaritmické pravítko násobí čísla na stupnicích tím, že sčítá délky stupnic. Na obrázku vidíme logaritmické pravítko s výpočtem 2 × 3 = 6. Pod číslem 2 na horní pevné stupnici je nastaven počátek posuvné stupnice (číslo 1, ukazuje, že 2 × 1 = 2). Ryska jezdce je nastavena na číslo 3 na posuvné stupnici. Nad číslem 3 čteme na pevné stupnici výsledek, a to je číslo 6. Tím, že sčítáme délky pevné a posuvné části pravítka, násobíme čísla vyznačená na logaritmických stupnicích. Všimněte si, jak se logaritmické stupnice zhušťují směrem zleva doprava.

Hledaná funkce, která změní násobení pravděpodobností na sčítání množství informace, je tedy logaritmus. Využijeme toho, že logaritmus má tyto vlastnosti:

- $\log_z (a \times b) = \log_z (a) + \log_z (b)$
  např. $\log_2 (8 \times 16) = \log_2 (8) + \log_2 (16) = 3 + 4 = 7$

- $\log_z (1/c) = -\log_z (c)$
  protože logaritmus jedné je nula a tedy $\log_z (1) - \log_z (c) = -\log_z (c)$
  např. $\log_2 (1/4) = -\log_2 (4) = -2$

Nyní už dokážeme spočítat množství informace, které nám přinesou dva nezávislé jevy:

$$I_1 + I_2 = (\log_2 (1/P_1 \times 1/P_2)) =$$
$$= (\log_2 (1/P_1)) + (\log_2 (1/P_2)) =$$
$$= -(\log_2 (P_1)) - (\log_2 (P_2))$$

Tento vztah ještě zobecníme pro N nezávislých jevů takto:

$$I = \Sigma (I_i) = \Sigma (\log_2 (1/P_i)) = -\Sigma (\log_2 (P_i))$$

kde I je množství informace,
P_i je četnost i-tého jevu a
Σ označuje součet všech hodnot výrazu uvedeného v závorce pro i od 1 do N

Příklad

```
nazdar
```

Obr.: zpráva „nazdar“

Spočítejme, kolik informace obsahuje zpráva „nazdar“. Délka zprávy N = 6 znaků. Znak „a“ se ve zprávě opakuje dvakrát, ostatní znaky se ve zpravě vyskytují jen jednou. Počty znaků ve zprávě (tj. absolutní četnosti) N_i vydělíme délkou zprávy N, tím získáme relativní četnosti P_i, ty zlogaritmujeme a logaritmy nakonec sečteme:

	abs. četnost	rel. četnost	logaritmus
	N_i	$P_i = N_i / N$	$\log_2(P_i)$
n	1	0,166666667	-2,58496
a	2	0,333333333	-1,58496
z	1	0,166666667	-2,58496
d	1	0,166666667	-2,58496
a	2	0,333333333	-1,58496
r	1	0,166666667	-2,58496
N = 6		–součet =	13,50978

Zpráva tedy nese asi 13,51 bitů informace. To je v průměru asi 2,25 bitu na 1 znak zprávy.

Stejných výsledků můžeme dosáhnout i jiným postupem, který bývá výhodnější. V řádcích tabulky uvedeme jednotlivé znaky abecedy, takže před sčítáním logaritmů pak ještě musíme vynásobit každý logaritmus počtem výskytů znaku. Výhoda spočívá v tom, že množství informace na jeden znak zprávy pak snadno spočítáme jako tzv. vážený průměr, jak je uvedeno v posledním sloupci tabulky:

	abs. četnost	rel. četnost	logaritmus		
	N_i	$P_i = N_i / N$	$\log_2(P_i)$	$N_i \times \log_2(P_i)$	$P_i \times \log_2(P_i)$
a	2	0,333333333	-1,58496	-3,169925001	-0,528320834
d	1	0,166666667	-2,58496	-2,584962501	-0,430827083
n	1	0,166666667	-2,58496	-2,584962501	-0,430827083
r	1	0,166666667	-2,58496	-2,584962501	-0,430827083
z	1	0,166666667	-2,58496	-2,584962501	-0,430827083
N=	6		–součet =	13,509775	2,251629167

Vážený průměr je jen jiný způsob, jak vypočteme aritmetický průměr – výsledek je stejný. Místo abychom součet všech průměrovaných hodnot dělili jejich počtem:

$$-\Sigma (N_i \times \log_2 (P_i)) / N,$$ případně ve tvaru

$$(1 / N) \times -\Sigma (N_i \times \log_2 (P_i)),$$

distribuujeme koeficient (1 / N) do jednotlivých sčítanců:

$$-\Sigma\ ((N_i / N) \times \log_2 (P_i))$$

a to pak ještě zjednodušíme tím, že dosadíme relativní četnost $P_i = N_i / N$ takto:

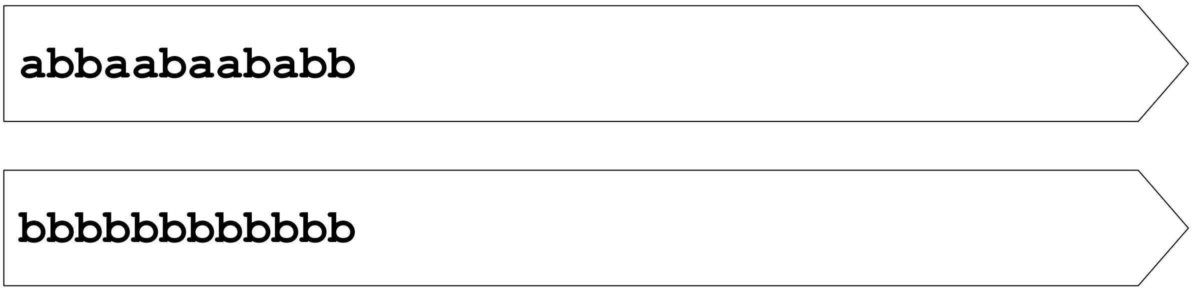
$$\mathbf{H = -\Sigma\ (P_i \times \log_2 (P_i))}$$

kde koeficient P_i se nazývá váha (a odtud pak „vážený“ průměr). Vážený průměr je jen takový malý fígl, který si vymysleli statistici, aby se jim lépe počítalo. Nicméně vzorcem, ke kterému jsme dospěli, se koncem 40. let 20. století proslavil jeden ze zakladatelů teorie informace Claude Elwood Shannon.[3]

A proč nás vůbec zajímá průměrné množství informace na jeden znak? To si ukážeme na dalším příkladu.

Příklad

Která ze dvou zpráv na obrázku obsahuje víc informace?



Obr.: dvě zprávy

První zpráva obsahuje

$$-(6 \times \log_2 (6/12)) - (6 \times \log_2 (6/12)) = -\ (6 \times (-1)) - (6 \times (-1)) = 6 + 6 = 12 \text{ bitů}$$

Druhá zpráva obsahuje

$$-(0 \times \log_2 (0/12)) - (12 \times \log_2 (12/12)) = -\ (0 \times \log_2 (0)) - (12 \times 0) = 0 + 0 = 0 \text{ bitů}$$

Všimněme si, že obě zprávy jsou složeny ze znaků stejné abecedy (pro jednoduchost připusťme, že druhá zpráva obsahuje nulový počet znaků

[3] viz Claude Shannon. Wikipedia. http://en.wikipedia.org/wiki/Claude_Shannon (26.5.2010) [**!!! odkaz !!!**]. Shannon ovšem měl na co navazovat: základy k teorii informace položil už v roce 1928 Ralph Hartley, který ovšem k měření informace používal logaritmů desítkových místo dvojkových – viz Wikipedia contributors. Ralph Hartley [Internet]. Wikipedia, The Free Encyclopedia; 2012 Apr 30, 18:53 UTC [cited 2012 Aug 4]. Available from: http://en.wikipedia.org/w/index.php?title=Ralph_Hartley&oldid=489993013. [**!!! odkaz !!!**]

„a“ – v tom případě je abeceda obou zpráv tvořena dvěma znaky, a to „a“ a „b“). Obě zprávy jsou stejně dlouhé. Jak to, že nenesou stejné množství informace?

Mohli bychom říci, že každá z obou zpráv by mohla přenášet stejné množství informace, které je dáno jednak délkou zprávy, jednak počtem znaků v abecedě (tj. kolik znaků dokážeme rozlišit, když přečteme jeden znak). Zpráva má tedy určitou kapacitu, ale ta není dokonale využita. Je to podobné jako když zkomprimujeme („zazipujeme“) soubor na disku: jeho velikost se několikanásobně zmenší, ale informace v něm zůstane všechna. Místo na disku zřejmě původně nebylo dokonale využito. Čím se liší zprávy v našem příkladu? V první zprávě mají prvky abecedy **stejnou četnost** (6 krát „a“ a 6 krát „b“), kdežto ve druhé zprávě mají **četnost různou** (žádné „a“ a 12 krát „b“).

Delší zpráva může přenášet víc informace. Zamysleme se nyní nad otázkou, jaký vliv na využití zprávy má její délka? Vezměme si teď pro srovnání ještě třetí zprávu:

abbaabaababbbaabbabbabaa

Obr.: dlouhá zpráva

Tato zpráva obsahuje

$$-(12 \times \log_2 (12/24)) - (12 \times \log_2 (12/24)) = 12 + 12 = 24 \text{ bitů.}$$

Množství informace ve třetí zprávě je větší, protože i její délka a tedy i kapacita je větší. Ale využití třetí zprávy není lepší než využití zprávy první. Chceme-li tedy mluvit o tom, jak dobře je zpráva využita, musíme abstrahovat od její délky. A toho právě dosáhneme, když spočítáme průměrné množství informace, které připadá na jeden znak zprávy.

Jak už jsme si ukázali, průměrné množství informace snadno spočítáme jako vážený průměr podle vzorce:

$$-\Sigma (P_i \times \log_2 (P_i))$$

První zpráva tedy obsahuje

$$\begin{aligned} &-(6/12 \times \log_2 (6/12)) - (6/12 \times \log_2 (6/12)) \\ &= -(1/2 \times (-1)) - (1/2 \times (-1)) \\ &= +(1/2) + (1/2) \\ &= 1 \text{ bit na znak.} \end{aligned}$$

Druhá zpráva obsahuje

$$\begin{aligned} &-(0/12 \times \log_2 (0/12)) - (12/12 \times \log_2 (12/12)) \\ &= -(0 \times \log_2 (0)) - (1 \times 0) \end{aligned}$$

= 0 + 0
= 0 bitů na znak.

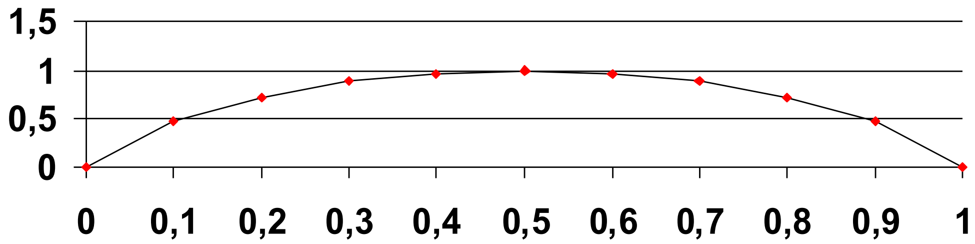
Třetí zpráva obsahuje

$$–(12/24 \times \log_2 (12/24)) – (12/24 \times \log_2 (12/24))$$
$$= –(1/2 \times (-1)) – (1/2 \times (-1))$$
$$= +(1/2) + (1/2)$$
$$= 1 \text{ bit na znak.}$$

Vidíme, že první a třetí zpráva obsahuje stejné průměrné množství informace. Shrňme: průměrné množství informace

- nezáleží na délce zprávy
- charakterizuje např. efektivitu zdroje informace, využití paměti nebo využití přenosové kapacity informačního kanálu

Jak souvisí využití kapacity zprávy s četností znaků abecedy? Abeceda našich tří zpráv je množina dvou znaků {a , b}, kde součet relativních četností znaků ve zprávě P(a) + P(b) = 1 (tj. $N_a/N + N_b/N = N/N$, protože $N_a + N_b = N$). Vynesme do grafu závislost průměrného množství informace H na relativní četnosti P(a) (pro P(b) bychom mohli nakreslit úplně stejný graf):



Obr.: závislost efektivity na četnosti znaků

Z grafu vidíme, že **průměrná informace je největší, když znaky abecedy jsou ve zprávě stejně četné.**[4] Toto pravidlo se dá zobecnit i pro abecedy, které obsahují víc než jenom dva znaky (avšak k tomu, abych se abgéboval k takovému zobecnění, se mi nedostává mentálního kojachu).

Využití zprávy nyní můžeme vyjádřit i číselně jako tzv. účinnost neboli **efektivitu**

$$E = H / H_{max}$$

To je poměr mezi skutečně dosaženou průměrnou informací zprávy a největší možnou průměrnou informací (tj. kapacitou) zprávy. Efektivita

[4] To se dá zjistit nejen odhadem z grafu, ale také výpočtem. Kdo umí počítat s derivacemi ví, že maximum funkce se nachází v bodě, kde funkce je konkávní a její derivace je nulová – tj. graf funkce je vypouklý a tečna ke grafu je vodorovná.

vlastně vyjadřuje, na kolik procent své původní délky by se dala zkrátit zpráva při použití dokonalé komprese dat.

Není-li kapacita zprávy zcela využita, je zřejmě část dat ve zprávě nadbytečná. Nadbytečnost neboli **redundanci** vyjadřujeme relativně vzhledem k délce zprávy (podobně jako efektivitu):

$$R = 1 - (H / H_{max})$$

Na první pohled by se mohlo zdát, že nadbytečnost jen plýtvá kapacitou zpráv. Ale není tomu tak. Když někdo na otázku hostitelky (např.: „Alfréde, dáte si ještě osmnáctý knedlík?“) slušně odpoví „ne ne!“ místo prostého „ne“, tak nesdělí hostitelce dvakrát víc informace. Sdělí právě tolik, jako by řekl „ne“ jenom jednou. Jeho zpráva má tedy poloviční nadbytečnost. Host má ovšem větší jistotu, že mu hostitelka správně porozumí. Kdyby náhodou první „ne“ přeslechla, může ještě zachytit to druhé „ne“.

Problém je v tom, že každý komunikační kanál trochu kazí zprávy, které přenáší (kazí je jen trochu, protože kdyby je kazil hodně, nebyl by to komunikační kanál). Všichni známe šum, který kazí telefonní hovory, rušení v rádiu nebo zrnitost fotografií exponovaných na citlivý materiál. A právě nadbytečnost ve zprávách umožňuje, aby příjemce dokázal správně rekonstruovat přijatou zprávu. Ovšem když zpráva není dost nadbytečná a příjemce ji proto nedokáže rekonstruovat, tak by měl aspoň poznat, že je pokažená šumem.

Je přece důležité, aby příjemce rozpoznal šum jako šum a neinterpretoval jej jako užitečnou informaci. Ve zprávách bez nadbytečnosti totiž bohužel musí dojít právě k tomu, že příjemce „rozumí“ i tomu, co odesílatel neposlal, co do zprávy chaoticky vnesl komunikační kanál. Teorie informace ovšem neodlišuje informaci od dezinformace – ani v případě šumu a tím méně pak v případě, že odesílatel nebo kanál úmyslně klame příjemce.

Nadbytečnost tedy může zlepšit odolnost zprávy proti rušení. Kromě toho můžeme nadbytečnost využít i k rychlejšímu vybavování uložené informace.

> **Příklad:** Představte si, že hledáme např. kožního lékaře v seznamu lékařských ordinací. Kožního lékaře stěží bude někdo potřebovat naléhavě, takže nám stačí, když bude uveden jenom jednou na vhodném místě (např. ve správné rubrice). Záchrannou službu však potřebujeme najít co nejrychleji. Proto bude vhodné, abychom ji v seznamu opakovali tak často, že ji každý okamžitě najde (např. na každé obrazovce webu nebo na každé stránce tištěného seznamu).

Také lidská mysl uchovává informace tak, aby optimalizovala na jednu stranu objem dat a na druhou stranu rychlost vybavování. Uvažování v obecných pojmech slouží ke kompresi dat.

> **Příklad:** Z toho, že Mikeš je kocour, mohu usoudit, že stejně jako všechny kočky i Mikeš bude chytat myši, lézt po stromech a nebude se rád koupat, zato bude rád spát v suchu a v teple. Ani u jiných koček a kocourů tohle nebude jinak a stačí, když si to zapamatuji jenom jednou. Ale to, že Mikeš je černý kocour, že ptáčky zásadně neloví a že spát chodí za pec s Pepíkem, to platí pro Mikeše, ale pro jiné kočky to nejspíš platit nebude – takové věci si musím zapamatovat o každé kočce zvlášť.

Tím, že si pamatuji jenom jednou, co mají všechny kočky společného, šetřím paměť a jsem schopen si zapamatovat víc, než kdybych si každý údaj pamatoval zvlášť a nevyužíval souvislostí – nebudoval si znalostní strukturu. Podobně to dělá i komprimační program v počítači, když „zipuje“ soubory. Na některé otázky nepotřebuji odpovědět honem a mohu i dost dlouze rozvažovat s pomocí obecných pojmů a analogií, než se doberu správné odpovědi (např. právě jsem dost dlouze vzpomínal, jak bych popsal kocoura Mikeše).

Dostanu-li se však do krizové situace, např. když se na výletě vážně zraní kamarád a bude potřebovat záchranku, nebudu mít čas na dlouhé úvahy – musí mi okamžitě naskočit číslo tísňového volání 112 nebo záchranné služby 155, protože na přemýšlení nebudu mít čas a ve stresu bych rozumného úsudku ani nebyl schopen. Podle mého názoru je možné, že právě pouhá optimalizace objemu dat a rychlosti vybavování stojí za celým lidským myšlením, za schopností abstrahovat a konkretizovat, uvažovat v obecných pojmech a v analogiích, indukovat a dedukovat.[5]

> **Cvičení:** Zprávu **abbaabaababb** bychom mohli zakódovat úsporněji. Nahradíme-li dílčí řetězce **abb** znakem **c** a dílčí řetězce **aab** znakem **d**, dostaneme tutéž zprávu překódovanou jako **cddc**. Tato zpráva obsahuje 4 bity informace. Jenže původní zpráva se svými 12 bity byla kódovaná zcela efektivně, neboť četnost znaků **a** a **b** byla stejná. Zkuste vysvětlit, jak je to možné. Kolik informace obsahují náhrady řetězců **c = abb**, **d = aab**?

O čem teorie informace nehovoří

Nyní se ovšem dostáváme k hranici, kterou teorie informace nemůže překročit. Však je taky výslovně omezena na jevení jevů a nepokouší se měřit informaci obsaženou ve „faktech“ nebo v řádu světa jako takového – tyto problémy přenechali otcové teorie informace kolegům filosofům. Tou nepřekročitelnou hranicí je stručně řečeno pragmatika: otázky kdo komu a proč něco sděluje, v jakých souvislostech a za jakých podmínek se sdělení přenáší a jak dalece je příjemce schopen nebo ochoten zprávě porozumět. Hodně záleží na účelu, proč odesílatel zprávu sděluje a proč příjemce naslouchá.

---

[5] Viz Guizzardi, Giancarlo: *Ontological Foundations for Structural Conceptual Models.* (CTIT PhD Thesis Series, No. 05-74, Telematica Instituut Fundamental Research Series, No. 015 (TI/FRS/015)). Enschede, The Netherlands: Centre for Telematics and Information Technology, 2005. str. 115 [**!!! odkaz !!!**] a tam odkaz na Markman, Ellen. *Categorization and naming in children.* MIT Press, 1989, str. 11. [**!!! odkaz !!!**]

Tak např. nápis MAT na skleněných dveřích znamená „táhni“, protože je určen příchozím z druhé strany dveří. Ti ovšem čtou tentýž nápis jako TAM a znamená pro ně „tlač“. Jiný příklad: slovo ANO může obsahovat 1 bit informace, když je to pro nás jedna z dvojice hodnot ano / ne. V tom případě je lhostejné, zda slyšíme slovo vyslovené nebo je vidíme napsané – tím spíš nezáleží ani na velikosti písma nebo na popsané ploše. Ale slovo ANO může obsahovat taky asi kolem 10 bitů informace, když nám slouží k rozlišení jazyka (IGEN maďarsky, DA bulharsky, rumunsky, rusky, KEN hebrejsky atd.) Samozřejmě že nemusí jít jenom o slova, ale informace je obsažena obecně v čemkoli pravidelném (co nám umožňuje rozlišovat). Záleží jak na *schopnosti* přijmout zprávu (příjemce musí rozumět jazyku zprávy), tak i na *ochotě* zprávu přijmout (jestli to příjemce považuje za účelné nebo zda se naopak rozhodne sdělení ignorovat).

Vidíme, že teorie informace neodpovídá úplně vyčerpávajícím způsobem na svoji otázku (kolik informace nese zpráva?) Na čem ještě tedy závisí množství informace? Ať už se jedná o množství informace obsažené v tomtéž systému, nebo přenášené v téže zprávě a nebo jevící se jako tentýž jev – množství informace závisí i na okolnostech, kterými se teorie informace nezabývá. Mezi takové okolnosti patří především

- účel (proč odesílatel odesílá zprávu, proč ji příjemce přijímá, k čemu informaci potřebuje)
- vlivy prostředí (za jakých okolností a v jakých souvislostech je zpráva interpretována)
- schopnosti a ochota komunikujících
- jakož i jiné okolnosti, které mě právě nenapadají

Informace a hmota, informace a energie

Zajímavé je, že Claude Shannon nazval průměrné množství informace na jeden znak abecedy slovem **entropie** – a došlo k tomu asi takhle: Po druhé světové válce pracoval v americkém Ústavu pokročilých studií v Princetonu[6] výkvět světové matematiky a fyziky: Albert Einstein, Kurt Gödel, Robert Oppenheimer, János von Neumann a také Claude Shannon. Život každého z nich by vydal na román a jejich společné působení na legendu srovnatelnou s kterýmkoli z řeckých nebo biblických mýtů. Způsobili tenkrát zásadní převrat v dalším vývoji civilizace. Mnoho z nich uteklo z Evropy před nacismem. Za války někteří z nich vymýšleli atomovou bombu. A po válce se v téhle společnosti začal líhnout nový obor, kterému dnes říkáme informatika – teorie informace, kybernetika,[7] teorie her (von Neumann). Další ze zakladatelů teorie her, Oskar

[6] Viz Institute for Advanced Study. Wikipedia. http://en.wikipedia.org/wiki/Institute_for_Advanced_Study (13.6.2010) [**!!! odkaz !!!**]

[7] s výhradou, že zakladatel kybernetiky Norbert Wiener působil na MIT, kde ovšem Shannon také studoval;
viz Norbert Wiener. Wikipedia. http://en.wikipedia.org/wiki/Norbert_Wiener (14.6.2010) [**!!! odkaz !!!**],
viz Claude Shannon. Wikipedia. http://en.wikipedia.org/wiki/Claude_Elwood_Shannon (14.6.2010) [**!!! odkaz !!!**],
viz Massachusetts Institute of Technology. Wikipedia. http://en.wikipedia.org/wiki/MIT (14.6.2010) [**!!! odkaz !!!**]

Morgenstern, působil hned v sousedství, na princetonské univerzitě. Dovolte ještě perličku k teorii her: její otcové úspěšně testovali svou novou teorii na sázkách v Las Vegas a potom s ještě větším úspěchem na obchodech s akciemi na burze – což nějaký Karel Janeček dělá i dnes.

> **Cvičení:** Přiřaďte ke jmenovaným osobám (tj. Albert Einstein, Kurt Gödel, Robert Oppenheimer, János von Neumann) tyto objevy:
> - meze zhroucení neutronových hvězd v důsledku jejich vlastní tíže
> - princip stroje, který dokáže sám sebe opravovat a rozmnožovat
> - neúplnost teorie přirozených čísel (jsou takové věty o přirozených číslech, že nedokážeme rozhodnout, jestli jsou pravdivé nebo nepravdivé)
> - statistické rozdělení fotonů a jiných bosonů (společně se Šotendronátem Bošúem) – v podstatě jeden ze způsobů, jak spočítat množství informace obsažené v hmotě

Tak se stalo, že když Shannon vymyslel svůj nový vzorec, zašel k němu na obvyklou návštěvu John von Neumann a jen tak mimochodem (snad než paní Shannonová přinesla hostovi a rodinnému příteli svůj báječný tea & pie) se přeptal: „Tak co zrovna děláš, Claude, starý brachu?" A Shannon mu ukázal svůj nový vzorec: „právě jsem objevil novou veličinu a nevím, jak bych ji měl pojmenovat." Von Neumman si prohlíží papír se vzorcem, chvíli přemýšlí a pak prohlásí: „Tenhle vzorec já znám." „Nemůžeš ho znát, já jsem ho právě vymyslel." „Kdepak Claude, tenhle vzorec znám z fyziky, tak se přece počítá entropie. Nikdo sice kloudně nechápe, co to vlastně je, ta entropie, ale možná že právě to by byl dobrý důvod, abys tak pojmenoval svoji novou veličinu – co myslíš?" Ať už se rozhovor odehrál tak nebo jinak, Claude Shannon přistoupil na to, že jeho *průměrné množství informace* je vlastně zobecněná fyzikální entropie, i když s opačným znaménkem: množství informace je totiž míra uspořádanosti v systému, kdežto **entropie je míra neuspořádanosti,** tj. nedostatek informace. Zdá se tedy, že množství informace musí nějak úzce souviset s hmotou, s energií nebo s jinými fyzikálními veličinami. Ale není to tak docela pravda – nezapomeňme, že se jedná o zobecnění, které se neomezuje na hranice fyziky. A ostatně ani ta fyzikální entropie není tak úplně „fyzikální" veličina, jak uvidíme v následující kapitole.

> **Cvičení:** ověřte pravdivost historky pomocí ještě aspoň jednoho nezávislého zdroje.

Informace je kvalitou *jevů,* nikoli nutně kvalitou něčeho objektivního nebo skutečného. Jev může být i klam… Může jít o kvalitu jevů nejen fyzikálních, ale také chemických, biologických nebo jazykových, historických, společenských (veřejné mínění právě tak jako davová psychóza) nebo psychických (smyslové vnímání právě tak jako sen). Může jít o jevy přirozené nebo taky nadpřirozené. Je tedy jedním ze základních požadavků na množství informace, že množství informace nemá nějak jednoduše nebo přímočaře souviset s množstvím hmoty.

A přece je zřejmé, že signál je *fyzikální* děj, který přenáší informaci. Ponecháme-li stranou telepatii, zjevení, vnuknutí, osvícení a další paranormální nebo mystické způsoby přenosu informace a omezíme se na technicky využitelné fyzikální principy, můžeme s velkou mírou jistoty (byť i s velkou dávkou sebevědomí) tvrdit, že bez hmotného nosiče informaci nepřeneseme. Hmotným nosičem může být např. papír a inkoust, magnetický materiál pevného disku, vrstva kompaktního disku, která odráží nebo neodráží světlo laserového paprsku, elektrony ve vodiči nebo fotony v optickém kabelu.

Kolik kilogramů hmoty potřebujeme k přenesení jednoho bitu informace? – odpověď není ani snadná, ani jednoznačná. Jedna věc je, že při záznamu informace nemusíme dokonale využít kapacitu média. List papíru můžeme popsat ze čtvrtiny nebo z poloviny a nebo taky můžeme mezi řádky zcela popsaného listu vpisovat řádky další.[8] V předchozí kapitole jsme si již vysvětlili, že záleží také na tom, jak efektivně kódujeme informaci: zazipovaný textový soubor bude několikanásobně kratší než originál, aniž by se z něj ztratil jediný bit informace.

Ovšemže se na dva CD může vejít až dvakrát víc informace než na jeden CD, ale na druhou stranu DVD představuje zhruba stejné množství hmoty a dá se na něj zaznamenat třeba i několikanásobně víc informace než na CD, protože tam dosáhneme vyšší hustoty záznamu. Množství zaznamenané informace tedy nezávisí jenom na množství hmoty, ale také na hustotě záznamu. Hustotu však nelze libovolně zvyšovat. Musíme brát v úvahu, že různé druhy a množství hmoty jsou různě odolné proti poškození nebo zničení, tj. proti rušení a šumu. Krajní hranicí pro zvyšování hustoty záznamu pak bude rozlišitelnost fyzikálních jevů na úrovni částic, např. schopnost rozlišit atomové orbitaly nebo spin elektronů. Proto taky spíš než hmotnost v kilogramech by nás měl zajímat počet (rozlišitelných) částic, případně množství látky v molech nebo kilomolech. Ale ani to není výstižné: co třeba nápis vytesaný do kamene? – tam informaci nese ta hmota, která chybí. Podobně socha, ze které sochař otesal přebytečný kámen. Nějaká souvislost mezi množstvím informace a množstvím hmoty tedy zřejmě bude, ale není přímočará (není to např. úměra).

Podobně nejasná je na první pohled **souvislost mezi množstvím informace a energie.** Technicky vzato, vyslání a přijetí zprávy je spojeno s vynaložením nebo spotřebou energie. Zdá se však, že potřebné množství energie můžeme téměř libovolně zmenšovat až k hranici, kdy šum komunikačního kanálu pohltí přenášenou informaci. Krajní hranicí pak

[8] To prý dělal ruský básník Velemír Chlebnikov za revoluce. Na lístek papíru napsal básničku. I z druhé strany napsal básničku. Když ho pak napadla další báseň, obrátil papírek vzhůru nohama a psal mezi řádky. Zcela popsané papírky pak cpal do svého slamníku. A když si chtěl zakouřit, vytáhl ze slamníku papírek s básničkami a ubalil si do něj cigaretu. Kolik informace se asi vešlo na jeden lístek papíru? Jak to záviselo na velikosti písmen? Co když v některých básničkách škrtal? A když básničku spálil, kolik tepla se uvolnilo? Souviselo nějak množství uvolněné energie s množstvím informace v básničce? (Neznáte Chlebnikova? – kdo rád Morgensternovy hříčky, toho jistě nadchne také Chlebnikov)

zase budou kvantové jevy, protože kvantum energie se nám jeví jako nedělitelné.

> **Poznámka:** Neměli bychom zapomenout ani na to, že informace má hodnotu energie potřebné k získání této informace, případně energie ušetřené za omyly, kterým se díky získané informaci vyhneme. To je hledisko spíš ekonomické než technické – jde spíš o porovnávání hodnot než o vysvětlení fyzikálních principů, nicméně i tak vyjadřuje dost důležitý vztah mezi informací a energií.

Vidíme, že se nedopátráme nějakého jednoduchého vzorečku, kterým bychom mohli přepočítávat kilogramy nebo kilowatthodiny na bity. A přece tušíme souvislost. Je to spíš otázka víry než rozumné úvahy nebo objektivních ověřitelných faktů, nicméně i tak, na základě pouhé víry, uvažujeme o světě a o věcech jako že jsou, že když pozorujeme jev, tak samozřejmě předpokládáme, že se nám jeví *něco* (a nikoli nic), že svět se nějak člení na věci sám od sebe a jestli ne zrovna na věci, tak přinejmenším musí být nějak pravidelný nebo uspořádaný, protože jinak by se nám věci nejevily jako součásti světa. Jestliže se neshodneme v tomto punktu, dosáhli jste *satori* – maximálního poznání. Odložte tedy prosím tuto knihu, protože jakákoli další četba je zbytečná, a odeberte se do zasloužené *nirvány.* Pokud se však shodneme, nezbývá nám než pokračovat v našem – ovšemže konec konců marném – pachtění. Zdá se nám tedy, že informace musí být nějak obsažena ve světě a ve věcech. Je jistě rozdíl mezi uspořádáním atomů v krystalu a chaotickým pohybem molekul plynu nebo mezi organizovaností organismu a dezorganizací hnijícího bahna. Za informaci můžeme považovat rozdíl mezi uspořádáním věcí ve světě a naprostou neuspořádaností nicoty, mezi řádem světa a primordiálním chaosem, právě tak jako mezi předvídatelnou pravidelností a nepředvídatelnou náhodností. Naproti tomu neuspořádanost či chaos ve srovnání s řádem či kosmem je opakem informace, je to nedostatek řádu a tedy i informace. Soustřeďme se na jádro předmětu, který zkoumáme a zformulujme otázku jdoucí co možná na dřeň: **je nějaká souvislost mezi samotnou pravidelností a energií?**

Představme si sněhovou vločku: je to pravidelně uspořádaný krystal ledu. Když vločku zahřejeme třeba v dlani, její uspořádání zmizí. Vločka se promění v kapku vody. Teplem se kapka odpaří a tím se molekuly vody chaoticky promísí s okolním vzduchem. Zdá se, že dodáním energie zmizelo uspořádání sněhové vločky.

Strom se ovšem chová jinak: pohlcuje sluneční energii a roste, zvětšuje se objem i složitost organizované hmoty.[9] Nemá-li strom přísun energie, zahyne. Složité uspořádání pomine a energie dosud vázaná v organické

[9] Viz Norbert Wiener, Kybernetika a společnost, Nakladatelství ČSAV, Praha 1963, str. 50 (z originálu The Human Use of Human Beings. Cybernetics and Society, druhé doplněné vydání, Doubleday, New York 1954 přeložil Karel Berka, poprvé vyšlo v nakladatelství Houghton Mifflin, Boston, 1950) [**!!! odkaz !!!**]

hmotě se může uvolnit.[10] Živé bytosti mají tendenci uspořádávat sebe i svět kolem sebe – a k tomu potřebují energii.

Nedá se jednoznačně říct, že by přísun energie vždy vedl k tomu, že se hmota uspořádá a že naopak při rozkladu uspořádaných struktur by se energie uvolňovala. Energie se uvolňuje jak při štěpení, tak i při fúzi atomů. Při hoření uhlíku nebo oxidu uhelnatého vzniká složitější sloučenina, kdežto při hoření etylalkoholu vznikají sloučeniny jednodušší (v obou případech se energie uvolňuje). Tímto způsobem asi žádnou zřejmou souvislost mezi energií a pravidelností nenajdeme.

Podobně: nasypeme-li do hrnku červené a bílé fazole, přiklopíme a důkladně protřepeme, fazole se promíchají. Čím déle protřepáváme, tím rovnoměrněji se fazole promíchají. A ať protřepáváme sebevíc nebo sebevětší silou, nikdy už neoddělíme červené fazole od bílých. Podobné je to s mícháním kořalek, koktejlů i jiných směsí – promícháváním se samovolně nerozloží na svoje složky. Přestaneme-li míchat, některé směsi se zase rozloží, ale ne samovolně: smetana se usadí na hladině mléka působením zemské tíže, zrno od plev oddělíme foukáním, zlato od hlušiny rýžováním. Také hrách promíchaný s popelem Popelčini holubi oddělili, ale museli vyzobávat zrnko po zrnku. Chceme-li oddělit složky směsi, musíme vždy vynaložit práci nebo energii. Zdá se, že energii musíme vynaložit i na míchání složek, aby vznikla směs…

A přece ne vždy. Např. plyny samovolně difundují, až vznikne rovnoměrně promíchaná směs. Nebo si představme kovovou lžičku v hrnku horké kávy: vyndáme lžičku z hrnku, jeden konec bude horký, druhý ne. Položíme-li lžičku tak, aby nepohlcovala ani neuvolňovala teplo nijak výrazně, můžeme pozorovat, že chladnější konec se bude ohřívat, zatímco teplejší konec bude chladnout, a to tak dlouho, až se teploty vyrovnají. Vidíme, že nemusíme dodávat ani odebírat energii a uspořádanost přesto pomine sama od sebe. Zajímavé je, že ještě nikdo nikdy nepozoroval, aby se kovová lžička sama od sebe na jednom konci zahřála a tím současně na druhém konci zchladla – **neuspořádanost (tj. entropie) sama od sebe nikdy nepomíjí.**

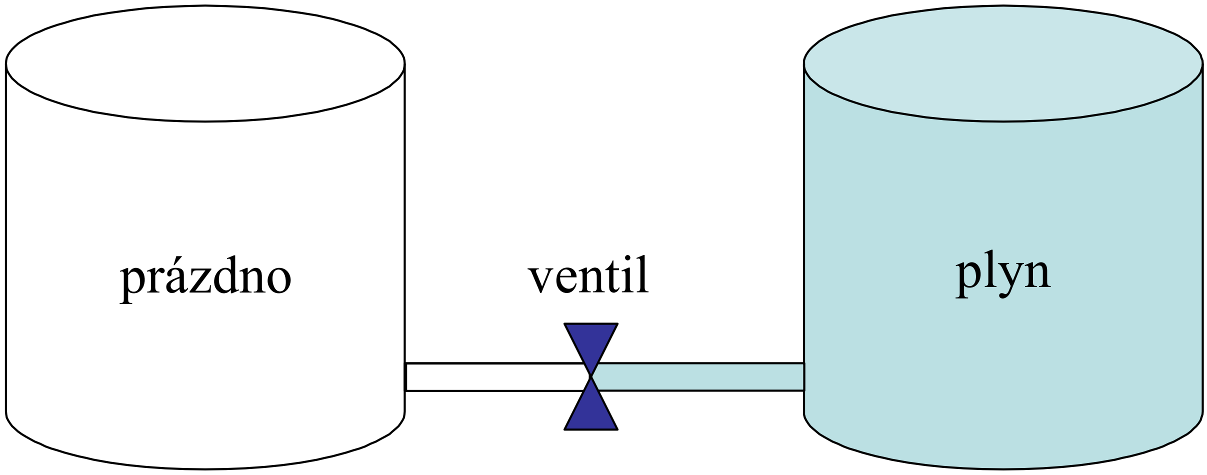
Tuto zákonitost pozoroval Joseph Gay-Lussac[11] na soustavě dvou nádob. Nádoby propojil trubkou s kohoutkem a z jedné nádoby vyčerpal vzduch. Nyní měl v jedné nádobě prázdno a podtlak, ve druhé nádobě vzduch a přetlak – to je ovšem pozorovatelný rozdíl, tedy informace. Pak otočil kohoutkem – a co se stalo? Vzduch z plné nádoby začal proudit do nádoby prázdné a proudil tak dlouho, až se tlak vyrovnal – informace o rozdílu tlaků se vytratila. Gay-Lussaca zajímalo, zda se při tom uvolnila nebo pohltila nějaká energie. Ponořil tedy soustavu nádob do vany s vodou a měřil teplotu vody. Čím přesněji měřil, tím stejnější zůstávala teplota vody ve vaně. Energie se ani neuvolnila, ani nepohltila. Množství informace se

---

[10] zatímco v teple a vlhku se strom rozloží a energie se uvolní, v suchu nebo silném mrazu se mrtvý strom nerozkládá a energie se neuvolňuje

[11] viz Joseph Louis Gay-Lussac, Wikipedia, http://cs.wikipedia.org/wiki/Joseph_Louis_Gay-Lussac (4.6.2010) [**!!! odkaz !!!**]

snížilo, entropie vzrostla. **Entropie v izolované soustavě samovolně roste.**



Obr.: Otevřeme ventil a plyn se sám od sebe rovnoměrně rozptýlí. Zvětší se entropie, ale energie se nepohltí ani neuvolní.

Jak obecně platí ve fyzice princip rostoucí entropie? Má nějaké meze? Na první pohled se zdá, že jedinou mezí je sama fyzika. Asi proto, že mimo pojmový systém fyziky ztrácejí fyzikální pojmy i principy svůj smysl i význam. Pokud však jde o entropii, tak ani tato mez zřejmě není nepřekročitelná – Claude Shannon přece zobecnil pojem entropie i do oborů za hranicí fyziky a fyzikálního myšlení. Nicméně pokud jde o fyziku, tak fyzikové považují princip rostoucí entropie za jeden z nejobecněji platných principů, obecnější než rozpínání vesmíru nebo samovolné shlukování hmoty vlivem gravitace a vytváření galaxií, planetárních soustav aj. systémů.[12] Je to podobně obecný princip jako zákon zachování energie.

A jestliže ve vesmíru platí zákon zachování energie, energie ve vesmíru nepřibývá ani neubývá a vesmír je tedy izolovaná soustava, tak ve vesmíru jako celku entropie nikdy neklesá, ale naopak neustále a nezadržitelně roste. A poroste tak dlouho, až se všechny teploty a tlaky ve vesmíru vyrovnají, až se z vesmíru vytratí jakékoli uspořádání nebo pravidelnost, možnost cokoli rozlišit od čehokoli jiného. Se zastavením růstu entropie se zastaví i tok času a nastane věčnost. Pak už nebudou žádné planety, hvězdy, komety ani žádné jiné věci, jen stejnorodá chaotická směs nicoty, která bývala hmotou – prostě konec světa.

Zatímco konec světa dokážu předpovědět nejen já, ale kdekdo (dokonce i fyzikové), *s počátkem* světa je to těžší: odkud a jak se objevil řád, rozlišitelnost věcí, která uvádí svět do pohybu provázeného nárůstem entropie? Kde se vzala ve vesmíru informace, tj. nedostatek entropie, informace, která umožňuje entropii, aby narůstala, aby čas tekl z minulosti do budoucnosti a aby se svět proměňoval, pohyboval, vyvíjel?

[12] s velkým třeskem může (ale nemusí) být trochu problém, protože v okamžiku velkého třesku ještě neplatily dnes známé fyzikální zákony – viz Singh: Velký třesk [**!!! odkaz !!!**]

Tady už překračujeme hranici tajemna, do kterého nevidí fyzikové a asi ani nikdo jiný z přirozených bytostí. Prostě jsou otázky, které sice dokážeme položit, ale odpovědět na ně nedokážeme.

Svět a jeho řád jsou nepravděpodobné, chaos je pravděpodobný. Řád se samovolně vytrácí ze světa. Proto:

Hmota je vlna nepravděpodobnosti valící se do nicoty.

Není zákon zachování informace. Informace může v hmotném světě samovolně pominout, a to nenávratně a bez náhrady. Pokud by měl platit zákon zachování informace, musela by se fyzikálně ztracená informace přesouvat do jiného světa mimo hmotu, prostor a čas. Sem patří např. představy o tom, že po smrti člověka se jeho nesmrtelná, věčná duše přestěhuje z hmotného světa do světa jiného. Ať už věříme nebo nevěříme na nesmrtelnost duše – každý z nás jak chce, umí a může – totožnost člověka zřejmě není spojena ani tak s některým orgánem nebo vůbec s tělesnou stránkou lidské bytosti, jako spíš s informací, která zahrnuje povahové rysy, životní zkušenost, znalosti, hodnotový systém a postoje, vazby na sociální a kulturní prostředí, tedy osobnost člověka (aniž bych tím jakkoli popíral významnou – ne-li rozhodující – roli dědičnosti, tělesnosti a ostatních materiálních vlivů na utváření všeho toho, co dělá člověka sebou samým).

Entropie je míra neuspořádanosti; pro entropii v informatice je charakteristické:

- měří se stejně jako průměrná informace
- entropie je informace chybějící nebo ztracená[13]
    - nevyužitá kapacita paměti nebo zprávy (čisté CD)
    - využitá, ale ztracená paměť (vyndané CD)
    - zapomenutá informace (RAM po výpadku napájení)

Shrnutí

- množství informace převráceně závisí na četnosti nebo pravděpodobnosti jevu, dá se sčítat a do značné míry nezávisí na hmotě nebo energii (to vše vyžaduje definice)
- informaci získáváme, přenášíme a uchováváme zejména pomocí hmoty a energie (signál, médium); přenos informace bez využití hmoty a energie (např. telepaticky) není technologicky zvládnut a možnost takového přenosu je sporná
- informace může samovolně pominout, a to nenávratně a bez náhrady (ve fyzikálním světě neplatí zákon zachování informace)
- informace souvisí také s cíli, se schopnostmi a s ochotou svého příjemce nebo uživatele...

Entropie (výprava do fyziky)

...bude doplněno...

[13] ve starší populární a odborné literatuře se informace také nazývala *negativní entropie* nebo *negentropie*

Ludwig Boltzmann byl fyzik a snažil se dělat co nejlepší fyziku. Když se ale podíváme na jeho nejslavnější formuli, vidíme, že fyzikální rozměr má jenom konstanta k – celá fyzika se smrskla do jediné konstanty, zatímco zbytek formule je bezrozměrná četnost jakéhosi jevu: ten jev vypovídá spíš něco o uspořádanosti nebo neuspořádanosti obecně[14] než o něčem, co by se dalo zařadit do fyziky. Na stará kolena chudák Boltzmann zjistil, že celý život vlastně dělal něco úplně jiného než svou milovanou fyziku, začal hledat útěchu ve filosofii, ale nenašel ji a zatrpkl na celý svět. Nevěděl, že na jeho práci navážou další fyzikové, kteří spočítají další statistiky, např. pro elektronový plyn ve vodiči nebo pro fotony ve světle. A už vůbec nemohl tušit, že svým výpočtem entropie položil základy pro teorii informace a tím i pro informatiku.

Retrospektivní shrnutí

- SHANNONOVO *množství informace* zobecňuje BOLTZMANNŮV pojem *entropie* i na jevy mimo hranice fyziky.
- MAXWELL, GIBBS a BOLTZMANN vyšli z CLAUSIOVA výpočtu *změny entropie při izotermických dějích v plynech,* ale výpočet zobecnili na libovolné termodynamické děje. Tohoto zobecnění dosáhl JOSIAH GIBBS tak, že nepopisoval plyn vcelku pomocí pouhé změny objemu, nýbrž uvažoval o pohybu molekul v plynu. Entropii plynu pak vyjádřil LUDWIG BOLTZMANN jako **statistiku poloh a hybností molekul** (které v souhrnu určují *objem, tlak* i *teplotu*). BOLTZMANNOVO statistické rozdělení počítá s pravděpodobnostmi jevů – jakých jevů? – totiž zda se molekula svou polohou a hybností nachází nebo nenachází v daném intervalu poloh a hybností.
- Souvislost mezi *polohami* molekul a *objemem* plynu je triviální a stejně zřejmá je i souvislost mezi *tlakem* plynu a *silou,* kterou narážejí molekuly na *plochu* stěn nádoby. Naproti tomu souvislost mezi *hybností* molekul a *teplotou* je složitější: *hybnost* odpovídá *kinetické energii* a souhrnná *kinetická energie* molekul určuje *vnitřní energii* plynu. Úměra mezi *vnitřní energií* a *teplotou* se pak názorně zjistí pomocí **kaskády jednotkových tepelných strojů** (pomocí které jsme zjišťovali také energetickou bilanci a účinnost *Carnotova cyklu*).
- Fyzikální rozměr BOLTZMANNOVĚ entropii dává *Boltzmannova konstanta,* která se spočítá jako poměr mezi *univerzální plynovou kontsantou* a *Avogadrovou konstantou.*
- Pro informatika *doby postindustriální* je zajímavé i to, že BOLTZMANNOVA fyzika překonává *moderní pozitivistické paradigma* devatenáctého století, typické pro *dobu industriální,* protože mluví **o jevech** a jejich pravděpodobnostech, a ne o měřitelných veličinách a evidentních *faktech.* Tím má daleko blíž k HUSSERLOVĚ *fenomenologii* než *k pozitivismu.*
- WALTHER NERNST objevil, že entropie látky při nulové teplotě je nulová. To umožnilo počítat s absolutními hodnotami, a ne jen se změnami entropie. Na podobném principu můžeme i v informatice

[14] ta uspořádanost a neuspořádanost se může týkat v podstatě čehokoli – není důležité, že Boltzmann počítal zrovna molekuly v ideálním plynu

mluvit *o informaci* nejen jako o přírůstku / úbytku znalosti, ale také absolutně jako o rozdílu mezi uspořádáním systému, které se nám jeví, a mezi rovnovážným stavem naprostého chaosu, který neobsahuje žádnou informaci (POZOR na rozdíl: naprostý chaos není stav s nulovou, nýbrž s maximální entropií).

- RUDOLF CLAUSIUS zavedl pojem *entropie* jako fyzikální veličinu – schopnost nebo neschopnost látky konat práci. Odvodil vztah *pro výpočet změny entropie při izotermických dějích,* ze kterého později BOLTZMANN odvodil svoji statistiku.
- CLAUSIUS zjistil, že *v rovnovážném stavu* nemohou probíhat žádné termodynamické děje – přišel na to, proč dosavadní výpočty založené *na rovnovážných stavech* a *vratných dějích* nemohou být přesné.
- CLAUSIUS také formuloval druhou větu termodynamickou, že totiž *teplo samovolně nepřechází z chladnějšího do teplejšího prostředí* (nebo jinými slovy: že **entropie izolované soustavy neklesá**). Na základě této věty pak kosmologové vysvětlují, proč **čas teče z minulosti do budoucnosti** (a nikdy naopak) a že **svět jednou zanikne v naprostém chaosu** při dosažení maximální entropie a rovnovážného stavu vesmíru.
- SADI CARNOT vysvětlil, jak parní stroj přeměňuje teplo na mechanickou práci. Na přelomu 18. a 19. století lidé sice už běžně používali parní stroje, ale fyzikální princip tepelného stroje vysvětlil až SADI CARNOT a tím položil základy *termodynamiky.* CARNOT už věděl i to, že samo teplo neurčuje, kolik práce vykoná tepelný stroj, ale že také záleží **na rozdílu mezi teplotou ohřívače a teplotou chladiče.** CARNOTOVI připadalo samozřejmé, že **teplo nemůže samo od sebe přecházet z chladnějšího prostředí do teplejšího,** a ve svých výpočtech tuto zásadu důsledně respektoval. Nicméně k tomu, aby se dala počítat práce (nebo naopak potřebné množství tepla), stále chyběla veličina, která by charakterizovala schopnost látky konat práci (a která by se dala např. integrovat).
- Konstanty: V daném množství *n* plynu se poměr pV/T nemění. To vyjadřuje *univerzální plynová konstanta R.* Počet částic v jednom molu látky je konstantní a nazývá se *Avogadrova konstanta* N_A.

Poděkování

Děkuji

- panu Mgr. Radku Štěpánovi za kontrolu kapitoly o entropii ve fyzice a za připomínky k této kapitole
- čtenářům portálů RVP a Česká škola za komentáře, náměty a připomínky v diskuzích k článkům věnovaným této učebnici
- panu PhDr. Ondřeji Neumajerovi, Ph.D. a ostatním pracovníkům portálu RVP za možnost průběžně publikovat a konzultovat na portálu

Literatura

ADAMS, Douglas. *Stopařův průvodce po galaxii.* **(!!! doplň !!!)**

FEYNMAN, Richard P.; Robert B. LEIGHTON; Matthew SANDS. *Feynmanovy přednášky z fyziky s řešenými příklady.* Havlíčkův Brod: Fragment, 2005. ISBN 80-7200-402-0.

HOFSTADTER, Douglas. *Gödel, Escher, Bach.* New York: Basic Books, 1979. ISBN 0-465-02656-7 (český překlad vydala nakladatelství Argo a Dokořán v roce 2012)

SINGH, Simon. *Velký třesk: nejdůležitější vědecký objev všech dob a proč o něm musíte vědět*. [z angl. orig. *The big bang* přel. Jiří Podolský a Martin Žofka]. Praha: Argo; Dokořán, 2007. ISBN 978-80-86569-62-8, ISBN 978-80-7203-894-7.

WIENER Norbert. *Kybernetika a společnost.* Praha: Nakladatelství ČSAV, 1963. (z originálu The Human Use of Human Beings. Cybernetics and Society, druhé doplněné vydání, Doubleday, New York 1954 přeložil Karel Berka, poprvé vyšlo v nakladatelství Houghton Mifflin, Boston, 1950). Str. 50.