Albert Pérez

Matematická teorie informace. I

Aplikace matematiky, Vol. 3 (1958), No. 1, 1–21

Persistent URL: `http://dml.cz/dmlcz/102599`

ČLÁNKY

MATEMATICKÁ TEORIE INFORMACE

ALBERT PEREZ

(Došlo dne 24. června 1957.) DT: 621.39.001:51

Po zavedení pojmů zprávy a entropie na základě úvah z thermodynamiky a v těsném spojení s pojmem rozlišitelnosti zavedeném pravděpodobnostním způsobem se podává obecná teorie informace (pro libovolnou abecedu), obsahující jako speciální případ diskretní teorii informace (pro konečnou abecedu). Zejména jsou v článku uvedeny zobecněné verse limitní věty McMillanovy a základního lemmatu Feinsteinova a pomocí nich se zkoumají otázky možnosti přenosu zdroje informace sdělovacím kanálem (věty Shannonova typu).

Obsah

[1]) Vyjde v příštím čísle tohoto časopisu.

Úvod

Tento článek obsahuje všeobecný přehled o současném stavu matematické teorie informace se zřetelem k aplikacím.

Pokud jde o t. zv. *diskretní* teorii informace (konečná abeceda) poskytne čtenáři zcela uspokojivý a přesný obraz přehledný článek [1] od A. Ja. CHINČINA a práce J. NEDOMY [2], obsahující jisté zajímavé myšlenky o pojmu diskretního sdělovacího kanálu a jeho kapacity.

Docela jiná je situace v *zobecněné* teorii informace (libovolná abeceda). Avšak v nejnovější době byly publikovány jisté práce [3], [4], [5], které představují vážné snahy o zobecnění výsledků diskretní teorie informace. Ačkoliv forma jejich podání není taková, aby dávala úplný obraz o jejich obsahu, můžeme nyní konstatovat, že jsou velmi podobné řadě výsledků sdělených na různých konferencích v ČSR v letech 1954, 1955 a 1956 [6], [7], [8].

V roce 1954 zavedl autor zobecněné pojmy *nejistoty*, *entropie* a *informace* a odvodil některé jejich základní vlastnosti. Jejich další studium umožnilo v roce 1955 nalézti jisté nové vlastnosti těchto pojmů a určit jejich vztah k *teorii martingalů* a *semimartingalů*, což jsou náhodné procesy speciálního typu (viz [23]). Počátkem roku 1956 již bylo možno zobecnit také *limitní větu* MCMILLANOVU [9] a [1], což dále umožnilo zobecnění *základní* FEINSTEINOVY *věty* [10], kterou Chinčin v [1] nazývá *základním lemmatem Feinsteinovým*. Speciálně toto zobecnění, které je podstatné pro teorii přenosu stacionárních zdrojů informace stacionárním sdělovacím kanálem, bylo umožněno použitím jisté *pravděpodobnostní koncepce pojmu rozlišitelnosti*. Pomocí pojmu rozlišitelnosti lze uvažovat a vyšetřovat diskretní a nediskretní případ s jednotného hlediska. Pokud je nám známo, právě toto sjednocení působilo autorům zabývajícím se obecnou teorií informace značné starosti. Nejobecnější verse základní věty Feinsteinovy v zahraničních publikacích je obsažena v práci ROSENBLAT-ROTA [5] a týká se nejvýše *spočetné* abecedy. Naproti tomu tvar Feinsteinovy věty uvedený v tomto článku platí pro *libovolnou* abecedu. Výsledky jsou obsaženy v publikaci *Transactions of the First Prague Conference on Information Theory, Statistical Decision Functions and Random Processes* (1957) v pracích [11], [12] a [13]. Tento článek se opírá hlavně o poslední tři odkazy, kde čtenář nalezne úplné důkazy všech formulovaných vět.

1. Teorie informace a rozlišitelnost

V tomto paragrafu se pokusíme osvětlit skutečný smysl a předmět teorie informace v užším smyslu, která je velmi mladým odvětvím vědy a v současné době se rychle rozvíjí. Mluvíme o teorii informace v užším smyslu, neboť v literatuře se často pod název teorie informace zahrnují všechny pravděpodobnostní problémy, které se týkají různých systémů sdělování a řízení.

Je mimo jakoukoliv pochybnost, že pravděpodobnostní hledisko dominuje v teorii informace, a to plným právem. Je tomu tak proto, že *zpráva a to co se s ní stane*, které jsou hlavními předměty této teorie, mají pravděpodobnostní charakter.

Podle směru načrtnutého N. WIENEREM [14], C. SHANNONEM [15] a jinými autory, kteří úspěšně pracují v tomto oboru, je uvedená koncepce teorie informace *věda, která se snaží proniknout do podstaty pojmu zprávy, při čemž abstrahuje od jejich různých fysikálních forem* (signál) *a snaží se jich využít k tomu, aby omezila rušivé vlivy* (šum), *které zprávy znehodnocují.*

Co je tedy zpráva?

Intuitivní smysl pojmu zpráva si každý uvědomuje. Pokusme se nyní vystihnout podstatný charakter tohoto pojmu.

Především jen tehdy má smysl mluvit o zprávě, když dva fysikální systémy jsou ve stavu vzájemného působení nebo interakce. (Při tom se nevylučuje živočich nebo dokonce člověk jako speciální případ takového fysikálního systému). Pro studium jistých aspektů tohoto vzájemného působení je pojem zprávy naprosto nutným a vzniká otázka, které jsou to aspekty.

Obecně každá interakce mezi dvěma fysikálními systémy se projevuje jednak vzájemnou výměnou energie, jednak změnami stavů těchto systémů. Jak známo, *energetická bilance* (podle prvního zákona thermodynamiky) nám poskytuje všeobecný obraz o jistých často důležitých aspektech interakce. Používá se tu pojmu *energie*, který byl získán abstrakcí ze všech možných fysikálních forem, jimiž může být realisován a který je jedním z nejdůležitějších pomocníků vědy. Energetická bilance popisuje jen některé aspekty interakce fysikálních systémů. Často se stává, že právě *entropická bilance* (podle druhého zákona thermodynamiky) spolu se svým dvojníkem *informační bilancí* jsou důležitější nebo alespoň stejně důležité jako bilance energetická. Tak je tomu v řadě problémů z thermodynamiky a zvláště v problémech sdělování a řízení. Pojmy *zpráva* resp. *entropie* a *informace* mají v entropické a informační bilanci stejnou úlohu jako pojem *energie* resp. *množství energie* v bilanci energetické.

Nyní můžeme naznačit jeden z podstatných rozdílů mezi pojmem energie a pojmem zprávy. Podle prvního zákona thermodynamiky jakákoliv transformace isolovaného fysikálního systému *zachovává jeho energii*. Naproti tomu podle druhého zákona thermodynamiky toliko t. zv. reversibilní transformace zachovávají entropii isolovaného fysikálního systému. *Každá jiná transformace entropii zvětšuje.* Ve statistické thermodynamice to znamená, že systém má tendenci se vyvíjet k takovým makroskopickým stavům, které odpovídají stále se zvětšujícím „objemům“ ve fázovém prostoru, až do okamžiku, kdy se dosáhne makroskopické rovnováhy. Jak známo, každý bod fázového prostoru představuje jeden mikroskopický stav uvažovaného systému a *systém má tendenci se vyvíjet k takovému makroskopickému stavu, jehož údaj nám čím dále* *tím méně dovoluje rozeznat mikroskopický stav systému.* To znamená, že „chaos“ vzrůstá.

Zatím co energie isolovaného fysikálního systému se zachovává, údaje o makroskopickém stavu systému uvažované jako zprávy o jeho mikroskopickém stavu se stávají čím dále tím chudšími, t. j. znehodnocují se v tom smyslu, že nám v průběhu vývoje systému čím dále tím méně dovolují rozlišit jeho mikroskopické stavy.

Pro takové znehodnocování zpráv je měřítkem vzrůst entropie systému, nebo jinými slovy informace o jeho mikroskopickém stavu, obsažená v údajích o jeho makroskopickém stavu, klesá podle toho, jak entropie vzrůstá v tom smyslu, že rozlišitelnost mikroskopických stavů stále klesá.

Druhý zákon thermodynamiky je zcela ve shodě s naší intuitivní představou, že žádná transformace přijaté zprávy nemůže zvětšit informaci, právě tak jako nemůže zvýšit rozlišitelnost.

V předešlých úvahách jsme předpokládali fysikální systém isolovaný a nikoliv v interakci s druhým takovým systémem. Je to snad v rozporu s naším předešlým tvrzením, že pojem zprávy má smysl jen při interakci fysikálních systémů?

Skutečně se zdá na prvý pohled, že tu jde o rozpor. Avšak když identifikujeme zprávu s údaji o makroskopickém stavu fysikálního systému, pak musíme nutně předpokládat, že před tím bylo provedeno měření tohoto makroskopického stavu, t. j. tento systém musel být přiveden do stavu interakce s jiným fysikálním systémem (měřicí přístroje). Toliko změnami stavu druhého systému se konkretisují údaje o stavu systému prvého. Věda s úspěchem používá takových abstrakcí. Tato diskuse má upoutat pozornost čtenáře na tu skutečnost, že informace, údaje, zprávy nemohou být nic jiného než výsledky pozorování, měření, tedy výsledky interakce daného systému a „pozorovatele“. Kdybychom nerespektovali tuto skutečnost, pak bychom snadno mohli dospět k absurdnímu závěru, že druhý zákon thermodynamiky může být porušen zásahem „inteligentní bytosti“, jako je Maxwellův demon [16], [17]. Nechceme v tomto článku udávat příklady, kdy neúplná entropická a informační bilance může vést k „porušení“ druhého zákona thermodynamiky. Avšak úplná bilance je vždy ve shodě s tímto zákonem a ani tomu nemůže být jinak. Obecně entropie celého systému, t. j. pozorovaného systému a pozorovatele vzrůstá v průběhu pozorování, jímž se získává informace, avšak tuto informaci lze opět využít pouze k *částečnému* kompensování vzrůstu entropie celého systému. Celková definitivní bilance je tedy vždy ve shodě s druhým zákonem thermodynamiky.

Shrneme nyní předešlé úvahy. Pojem zprávy je nutný ke studiu jistých aspektů vzájemného působení mezi dvěma fysikálními systémy, zcela obdobně jako je nutný pojem energie. Tyto aspekty se týkají věrnosti, s jakou jsou stavy pozorovaného systému odráženy nebo registrovány druhým systémem (pozorovatel). Tento druhý systém se může popřípadě přizpůsobovat resp. může reagovat na pozorovaný systém (zpětná vazba). Jinak řečeno jde o rozlišování stavů pozorovaného systému pozorovatelem, při čemž úlohy obou systémů se mohou vzájemně vyměnit.

Když na př. pozorovaný systém je plyn v rovnovážném stavu a pozorovatelem je indikátor tlaku, pak zřejmě tento indikátor zahrne v jednom svém stavu celou řadu mikroskopických stavů plynu, které dávají stejný tlak. Tento počet mikroskopických stavů plynu se ještě zvětší, když přihlédneme k přesnosti údajů indikátoru. Rozlišovací schopnost indikátoru se ještě dále zmenší, když tlak plynu se mění rychle ve srovnání s časovou konstantou, takže indikátor nemůže přesně sledovat tyto rychlé změny tlaku.

Analogická je situace, uvažujeme-li dvě spolu hovořící osoby, mezi nimiž je tedy akustická interakce. V tomto případě akustické vlny řízené mluvícím člověkem (pozorovaný systém) přicházejí do ucha naslouchajícího člověka (pozorovatel). Zanedbáme-li ostatní možnosti interakce, pak pozorovatel bude moci tím lépe svým sluchovým orgánem rozlišovat vysílané zprávy, čím lépe mu bude jazyk (kod) zpráv srozumitelnější, čímž se rozumí schopnost reagovat a přizpůsobovat se přijímaným signálům tím lépe, čím jeho časem nahromaděná zkušenost je bohatší.

2. Pravděpodobnostní koncepce pojmu rozlišitelnosti

Dříve než podáme pravděpodobnostní pojetí pojmu rozlišitelnosti je účelné alespoň naznačit, jak se tento pojem často chápe bez přímého použití pojmů z teorie pravděpodobnosti. V tomto směru je velmi poučná práce [18] A. N. Kolmogorova.

Když prostor X má konečnou mohutnost N_X, t. j. obsahuje konečný počet N_X bodů, pak stačí $\log N_X + 1$ údajů (ano, ne) k určení každého bodu tohoto prostoru. Proto můžeme považovat $I_X = \log N_X$ za míru pro množství informace, jež obsahuje určení jednoho bodu prostoru X.

V nekonečných prostorech je účelné a zcela přirozené zavést pojem *přibližné rozlišitelnosti* bodů. Nejnázornější je příklad metrického prostoru. Určení bodu x prostoru X s „přesností" ε lze chápat dvěma způsoby:

(a) bod $x \in X$ určíme pomocí bodu x^* z pevné množiny $X^* \subset X$ tak, že vzdálenost $\varrho(x, x^*)$ bodů x a x^* je menší než ε,

(b) bod $x \in X$ určíme pomocí pevného rozkladu $\boldsymbol{P}$ prostoru X na podmnožiny o průměru menším než ε tak, že udáme množinu z $\boldsymbol{P}$, která obsahuje bod x.

Podle těchto dvou hledisek zavádí Kolmogorov dvě funkce $N_X^a(\varepsilon)$ a $N_X^b(\varepsilon)$, které jsou definovány pro všechna kladná ε takto:

(a) funkce $N_X^a(\varepsilon)$ je definovaná jako minimum mohutností všech podmnožin

X^* metrického prostoru X, které mají tu vlastnost, že ke každému $x \in X$ lze přiřadit bod $x^* \in X^*$ tak, že $\varrho(x, x^*) < \varepsilon$.

(b) funkce $N_X^b(\varepsilon)$ je definovaná jako minimum mohutností všech rozkladů $\boldsymbol{P}$ metrického prostoru X na množiny o průměru menším než ε.

Nalezení vhodných metod pro přibližné rozlišování bodů prostoru má nepochybně základní důležitost v teorii informace. Bylo by však účelné uplatnit při hledání těchto metod *pravděpodobnostní hledisko*, které zcela přirozeně převládá v této teorii. Další výklad vychází z tohoto hlediska.

Nechť $(X, \mathfrak{X})$ je *měřitelný prostor*, t. j. dvojice, kterou tvoří neprázdný prostor X a *σ-algebra* $\mathfrak{X}$ podmnožin prostoru X. Jak známo, *σ-algebra* $\mathfrak{X}$ *je neprázdný systém podmnožin prostoru* X *obsahující s každou množinou její komplement a se spočetně mnoha svými množinami jejich sjednocení*. V řeči teorie pravděpodobnosti jsou množiny z $\mathfrak{X}$ *náhodnými jevy*. Reálnou funkci μ definovanou v $\mathfrak{X}$ nazýváme *pravděpodobnostní mírou*, když je nezáporná, σ-aditivní a splňuje podmínku $\mu(X) = 1$, při čemž *σ-aditivita* znamená, že pravděpodobnost uskutečnění alespoň jednoho ze spočetně mnoha navzájem se vylučujících (po dvou disjunktních) náhodných jevů je rovna součtu pravděpodobností těchto náhodných jevů. Trojice $(X, \mathfrak{X}, \mu)$ nazýváme *pravděpodobnostním polem*. Řekneme, že reálná funkce f definovaná v X je *$\mathfrak{X}$-měřitelná*, když pro každou borelovskou množinu M reálných čísel, t. j. vytvořenou jistým způsobem pomocí otevřených nebo uzavřených intervalů, množina $f^{-1}(M) = \{x : x \in X, f(x) \in M\}$ náleží do $\mathfrak{X}$. Je známo, že měřitelnost funkce f je ekvivalentní podmínce $\{x : x \in X, f(x) < c\} \in \mathfrak{X}$, která musí být splněna pro každé reálné číslo c. Abychom intuitivně pochopili smysl měřitelnosti, uvažujme jednoduchý speciální případ, kdy σ-algebra $\mathfrak{X}$ je *vytvořena* rozkladem prostoru X, t. j. $\mathfrak{X}$ je minimální σ-algebra podmnožin prostoru X, obsahující všechny množiny rozkladu. Funkce f je v tomto speciálním případě tehdy a jen tehdy $\mathfrak{X}$-měřitelná, když je konstantní na každé množině rozkladu prostoru X. Když tedy f nabývá více hodnot než je množin v rozkladu prostoru X, pak samozřejmě není možné, aby byla $\mathfrak{X}$-měřitelná. Čtenář, kterému takové pojmy jako je σ-algebra, míra, měřitelnost a příslušná symbolika nejsou běžné, najde velmi přístupné poučení v [19].

Nechť $(X, \mathfrak{X})$ je měřitelný prostor. Požadavek, aby byl přímo přístupný pozorování, klademe jako rovnocenný požadavku, aby příslušné náhodné jevy byly přímo pozorovatelné. Řekneme, že množina $\{x_\gamma : x_\gamma \in X, \gamma \in \Gamma\}$ je *přímo rozlišitelná*, když ke každému $\gamma \in \Gamma$ lze přiřadit množinu $E_\gamma \in \mathfrak{X}$ tak, že množiny E_γ jsou po dvou disjunktní a $x_\gamma \in E_\gamma$.

Předpokládejme nyní, že $(X, \mathfrak{X})$ a $(Y, \mathfrak{Y})$ jsou měřitelné prostory a každému bodu $x \in X$ je přiřazena pravděpodobnostní míra ν_x v $\mathfrak{Y}$ tak, že pro každou množinu $F \in \mathfrak{Y}$ je $\nu_x(F)$ (jako funkce proměnné x) $\mathfrak{X}$-měřitelná. Množinu $\{\nu_x : x \in X\}$ nazýváme *kanálem* se *vstupní* σ-algebrou $\mathfrak{X}$ a s *výstupní* σ-algebrou $\mathfrak{Y}$. Kanál budeme označovat trojicí $(\mathfrak{X}, \nu_x, \mathfrak{Y})$.

Řekneme, že měřitelný prostor $(X, \mathfrak{X})$ je *pozorovatelný kanálem* $(\mathfrak{X}, \nu_x, \mathfrak{Y})$, když měřitelný prostor $(Y, \mathfrak{Y})$ je přímo pozorovatelný.

Když $\varepsilon > 0$, pak řekneme, že množina $\{x_\gamma : x_\gamma \in X, \gamma \in \Gamma\}$ je *ε-rozlišitelná* kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$, když každému $\gamma \in \Gamma$ lze přiřadit neprázdnou množinu $F_\gamma \in \mathfrak{Y}$ tak, že množiny F_γ jsou po dvou disjunktní a platí

$$\nu_{x_\gamma}(F_\gamma) > 1 - \varepsilon \, . \tag{2.1}$$

Tato definice ε-rozlišitelnosti kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ byla zvolena při studiu Chinčinova důkazu základního lemmatu Feisteinova [1].

Poznamenejme, že přímá rozlišitelnost množiny $\{x_\gamma : x_\gamma \in X, \gamma \in \Gamma\}$ znamená totéž co ε-rozlišitelnost kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$, kde ν_x je charakteristická funkce χ_E množiny $E \in \mathfrak{X}$, t. j. $\chi_E(x) = 1$ nebo 0 podle toho, je-li $x \in X$ nebo $x \in E' = X - E$ a číslo $\varepsilon > 0$ je libovolně malé.

Z hlediska rozlišovací schopnosti lze kanál charakterisovat *mohutností ε-rozlišitelnosti*, což podle definice je supremum mohutností všech množin ε-rozlišitelných daným kanálem. Měříme-li *stupeň* (přesnost) *rozlišitelnosti* převratnou hodnotou čísla ε, pak *obecně vyššímu stupni rozlišitelnosti odpovídá menší mohutnost rozlišitelnosti.*

Řekneme, že kanál $(\mathfrak{Z}, \nu_z, \mathfrak{V})$ je *v kaskádě* s kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$, když výstupní měřitelný prostor $(V, \mathfrak{V})$ prvního kanálu se shoduje se vstupním měřitelným prostorem $(X, \mathfrak{X})$ kanálu druhého a zmíněná dvojice kanálů určuje nový kanál $(\mathfrak{Z}, \tilde{\nu}_z, \mathfrak{Y})$, kde pravděpodobnostní míry $\tilde{\nu}_z$ jsou definovány vztahem

$$\tilde{\nu}_z(E) = \int_X \nu_x(E) \, d\nu_z \, , \tag{2.2}$$

pro $E \in \mathfrak{Y}$, $z \in Z$, při čemž *integrál je Lebesgueův* (viz [19]).

Uvažujme nyní kanály $(\mathfrak{X}, \nu'_x, \mathfrak{X})$ a $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ v kaskádě a nechť $(\mathfrak{X}, \tilde{\nu}_x, \mathfrak{Y})$ je výsledný kanál. Pak zřejmě mohutnost ε-rozlišitelnosti kanálu $(\mathfrak{X}, \tilde{\nu}_x, \mathfrak{Y})$ je nejvýše rovna mohutnosti ε-rozlišitelnosti kanálu $(\mathfrak{X}, \nu_x, \mathfrak{Y})$. Skutečně, v nejlepším případě může být kanál $(\mathfrak{X}, \nu'_x, \mathfrak{X})$ *kanálem přímého pozorování*, t. j. takový, že jeho mohutnost ε-rozlišitelnosti se shoduje s mohutností jeho vstupního měřitelného prostoru pro každé $\varepsilon > 0$, při čemž *mohutností měřitelného prostoru* rozumíme supremum mohutností všech přímo rozlišitelných množin tohoto prostoru. Speciálním případem je ten, že $\nu_x(E) = \chi_E(x)$, kde χ_E je charakteristická funkce množiny $E \in \mathfrak{X}$. V tomto případě ovšem jsou mohutnosti obou kanálů stejné.

Pokusme se nyní interpretovat právě zavedené definice v řeči použité v předcházejícím paragrafu.

Měřitelný prostor $(X, \mathfrak{X})$ nechť představuje pozorovaný fysikální systém. Body $x \in X$ nechť představují mikroskopické stavy systému a každá množina $E \in \mathfrak{X}$ nechť představuje údaj (zprávu) o daném fysikálním systému. Na př. údaj o energii nebo tlaku a pod. plynu vymezuje množiny jeho mikroskopických stavů. Volba σ-algebry $\mathfrak{X}$ závisí na tom, jak podrobně chceme daný fysikální systém zkoumat. Když na př. se zajímáme jen o makroskopický stav systému, pak je zbytečné a někdy dokonce zcela neúčelné snažit se o rozlišování každého mikroskopického stavu zvlášť, t. j. volit σ-algebru $\mathfrak{X}$ tak, aby obsahovala na př. všechny jednobodové množiny. Je tedy účelné volit strukturu σ-algebry podle charakteru zkoumaných jevů a procesů.

Z hlediska teorie přenosu zpráv měřitelný prostor $(X, \mathfrak{X})$ lze interpretovat jako prostor všech možných zpráv se σ-algebrou jistých množin zpráv.

Avšak fysikální systém není vždy pozorován přímo, resp. ani není přístupný přímému pozorování. Jinak řečeno, když měřitelný prostor $(X, \mathfrak{X})$ představuje pozorovaný fysikální systém a měřitelný prostor $(Y, \mathfrak{Y})$ pozorovatele, pak existence různých rušivých vlivů náhodného charakteru, které mají svůj původ jednak v samotných systémech, které jsou v interakci, jednak v prostředí mezi oběma systémy (thermický šum, šum způsobený nespojitým průtokem elektrického proudu po elektronech a pod.) způsobuje, že interakce mezi systémy nelze vyjádřit pevnou transformací mezi $(X, \mathfrak{X})$ a $(Y, \mathfrak{Y})$, nýbrž transformací náhodnou. To znamená, že každému bodu $x \in X$ neodpovídá jeden bod $y \in Y$, nýbrž pravděpodobnostní míra ν_x v $\mathfrak{Y}$. Tím je vyjádřeno, že pro daný bod x může nastat každý z náhodných jevů z $\mathfrak{Y}$, avšak s pravděpodobností, která je *podmíněna* bodem x. Nazýváme proto $\nu_x(F)$ *podmíněnou pravděpodobností* náhodného jevu $F \in \mathfrak{Y}$ za předpokladu, že systém vyjádřený měřitelným prostorem $(X, \mathfrak{X})$ byl ve stavu $x \in X$. Zavedení pojmu kanálu vyjadřuje právě tuto skutečnost. Tak měřitelný prostor $(X, \mathfrak{X})$ není pozorován přímo, nýbrž kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$. Tento kanál může ovšem být ve speciálním případě kanálem přímého pozorování.

Přejděme nyní k pojmu rozlišitelnosti. Pozorovatel rozlišuje stavy pozorovaného systému $(X, \mathfrak{X})$ kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$, t. j. pozorovacím kanálem. Jelikož každému bodu $x \in X$ je přiřazena pravděpodobnostní míra ν_x v $\mathfrak{Y}$, je účelné zavést pojem ε-rozlišitelnosti jako zobecnění přímé rozlišitelnosti. Když pozorovatel zaznamenal stav $y \in Y$ a $y \in F_\gamma$, t. j. jako výsledek pozorování byla realisována množina $F_\gamma \in \mathfrak{Y}$, pak podle vztahu (2.1) a pro $\varepsilon < \frac{1}{2}$ (což budeme vždy předpokládat), je účelné zvolit jako odhad stavu systému $(X, \mathfrak{X})$ bod x_γ a tento odhad je tím přijatelnější, čím je ε menší.

Poznamenejme, že když $\varepsilon < \frac{1}{2}$, pak množiny

$$E_\gamma = \{x : \nu_x(F_\gamma) > 1 - \varepsilon\} , \tag{2.3}$$

kde γ probíhá množinu Γ, jsou po dvou disjunktní a podle definice kanálu patří všechny do σ-algebry $\mathfrak{X}$. Pozorovatel tedy může místo množiny $\{x_\gamma : x_\gamma \in \in X, \gamma \in \Gamma\}$ uvažovat systém $\{E_\gamma : E_\gamma \in \mathfrak{X}, \gamma \in \Gamma\}$, neboť stačí zvolit v každé z množin E_γ přesně jeden bod, abychom dostali ε-rozlišitelnou množinu podobně jako je původní množina $\{x_\gamma : x_\gamma \in X, \gamma \in \Gamma\}$.

Jak vidíme, pojem rozlišitelnosti má dvě stránky, a to *kvalitativní,* určenou číslem ε (stupeň rozlišitelnosti) a *kvantitativní,* měřenou mohutností množiny bodů rozlišitelných s přesností ε. Pojem mohutnosti ε-rozlišitelnosti pozorovacího kanálu byl zaveden právě proto, aby se vyjádřila jeho rozlišovací schopnost.

Uvažujme nyní kanál $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ a předpokládejme, že je dána pravděpodobnostní míra μ v σ-algebře $\mathfrak{X}$. V řeči teorie informace to znamená, že *zdroj informace* (pravděpodobnostní pole) $(X, \mathfrak{X}, \mu)$ *je přímo připojen na vstup kanálu* $(\mathfrak{X}, \nu_x, \mathfrak{Y})$, nebo že *zdroj informace je pozorován tímto kanálem,* t. j. *pozorovacím kanálem.*

Identifikace zdroje informace s pravděpodobnostním polem je oprávněná, kdykoliv výskyt různých stavů fysikálního systému (výskyt různých signálů) má pravděpodobnostní charakter. Tak je tomu u všech fysikálních systémů, které tvoří předmět statistické thermodynamiky, u systémů sdělování zpráv a pod. Tak na př. v dlouhém textu mají relativní četnosti výskytu různých písmen abecedy jistou charakteristickou stabilitu. Samozřejmě hluboké důvody pro tuto stabilitu relativních četností je třeba hledat v samotné podstatě příslušných fysikálních systémů (viz na př. [20]).

Ze všech pozorovacích kanálů zdroje informace $(X, \mathfrak{X}, \mu)$ lze vybrat jeden, který je nejlépe přizpůsoben možnostem pozorovatele nebo adresáta (člověk, stroj atd.) a jeho vlastní podstatě. Takový privilegovaný pozorovací kanál se vždycky nutně zařazuje mezi měřitelný prostor $(X, \mathfrak{X})$ zdroje a měřitelný prostor $(Y, \mathfrak{Y})$ přímo pozorovatelný. Tento kanál $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ nazýváme *kanálem příjemce zpráv.*

Kanál příjemce zpráv je jen ve výjimečných případech kanálem přímého pozorování. Daleko nejčastěji má kanál příjemce zpráv rozlišovací schopnost (mohutnost) podstatně menší než je mohutnost pozorovaného měřitelného prostoru. Uveďme alespoň jeden příklad. Bylo zjištěno, že lidské ucho nemůže rozlišit více než 250 hladin intensity zvuku vzdálených od sebe o půl decibelu a něco analogického platí o rozlišování akustických frekvencí.

Značný počet výzkumů v různých zemích (viz na př. [21]) o smyslovém vnímání a o srozumitelnosti z hlediska teorie informace je možno si vysvětlit jejich velkou důležitost pro aplikace. Avšak jsme přesvědčeni, že tyto výzkumy by byly mnohem plodnější, kdyby byly prováděny s plným vědomím té skutečnosti, že to co se snažíme určit není vlastně nic jiného, než kanál příjemce zpráv.

V dalším, při zkoumání možnosti přenosu zdroje informace sdělovacím kanálem, budeme dále rozvíjet pravděpodobnostní koncepci pojmu rozlišitelnosti, zavedenou v tomto paragrafu. V následujícím paragrafu zavedeme pojem entropie jako automatický důsledek té skutečnosti, že opakováním pozorování se zvětšuje rozlišovací schopnost.

3. Opakování a stupeň rozlišitelnosti. Pojem entropie

Intuitivně usuzujeme, že opakování pozorování zlepšuje rozlišitelnost. Budeme tento zjev studovat matematicky pomocí pojmů zavedených v předcházejícím paragrafu.

Předpokládejme, že pozorovaný měřitelný prostor je $(X, \mathfrak{X})$, kde $X = \{1, 2\}$ a $\mathfrak{X} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$, t. j. pozorovaný systém má dva možné stavy. Nechť dále $(Y, \mathfrak{Y})$ je měřitelný prostor pozorovatele, který odpovídá jednomu pokusu, t. j. jednomu pozorování. Potom měřitelný prostor odpovídající počtu n pokusů (n pozorování) je n-násobný kartézský součin $(Y \times Y \times \ldots \times Y, \mathfrak{Y} \times \mathfrak{Y} \times \ldots \times \mathfrak{Y}) = (Y^n, \mathfrak{Y}^n)$ (viz podrobnější vysvětlení v [19]).

Nechť $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ je pozorovací kanál odpovídající jednomu pokusu. Potom pozorovací kanál, který odpovídá n *nezávislým* (ve smyslu teorie pravděpodobnosti) pokusům je $(\mathfrak{X}, \nu_x^n, \mathfrak{Y}^n)$, kde ν_x^n značí pravděpodobnostní míru v $\mathfrak{Y}^n$ pro $x = 1, 2$, indukovanou pravděpodobnostní mírou ν_x v $\mathfrak{Y}$.

Označíme-li μ_x pravděpodobnostní míru v měřitelném prostoru $(Y^N, \mathfrak{Y}^N) = (Z, \mathfrak{Z})$ nekonečných posloupností nezávislých pokusů indukovanou pravděpodobnostní mírou ν_x v $\mathfrak{Y}$ pro $x = 1, 2$, pak místo $(\mathfrak{X}, \nu_x^n, \mathfrak{Y}^n)$ můžeme psát $(\mathfrak{X}, \mu_x, \mathfrak{Y}^n)$ pro $n = 1, 2, 3, \ldots$, neboť ν_x^n lze identifikovat s pravděpodobnostní mírou μ_x pro $x = 1, 2$ na σ-algebře $\mathfrak{Y}^n$, kterou zřejmě můžeme identifikovat se σ-algebrou všech n-rozměrných válců σ-algebry $\mathfrak{Z}$ bez obavy z nedorozumění. Můžeme tedy psát $\mathfrak{Y}^n \subset \mathfrak{Z}$.

Problém je nyní zjistit, jak vzrůstá asymptoticky stupeň rozlišitelnosti posloupnosti kanálů $(\mathfrak{X}, \mu_x, \mathfrak{Y}^n)$ pro mohutnost rozlišitelnosti rovnou *dvěma*, odpovídající dvěma bodům prostoru X, které je třeba rozlišit.

Budeme nyní předpokládat, že pravděpodobnostní míra ν_1 je absolutně spojitá vzhledem k ν_2, z čehož plyne, že μ_1 je absolutně spojitá vzhledem k μ_2 na σ-algebře $\mathfrak{Y}^n$ a budeme psát $\mu_1 \ll \mu_2$ $(\mathfrak{Y}^n)$ pro $n = 1, 2, 3, \ldots$ Připomeňme, že ν_1 je *absolutně spojitá* vzhledem k ν_2, když každá množina z $\mathfrak{Y}$, která je ν_2-míry nula, je také ν_1-míry nula. Říkáme také, že ν_2 *dominuje* ν_1.

Je známo (viz na př. [19]), že potom existuje nezáporná funkce $f(y)$ měřitelná vzhledem k σ-algebře $\mathfrak{Y}$ a splňující vztah

$$\nu_1(E) = \int_E f(y)\, d\nu_2$$

pro všechny $E \in \mathfrak{Y}$. Funkci f nazýváme Radon-Nikodymovou *hustotou* míry ν_1 podle ν_2. Je jednoznačně určena až na množinu ν_2-míry nula, což budeme stručně značit $[\nu_2]$.

Označme $z = (y_1, y_2, y_3, \ldots)$ nekonečnou posloupnost pozorování a nechť $f_n(z)$ je hustota pravděpodobnostní míry μ_1 podle μ_2 na σ-algebře $\mathfrak{Y}^n$ pro $n = 1, 2, 3, \ldots$. Z předpokladu nezávislosti pokusů plyne

$$f_n(z) = f_n(y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} f(y_i) \,. \tag{3.1}$$

Podle klasické ergodické věty (platné pro posloupnosti nezávislých náhodných proměnných se stejnými distribucemi) je

$$\lim_{n\to\infty} \frac{1}{n} \log f_n(z) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log f(y_i) = H \tag{3.2}$$

podle μ_1-středu za předpokladu, že střední hodnota funkce $\log f(y)$ (*zobecněná nejistota* [11]), t. j.

$$\int_Y \log f(y) \, d\nu_1(y) = \int_Y f(y) \log f(y) \, d\nu_2(y) = H \tag{3.3}$$

existuje.

Připomeňme, že posloupnost $g_1, g_2, g_3, \ldots$, μ-integrovatelných funkcí konverguje podle středu k μ-integrovatelné funkci g, když $\int |g_n - g| \, d\mu \to 0$ pro $n \to \infty$.

Speciálně je

$$\lim_{n\to\infty} \frac{1}{n} \int \log f_n \, d\mu_1 = H . \tag{3.4}$$

Jak známo z *konvergence podle středu plyne konvergence podle pravděpodobnosti*. Tedy když ε a η jsou libovolně malá kladná čísla a

$$F_1 = \{z : e^{n(H-\eta)} \leqq f_n(z) \leqq e^{n(H+\eta)}\} , \tag{3.5}$$

(kde e je báze logaritmů, avšak nikoliv nutně přirozených), pak podle (3.2) existuje celé kladné číslo $n_0(\varepsilon, \eta)$ tak, že pro $n \geqq n_0(\varepsilon, \eta)$ platí

$$\mu_1(F_1) \geqq 1 - \varepsilon . \tag{3.6}$$

Integrujeme-li podle μ_2 přes množinu F_1 nerovnosti v (3.5), dostáváme

$$\mu_2(F_1) \, e^{n(H-\eta)} \leqq \int_{F_1} f_n(z) \, d\mu_2 = \mu_1(F_1) \leqq \mu_2(F_1) \, e^{n(H+\eta)} \tag{3.7}$$

a tedy podle (3.6)

$$(1 - \varepsilon) \, e^{-(H+\eta)} \leqq \mu_2(F_1) \leqq e^{-(H-\eta)} . \tag{3.8}$$

První z nerovností (3.8) říká speciálně, že $H \geqq 0$, kteroužto vlastnost lze také odvodit přímo (viz [11], věta 5).

Skutečně kdyby bylo $H < 0$, pak pro $\eta < |H|$ a dostatečně velké n by levá strana první nerovnosti byla větší než pravá, která je nejvýše rovna jedné, což je spor. V našem případě předpokládáme $\nu_1 \neq \nu_2$ a tedy je $H > 0$.

Položíme-li $F_2 = Z - F_1$, pak podle druhé z nerovností (3.8) pro dostatečně velké n, které může být tím menší čím větší je H, platí nerovnost

$$\mu_2(F_2) \geqq 1 - e^{-n(H-\eta)} . \tag{3.9}$$

Tím jsme dostali rozklad měřitelného prostoru $(Y^n, \mathfrak{Y}^n)$ posloupností $y^n = (y_1, y_2, \ldots, y_n)$ ve dvě množiny F_1^n a F_2^n tak, že vzhledem k nim (viz definice ε-rozlišitelné množiny) jsou oba body prostoru X kanálem $(\mathfrak{X}, \mu_x, \mathfrak{Y}^n)$ pro dostatečně velké n ε-rozlišitelné. *Tento počet n nutných pokusů je podle*

(3.9) *tím menší čím H je větší*, takže H v jistém smyslu měří „snadnost", s jakou lze rozlišit body 1 a 2 prostoru X (resp. ν_1 a ν_2 nebo μ_1 a μ_2, což v podstatě znamená totéž). Vidíme tedy, že pro vztah mezi počtem opakování a rozlišitelností má číslo H důležitou úlohu. Až na znaménko je H t. zv. *zobecněná entropie* pravděpodobnostní míry ν_1 podle ν_2, která v [11] byla zavedena axiomaticky pomocí pojmu nejistoty: $H_{\nu_2}(\nu_1, \mathfrak{Y}) = - H$.

Než vyjmenujeme některé další důležité vlastnosti zobecněné entropie, poznamenejme, že entropie obvykle definovaná v případě *konečného* pravděpodobnostního pole, určeného konečným systémem $\{E_1, E_2, \ldots, E_n\}$ n náhodných jevů po dvou disjunktních, jejichž pravděpodobnosti $p_1, p_2, \ldots, p_n$ dávají součet jedna, je speciálním případem. Skutečně tato entropie je určena formulí

$$H(p_1, p_2, \ldots, p_n) = - \sum_{i=1}^{n} p_i \log p_i \,. \tag{3.10}$$

Když v (3.3) míra ν_1 je určena pomocí pravděpodobností $p_1, p_2, \ldots, p_n$ a dominující míra ν_2 (podle níž je určena hustota f) je taková, že každému z náhodných jevů $E_1, E_2, \ldots, E_n$ přiřazuje pravděpodobnost $\frac{1}{n}$, pak zobecněná entropie je

$$H_{\nu_2}(\nu_1) = - \int \log f \, d\nu_2 = - \sum_{i=1}^{n} p_i \log (np_i) = - \sum_{i=1}^{n} p_i \log p_i - \log n \tag{3.11}$$

a vidíme, že až na konstantu $\log n$ se shoduje se zobecněnou entropii (3.10).

Ostatně tuto konstantu lze úplně vyloučit, když netrváme na tom, aby ν_2 byla pravděpodobnostní míra, nýbrž žádáme, aby přiřazovala každé množině $E_1, E_2, \ldots, E_n$ míru jedna.

Jak se snadno přesvědčíme, nabývá obyčejná entropie (3.10) svého maxima pro $p_1 = p_2 = \ldots = p_n = \frac{1}{n}$, t. j. pro takovou míru ν_1, která každému náhodnému jevu $E_1, E_2, \ldots, E_n$ přiřazuje pravděpodobnost $\frac{1}{n}$, což znamená, že $\nu_1 = \nu_2$. Jak uvidíme v následujícím paragrafu je podmínka $\nu_1 = \nu_2$ nutná a zároveň postačující k tomu, aby zobecněná entropie byla nulová. Ve všech ostatních případech je záporná (viz speciálně (3.11)).

4. Vlastnosti zobecněné entropie a pojem informace

Pro důkaz vět, které budou v dalším formulovány, se odvoláváme na [11].

Věta 4.1. *Zobecněná entropie $H_\lambda(\mu, \mathfrak{X})$ pravděpodobnostní míry μ podle pravděpodobnostní míry λ, která dominuje μ na σ-algebře $\mathfrak{X}$* (t. j. $\mu \ll \lambda\, (\mathfrak{X})$) *splňuje nerovnost*

$$H_\lambda(\mu, \mathfrak{X}) \leqq 0 \,, \tag{4.1}$$

při čemž rovnost platí tehdy a jen tehdy, když μ a λ jsou identické na $\mathfrak{X}$ (t. j. $\mu = \lambda(\mathfrak{X})$).

Věta 4.2. *Když $\mu \ll \lambda(\mathfrak{X})$ a $\mathfrak{X}_1, \mathfrak{X}_2, \mathfrak{X}_3, \ldots$ je neklesající posloupnost σ-podalgeber σ-algebry $\mathfrak{X}$, t. j. $\mathfrak{X}_1 \subset \mathfrak{X}_2 \subset \ldots \subset \mathfrak{X}$ pak posloupnost pžíslušných zobecněných entropií je nerostoucí,* t. j.

$$H_\lambda(\mu, \mathfrak{X}_1) \geqq H_\lambda(\mu, \mathfrak{X}_2) \geqq \ldots \geqq H_\lambda(\mu, \mathfrak{X}) , \tag{4.2}$$

a rovnost mezi dvěma za sebou následujícími členy, t. j. $H_\lambda(\mu, \mathfrak{X}_n) = H_\lambda(\mu, \mathfrak{X}_{n+1})$ *platí tehdy a jen tehdy, když σ-algebra $\mathfrak{X}_n$ je suficientní* [24] *pro množinu měr $\{\mu, \lambda\}$ na $\mathfrak{X}_{n+1}$, jinak řečeno rovnost platí tehdy a jen tehdy, když existuje hustota míry, μ podle λ měřitelná vzhledem $\mathfrak{X}_{n+1}$, která je také měřitelná vzhledem k menší σ-algebře $\mathfrak{X}_n$.*

Právě uvedený výsledek lze také vyjádřit v řeči transformací. Nechť tedy T je *měřitelná transformace* měřitelného prostoru $(X, \mathfrak{X})$ do měřitelného prostoru $(Y, \mathfrak{Y})$, t. j. taková transformace prostoru X do prostoru Y, že množina $T^{-1}(F) = \{x : T(x) \in F\}$ patří do σ-algebry $\mathfrak{X}$ pro všechny množiny F ze σ-algebry $\mathfrak{Y}$, jinak řečeno σ-algebra $T^{-1}(\mathfrak{Y}) = \{T^{-1}(F) : F \in \mathfrak{Y}\} = \mathfrak{X}'$ je částí σ-algebry $\mathfrak{X}$. Nechť dále μT^{-1} resp. λT^{-1} je míra indukovaná v σ-algebře $\mathfrak{Y}$ mírou μ resp. λ, t. j. $\mu T^{-1}(F) = \mu(T^{-1}(F))$ resp. $\lambda T^{-1}(F) = \lambda(T^{-1}(F))$ pro $F \in \mathfrak{Y}$. Z předešlé věty dostaneme ihned tento důsledek:

Korolár 4.1. *Žádná měřitelná transformace T nemůže zmenšit entropii,* t. j.

$$H_{\lambda T^{-1}}(\mu T^{-1}, \mathfrak{Y}) \geqq H_\lambda(\mu, \mathfrak{X}) , \tag{4.3}$$

při čemž rovnost platí tehdy a jen tehdy, když transformace T je suficientní pro množinu měr $\{\mu, \lambda\}$, jinak řečeno, když σ-algebra $T^{-1}(\mathfrak{Y}) = \mathfrak{X}'$ je suficientní [24] *pro $\{\mu, \lambda\}$.*

Dříve než budeme pokračovat, pokusíme se interpretovat výsledky. Především je zcela jasné, že každá měřitelná transformace má tendenci zmenšovat „jemnost" struktury jevů, má tendenci potlačovat „detaily" pozorovatelné před použitím transformace. To je vyjádřeno „ochuzením" původní σ-algebry, která se stane „hrubší" ($\mathfrak{X}' \subset \mathfrak{X}$). Avšak čím je systém pozorovatelných jevů chudší, tím je obtížnější rozlišování mezi dvěma pravděpodobnostními mírami definovanými na σ-algebře obsahující tento systém. Toliko ve speciálních případech (když uvažovaná σ-algebra resp. příslušná transformace je suficientní pro uvažovanou množinu pravděpodobnostních měr) je tato rozlišitelnost stejně obtížná jako před transformací a nestane se obtížnější, jako je tomu v obecném případě.

Toto konstatování je v naprosté shodě s výsledky předcházejícího paragrafu, kde čím více se σ-algebra zvětšovala (zvětšováním počtu n nezávislých pokusů, zkrátka opakováním) tím snadnější se stávalo rozlišování mezi mírami μ_1 a μ_2, t. j. mezi body 1 a 2. Na druhé straně jsme zjistili při této příležitosti, že pro dané n bylo rozlišování tím snadnější, čím větší bylo H (podle (3.3)), t. j. čím menší byla zobecněná entropie $H_{\nu_2}(\nu_1, \mathfrak{Y})$ míry ν_1 podle ν_2. S ohledem na to, že v (3.9) vystupuje součin nH a že až na znaménko tento součin je roven zobecněné entropii $H_{\mu_2}(\mu_1, \mathfrak{Y}^n)$ míry μ_1 podle μ_2 na σ-algebře $\mathfrak{Y}^n$, tedy uvedené dva výsledky jsou obsaženy v tvrzení, že *čím větší je zobecněná entropie, tím obtížnější se stane rozlišování mezi body 1 a 2.*

Z tohoto hlediska je zřejmé, že zobecněná entropie jako míra obtížnosti rozlišování nutně klesá, když příslušná σ-algebra se zvětšuje (věta 4.2) anebo, což je totéž, musí se po transformaci zvětšit (korolár 4.1).

Můžeme si položit otázku, v jakém vztahu jsou předešlé výsledky s poznámkami v paragrafu 1 o druhém zákonu thermodynamiky, podle kterého entropie isolovaného fysikálního systému se zachovává jen pro t. zv. reversibilní transformace a jinak vždycky stoupá, t. j. podle interpretace, kterou dává statistická thermodynamika, že systém má tendenci se vyvíjet k takovým makroskopickým stavům, jež odpovídají čím dále tím větším ,,objemům" ve fázovém prostoru (až do okamžiku dosažení rovnováhy), takže údaje o těchto stavech nám čím dále tím méně dovolují rozlišit příslušný mikroskopický stav.

Abychom mohli odpovědět na tuto otázku, je třeba napřed precisovat, co znamená ,,údaj o makroskopickém stavu". S hlediska statistické thermodynamiky to znamená, že zvolíme jistou pravděpodobnostní míru ve fázovém prostoru, jehož body představují jdenotlivé mikroskopické stavy uvažovaného systému. Když fázový prostor, jak tomu často bývá, je euklidovský prostor s jistým počtem rozměrů, pak uvažovaná σ-algebra je σ-algebra borelovských množin (obsahující speciálně všechny množiny, které mají ,,objem" v obyčejném geometrickém smyslu). Vyjdeme-li z pohybových rovnic mechaniky, pak zjistíme, že pro konservativní systém každá měřitelná množina (která má objem) fázového prostoru zachovává Lebesgueovu míru (objem) vzhledem k transformacím mechaniky (které zobrazují fázový prostor do sebe), zkrátka tyto transformace zachovávají Lebesgueovu míru.

Na druhé straně rovnice mechaniky nám umožňují z počátečního makroskopického stavu (z počáteční pravděpodobnostní míry) vypočítat makroskopický stav v každém pozdějším okamžiku, takže údaj (zpráva) o počátečním makroskopickém stavu lze považovat za rovnocenný údajům o všech pozdějších makroskopických stavech. Za předpokladu, že počáteční pravděpodobnostní míra není singulární, nýbrž absolutně spojitá vzhledem k Lebesgueově míře, je snaha ukázat, že isolovaný systém se vyvíjí a přibližuje se k rovnovážnému stavu, při kterém pravděpodobnostní míra (charakterisující tento stav) se shoduje s mírou Lebesgueovou. Právě toto chceme vyjádřit tvrzením, že systém se vyvíjí k makroskopickým stavům odpovídajícím ,,objemům" ve fázovém prostoru, které jsou čím dále tím větší až do okamžiku dosažení rovnováhy.

Druhý zákon thermodynamiky vyjadřuje tuto skutečnost pomocí thermodynamické entropie, která matematicky není ničím jiným než právě speciální případ zobecněné entropie, při čemž dominující míra je v našem případě míra Lebesgueova. Volba této míry jakožto míry dominující je pro rozsáhlou kategorii fysikálních systémů diktována přírodními zákony (připomeňme speciálně její invarianci vzhledem k transformacím mechaniky, jakož i tendenci fysikálních systémů vyvíjet se k takovým makroskopickým stavům, odpovídajícím čím dále tím více Lebesgueově míře v jejich fázovém prostoru, jinými slovy k makroskopickým stavům čím dále tím většího chaosu). Avšak jakmile jsme jednou zvolili dominující míru, pak věta 4.1 nám zajišťuje, že zobecněná entropie nabude maxima, když a jen když dominovaná pravděpodobnostní míra se shoduje s mírou dominující (po případě až na multiplikativní konstantu). Řekneme-li, že dominující míra je invariantní vzhledem k uvažovaným transformacím, pak to neznamená nic jiného, než že makroskopická rovnováha je dosažena současně s maximem zobecněné entropie právě vzhledem k této dominující míře.

Obecně, časté tvrzení ve statistické thermodynamice, že *rovnováhy se dosáhne současně s maximem entropie, předpokládá, že tato entropie je počítána vzhledem k vhodné dominující míře, jejichž volba je diktována přírodními zákony.*

Že vzrůst entropie znamená v thermodynamice přiblížení se makroskopického stavu systému ke svému rovnovážnému stavu, to snadno vyplývá z toho, co jsme konstatovali dříve, že totiž čím větší je entropie, tím méně je rozlišitelná dominující míra od míry dominované, t. j. tím obtížnější je rozlišování mezi oběma odpovídajícími makroskopickými stavy.

Není naším úkolem, abychom tu dále rozváděli naše úvahy zvláště pokud jde o problémy spojené s důkazem druhého zákona thermodynamiky na základě jistých hypotéz o obecném charakteru elementárních procesů v makroskopickém systému (princip detailní bilance, atd.).

Zdá se nám však, že je na místě poznámka o obyčejné entropii v konečném případě (viz (3.10)). V § 3 jsme konstatovali, že je to zobecněná entropie vzhledem k rovnoměrnému rozložení pravděpodobnosti. Z tohoto hlediska je tvrzení, že obyčejná entropie nabývá svého maxima pro rovnoměrné rozložení pravděpodobnosti, ve shodě s větou 4.1. Při této příležitosti je poučné konstatovat, že jeden z postulátů na kterém Chinčin [22] založil svoji větu o jednoznačnosti entropie zavedené jako míra střední nejistoty, pozůstává právě v předpokladu, že tato míra nejistoty nabývá svého maxima pro rovnoměrné rozložení pravděpodobnosti.

Mezi základními vlastnostmi entropie zaujímá význačné místo *aditivita,* která v axiomatické definici entropie tvoří základní postulát (přímá definice v konečném případě viz [22], definice pomocí pojmu nejistoty v obecném případě viz [11]). Setkali jsme se s použitím aditivity na začátku tohoto paragrafu, kde zobecněné entropie $H_{\nu_2}(\nu_1, \mathfrak{Y})$ a $H_{\mu_2}(\mu_1, \mathfrak{Y}^n)$ pro $n = 1, 2, 3, \ldots$, splňují vztah $H_{\mu_2}(\mu_1, \mathfrak{Y}^n) = nH_{\nu_2}(\nu_1, \mathfrak{Y})$, což je zřejmé z logaritmické formy entropie a z formule (3.1), jejíž platnost plyne z předpokladu nezávislosti pokusů.

V obecnějším případě závislosti se aditivita vyjádří pomocí pojmu *střední podmíněné entropie*, jejíž obvyklou definici pro konečný případ lze najít na př. v citovaném Chinčinově článku [22]. Tady jenom stručně naznačíme, jak se tento pojem zavede v obecném případě. Pokud jde o podrobnosti, odkazujeme na práci [11].

Použijeme terminologie zavedené v § 2 a budeme značit $(X, \mathfrak{X}, \mu)$ zdroj informace připojený na vstup kanálu $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ přímo. Předpokládejme, že množina $\{\nu_x : x \in X\}$ pravděpodobnostních měr je dominovaná mírou λ_2 v tom smyslu, že každá míra ν_2 z této množiny je absolutně spojitá vzhledem k λ_2 na σ-algebře $\mathfrak{Y}$. Hustotu míry ν_x podle λ_2 pro $x \in X$ označíme $g(x, y)$ a zobecněná entropie je

$$H_{\lambda_2}(\nu_x, \mathfrak{Y}) = -\int_Y g(x, y) \log g(x, y) \, d\lambda_2 .$$

Příslušná *střední podmíněná entropie* je pak definovaná jako střední hodnota podle μ zobecněné entropie vztahem

$$\bar{H}_{\lambda_2}(\nu_x, \mathfrak{Y}) = \int_X H_\lambda(\nu_x, \mathfrak{Y}) \, d\mu = -\int_X \int_Y g(x, y) \log g(x, y) \, d\mu \, d\lambda_2 . \tag{4.4}$$

Určeme nyní pravděpodobnostní míru ω, indukovanou zdrojem $(X, \mathfrak{X}, \mu)$ a kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ na měřitelném prostoru $(X \times Y, \mathfrak{X} \times \mathfrak{Y})$, který je kartézským součinem měřitelných prostorů $(X, \mathfrak{X})$ a $(Y, \mathfrak{Y})$. Jak známo, je tato míra jednoznačně určena již na systému měřitelných obdélníků $E \times F = \{(x, y) : x \in E, y \in F\}$, kde $E \in \mathfrak{X}$ a $F \in \mathfrak{Y}$ (připomeňme, že σ-algebra $\mathfrak{X} \times \mathfrak{Y}$ je podle definice vytvořena právě těmito obdélníky). Avšak pro tyto obdélníky platí

$$\omega(E \times F) = \int_E \nu_x(F) \, d\mu . \tag{4.5}$$

V dalším budeme nazývat pravděpodobnostní pole $(X \times Y, \mathfrak{X} \times \mathfrak{Y}, \omega)$ *dvojitým zdrojem* vytvořeným zdrojem $(X, \mathfrak{X}, \mu)$ a kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$.

Nechť ν je pravděpodobnostní míra indukovaná pravděpodobnostní mírou ω na měřitelném prostoru $(Y, \mathfrak{Y})$, t. j. pro $F \in \mathfrak{Y}$ je $\nu(F) = \omega(X \times F)$.

Pravděpodobnostní pole $(Y, \mathfrak{Y}, \nu)$, které jsme takto dostali, budeme nazývat *výstupním zdrojem*, indukovaným *vstupním zdrojem* $(X, \mathfrak{X}, \mu)$ a kanálem $(\mathfrak{X}, \nu_x, \mathfrak{Y})$.

Jestliže $\omega \ll \lambda_1 \times \lambda_2$ na $\mathfrak{X} \times \mathfrak{Y}$, kde λ_1 je míra na $\mathfrak{X}$ a λ_2 míra na $\mathfrak{Y}$, lze ukázat, že existuje podmíněná pravděpodobnost $p(\cdot \mid \cdot)$ tak, že pro každé $y \in Y$ je $p(\cdot \mid y)$ pravděpodobnostní míra v $\mathfrak{X}$ (viz [19]). Podmíněná pravděpodobnost p určuje kanál $(\mathfrak{Y}, p(\cdot \mid y), \mathfrak{X})$, který nazýváme *inversním kanálem*. Tento inversní kanál spolu s výstupním zdrojem $(Y, \mathfrak{Y}, \nu)$ indukuje vstupní zdroj, který se shoduje s původním zdrojem $(X, \mathfrak{X}, \mu)$. Poznamenejme, že původní kanál $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ je určen množinou pravděpodobnostních měr, které nejsou ničím jiným, než podmíněnou pravděpodobností $q(\cdot \mid \cdot)$ tak, že $q(\cdot \mid x)$ je pravděpodobnostní míra na $\mathfrak{Y}$ pro každé $x \in X$.

Nechť za uvedených předpokladů je $h(x, y)$ hustota míry ω podle $\lambda_1 \times \lambda_2$ a $e(x)$ hustota míry μ podle λ_1. Pak příslušné zobecněné entropie jsou

$$H_{\lambda_1 \times \lambda_2}(\omega, \mathfrak{X} \times \mathfrak{Y}) = - \int_X \int_Y h(x, y) \log h(x, y) \, d\lambda_1 \, d\lambda_2 \,, \tag{4.6}$$

$$H_{\lambda_1}(\mu, \mathfrak{X}) = H_{\lambda_1 \times \lambda_2}(\omega, \mathfrak{X}) = - \int_X e(x) \log e(x) \, d\lambda_1 \,. \tag{4.7}$$

Z předpokladu $\omega \ll \lambda_1 \times \lambda_2$ $(\mathfrak{X} \times \mathfrak{Y})$ plyne také, že $\nu_x \ll \lambda_2$ $(\mathfrak{Y})$ pro $x \in X$, takže existuje příslušná hustota $g(x, y)$ a tedy také střední podmíněná entropie daná formulí (4.4). — Jelikož, jak se snadno přesvědčíme, $h(x, y) = g(x, y) \,.$ $. \, e(x)$ $[\omega]$, tedy podle (4.4), (4.6) a (4.7) platí

$$H_{\lambda_1 \times \lambda_2}(\omega, \mathfrak{X} \times \mathfrak{Y}) = H_{\lambda_1}(\mu, \mathfrak{X}) + \overline{H}_{\lambda_2}(\nu_x, \mathfrak{Y}) \,, \tag{4.8}$$

jinak řečeno, *zobecněná entropie, odpovídající dvojitému zdroji, je rovna zobecněné entropii, odpovídající vstupnímu zdroji, zvětšené o střední podmíněnou entropii, odpovídající výstupnímu zdroji indukovanému* (podobně jako dvojitý zdroj) *původním zdrojem a kanálem.*

Je zřejmé, že podle aditivity entropie (4.8), vstupní zdroj a původní kanál lze nahradit výstupním zdrojem a inversním kanálem.

Ve speciálním případě *nezávislosti* mezi vstupem a výstupem, t. j. v případě, že všechny pravděpodobnostní míry ν_x kanálu $(\mathfrak{X}, \nu_x, \mathfrak{Y})$ jsou identické (tedy rovny ν), dostává vztah (4.8) tvar

$$H_{\lambda_1 \times \lambda_2}(\omega, \mathfrak{X} \times \mathfrak{Y}) = H_{\lambda_1}(\mu, \mathfrak{X}) + H_{\lambda_2}(\nu, \mathfrak{Y}) \,, \tag{4.9}$$

se kterým jsme se již setkali ve speciálním tvaru, kde místo σ-algebry $\mathfrak{X} \times \mathfrak{Y}$ jsme měli $\mathfrak{Y}^n$ (příslušnou formuli jsme mohli získat opakovaným použitím formule (4.9)).

Uvažujme nyní speciální případ, že λ_1 se shoduje s μ a λ_2 se shoduje s ν. V tomto případě můžeme místo (4.8) psát

$$I(\omega, \mathfrak{X} \times \mathfrak{Y}) = - H_{\mu \times \nu}(\omega, \mathfrak{X} \times \mathfrak{Y}) = - \overline{H}_{\nu}(\nu_x, \mathfrak{Y}) \tag{4.10}$$

a toto číslo je podle věty (4.1) nezáporné a je rovno nule tehdy a jen tehdy, když $\omega = \mu \times \nu$, t. j. když a jen když je stochastická nezávislost mezi vstupem a výstupem (mezi x a y).

Snadno zjistíme, že *nezávisle na dominujících mírách λ_1 a λ_2 platí*

$$I(\omega, \mathfrak{X} \times \mathfrak{Y}) = H_{\lambda_2}(\nu, \mathfrak{Y}) - \overline{H}_{\lambda_2}(\nu_x, \mathfrak{Y}) \,, \tag{4.11}$$

jinak řečeno číslo $I(\omega, \mathfrak{X} \times \mathfrak{Y})$ je rovno zobecněné entropii, odpovídající výstupnímu (vstupnímu) zdroji, zmenšené o střední podmíněnou zobecněnou entropii (podle stejné dominující míry), odpovídající témuž zdroji. Podle (4.11) *je toto číslo vždy menší nebo rovno zobecněné entropii, odpovídající výstupnímu (vstupnímu) zdroji*, t. j. apriorní entropii, *a je rovno této apriorní entropii, když a jen když je stochastická nezávislost mezi vstupem a výstupem* (mezi x a y).

Za těchto okolností *je zcela přirozené považovat I za míru informace o vstupním nebo vyslaném signálu x, získanou údajem o výstupním nebo přijatém signálu y a opačně.* Právě tuto cestu jsme nastoupili v práci [11] pro zavedení pojmu informace (zobecněné informace), podobně jako se zavádí informace v konečném případě (viz na př. [1]).

V tomto článku použijeme pro zavedení pojmu informace ještě jiného způsobu, který je analogický způsobu zavedení entropie v paragrafu 3 a podle kterého dostaneme pojem informace zcela automaticky jako důsledek jistých úvah o rozlišovací schopnosti kanálu. Tato druhá metoda bude použita v paragrafu 5. Je účelné udat již nyní některé vlastnosti míry informace $I(\omega, \mathfrak{X} \times \mathfrak{Y})$ jako přímé důsledky její definice (4.10) a to tak, že budeme identifikovat I s jistou zobecněnou entropií až na znaménko. Podle věty 4.1 platí

Věta 4.3. *Míra informace I definovaná pomocí* (4.10) *a* (4.1) *splňuje nerovnost*

$$I(\omega, X \times \mathfrak{Y}) \geqq 0 , \tag{4.12}$$

při čemž znaménko rovnosti platí tehdy a jen tehdy, když $\omega = \mu \times \nu$, t. j. v případě stochastické nezávislosti mezi x a y.

Nechť $\mathfrak{X}'$ je σ-podalgebra σ-algebry $\mathfrak{X}$ a $\mathfrak{Y}'$ σ-podalgebra σ-algebry $\mathfrak{Y}$, takže $\mathfrak{X}' \times \mathfrak{Y}' \subset \mathfrak{X} \times \mathfrak{Y}$. Z věty 4.2 plyne

Věta 4.4. *Jsou splněny nerovnosti*

$$I(\omega, \mathfrak{X} \times \mathfrak{Y}') \leqq I(\omega, \mathfrak{X} \times \mathfrak{Y}) , \tag{i}$$

$$I(\omega, \mathfrak{X}' \times \mathfrak{Y}) \leqq I(\omega, \mathfrak{X} \times \mathfrak{Y}) , \tag{ii}$$

$$I(\omega, \mathfrak{X}' \times \mathfrak{Y}') \leqq I(\omega, \mathfrak{X} \times \mathfrak{Y}) , \tag{iii}$$

a znaménko rovnosti platí: v (i) *tehdy a jen tehdy, když σ-algebra $\mathfrak{Y}'$ je suficientní* [24] *pro množinu měr $\{\nu_x : x \in X\}$ kanálu $(\mathfrak{X}, \nu_x, \mathfrak{Y})$, v* (ii) *tehdy a jen tehdy, když σ-algebra $\mathfrak{X}'$ je suficientní pro množinu měr $\{\tilde{\nu}_y : y \in Y\}$ inversního kanálu $(\mathfrak{Y}, \tilde{\nu}_y, \mathfrak{X})$ a v* (iii) *tehdy a jen tehdy, když σ-algebra $\mathfrak{X}' \times \mathfrak{Y}'$ je suficientní pro množinu měr $\{\omega, \mu \times \nu\}$. Tato poslední podmínka implikuje předcházející dvojici podmínek, avšak obecně s ní není ekvivalentní.*

Podobně jako v koroláru 4.1 můžeme také tvrzení předešlé věty vyjádřit v řeči transformací, avšak nebudeme tuto formulaci věty 4.4 uvádět explicitně.

Interpretace předešlých výsledků je snadná. Tak věta 4.3 vyjadřuje jistou vlastnost, která jak vidíme musí být obsažena v každé adekvátní definici pojmu informace, totiž že informace je rovna nule v případě nezávislosti mezi vyslaným signálem a signálem přijatým a právě jen v tomto případě. *Z tohoto hlediska může naopak každá adekvátní definice informace sloužit jako vhodná míra stochastické závislosti mezi dvěma náhodnými elementy, mezi vyslaným a přijatým signálem* (viz [11]). Koeficient korelace a různé jeho zdokonalené verse, kterých lze ostatně použít toliko ve speciálních případech, obecně nemají uvedenou vlastnost a nejsou tedy adekvátní mírou závislosti. Ukážeme hned, že koeficient korelace ani zdaleka nepředstavuje universální míru závislosti.

Jedno-jednoznačné transformace každého ze dvou náhodných elementů, jejichž stochastickou závislost měříme, obecně nezachovávají hodnotu koeficientu korelace, což je závažným nedostatkem. *Skutečně každá adekvátní míra závislosti musí být invariantní vzhledem k takovým jedno-jednoznačným transformacím.*

Zobecněná informace, definovaná v tomto paragrafu pomocí formule (4.10) *splňuje uvedený požadavek invariance,* jak plyne z věty 4.1, neboť každá jedno-jednoznačná transformace je suficientní.

Věta 4.4 vyjadřuje ještě jinou vlastnost, kterou musíme zcela přirozeně vyžadovat od adekvátní míry informace nebo závislosti, totiž že *informace* resp. *závislost se musí obecně zmenšovat, když náhodné elementy jsou popisovány čím dále tím hrubším způsobem,* což matematicky znamená postupné zmenšování příslušných σ-algeber.

Abychom mohli formulovat jisté další výsledky obsažené v práci [11], připomeňme definici *rozkladu* měřitelného prostoru $(X, \mathfrak{X})$. Je to systém po dvou disjunktních množin z $\mathfrak{X}$, jejichž součet je roven prostoru X.

Řekněme, že rozklad $\boldsymbol{Q}$ měřitelného prostoru $(X, \mathfrak{X})$ je *jemnější* než rozklad $\boldsymbol{P}$ a píšeme $\boldsymbol{Q} \prec \boldsymbol{P}$, když každá množina rozkladu $\boldsymbol{Q}$ je obsažena v nějaké množině rozkladu $\boldsymbol{P}$. Pomocí relace „$\prec$" můžeme definovat *nestoupající* posloupnost rozkladů $\boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3, \ldots,$ měřitelného prostoru $(X, \mathfrak{X})$ t. j. $\boldsymbol{P}_{n+1} \prec \boldsymbol{P}_n$ pro $n = 1, 2, 3, \ldots$.

Když $\boldsymbol{S}(\boldsymbol{P}_n)$ je σ-algebra vytvořená rozkladem $\boldsymbol{P}_n$ pro $n = 1, 2, 3, \ldots$, pak posloupnost $\boldsymbol{S}(\boldsymbol{P}_1), \boldsymbol{S}(\boldsymbol{P}_2), \boldsymbol{S}(\boldsymbol{P}_3), \ldots,$ σ-algeber je *neklesající*, t. j.

$$\boldsymbol{S}(\boldsymbol{P}_1) \subset \boldsymbol{S}(\boldsymbol{P}_2) \subset \boldsymbol{S}(\boldsymbol{P}_3) \subset \ldots \subset \mathfrak{X} \, .$$

Věta 4.5. *Když $\mu \ll \lambda$ $(\mathfrak{X})$, pak existuje nestoupající posloupnost $\boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3, \ldots,$ konečných rozkladů měřitelného prostoru $(X, \mathfrak{X})$ tak, že (nestoupající) posloupnost $H_\lambda(\mu, \boldsymbol{P}(\boldsymbol{S}_1)), H_\lambda(\mu, \boldsymbol{S}(\boldsymbol{P}_2)), \ldots,$ entropií konverguje k zobecněné entropii $H_\lambda(\mu, \mathfrak{X})$,* t. j.

$$\lim_{n \to \infty} H_\lambda(\mu, \boldsymbol{S}(\boldsymbol{P}_n)) = H_\lambda(\mu, \mathfrak{X}) \, . \tag{4.13}$$

Tato posloupnost konečných rozkladů obecně závisí na μ a λ, avšak existují důležité speciální případy (separabilní metrické prostory a pod.), ve kterých posloupnost rozkladů lze volit *nezávisle* na μ a λ.

Věta 4.6. *Když $\omega \ll \mu \times \nu$ $(\mathfrak{X} \times \mathfrak{Y})$, pak existuje nestoupající posloupnost konečných rozkladů $\boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3, \ldots,$ měřitelného prostoru $(X, \mathfrak{X})$ a nestoupající posloupnost konečných rozkladů $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \boldsymbol{Q}_3, \ldots,$ měřitelného prostoru $(Y, \mathfrak{Y})$ tak, že (neklesající) posloupnost $I(\omega, \boldsymbol{S}(\boldsymbol{P}_1) \times \boldsymbol{S}(\boldsymbol{Q}_1)), I(\omega, \boldsymbol{S}(\boldsymbol{P}_2) \times \boldsymbol{S}(\boldsymbol{Q}_2)), \ldots,$ (obyčejných) informací konverguje k zobecněné informaci $I(\omega, \mathfrak{X} \times \mathfrak{Y})$,* t. j.

$$\lim_{n \to \infty} I(\omega, \boldsymbol{S}(\boldsymbol{P}_n) \times \boldsymbol{S}(\boldsymbol{Q}_n)) = I(\omega, \mathfrak{X} \times \mathfrak{Y}) \, . \tag{4.14}$$

Existují důležité speciální případy, kdy uvedené konečné rozklady lze volit *nezávisle* na ω.

Důkazy těchto vět a osvětlení jejich vztahu k teorii martingalových náhodných procesů [23] jsou obsaženy v práci [11]. Pokud jde o význam těchto vět pro otázky aproximace zobecněné entropie resp. informace pomocí výběrových entropií resp. informací, odvoláváme se na práci [13].

Zakončíme tento paragraf následujícími dvěma výsledky:

Věta 4.7. *Platí identita*

$$H_\lambda(\mu, \mathfrak{X}) = \inf_{\boldsymbol{P}} H_\lambda(\mu, \boldsymbol{S}(\boldsymbol{P})) , \tag{4.15}$$

kde $\boldsymbol{P}$ *probíhá všechny konečné rozklady měřitelného prostoru* $(X, \mathfrak{X})$.

Věta 4.8. *Platí identita*

$$I(\omega, \mathfrak{X} \times \mathfrak{Y}) = \sup_{\boldsymbol{P},\boldsymbol{Q}} I(\omega, \boldsymbol{S}(\boldsymbol{P}) \times \boldsymbol{S}(\boldsymbol{Q})) , \tag{4.16}$$

kde $\boldsymbol{P}$ *a* $\boldsymbol{Q}$ *probíhají všechny konečné rozklady měřitelných prostorů* $(X, \mathfrak{X})$ *resp.* $(Y, \mathfrak{Y})$.

Vztahy (4.15) resp. (4.16) by mohly sloužit jako definice zobecněné entropie resp. zobecněné informace, ale je jasné, že když μ není absolutně spojitá vzhledem k λ resp. když ω není absolutně spojitá vzhledem k $\mu \times \nu$, že jsou obě uvažované hodnoty nekonečné.

LITERATURA

[1] *А. Я. Хинчин:* Об основных теоремах теории информации (УМН, XI, 1956, стр. 17–75).

[2] *J. Nedoma:* Capacity of a discrete channel (Transactions of the First Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, 1957, str. 143–181).

[3] *И. М. Гельфанд, А. Н. Колмогоров* и *А. М. Яглом:* К общему определению количества информации (ДАН СССР, *111*, 1956, № 4, стр. 745–748).

[4] *М. Розенблат-Рот:* Энтропия стохастических процессов (ДАН СССР, *112*, 1957, № 1, стр. 16–19).

[5] *М. Розенблат-Рот:* Теория передачи информации через стохастические каналы связи (ДАН СССР, *112*, 1957, № 2, стр. 202–205).

[6] *A. Perez:* Nejistota, entropie, informace (sdělení na I. pracovní konferenci čsl. matematických statistiků, Praha, červen 1954).

[7] *A. Perez:* O konvergenci posloupností nejistot, entropií a informací odpovídajících rostoucím posloupnostem σ-algeber (sdělení na IV. sjezdu čsl. matematiků, Praha, září 1955).

[8] *A. Perez:* O teorii informace v případě abstraktní abecedy (sdělení na I. pražské konferenci o teorii informace, statistických rozhodovacích funkcích a náhodných procesech, Liblice, prosinec 1956).

[9] *A. McMillan:* The Basic Theorems of Information Theory (AMS, *24*, 1953, str. 196—219).
[10] *A. Feinstein:* A New Basic Theorem of Information Theory (Trans. IRE, PGIT, 1954, str. 2—22).
[11] *A. Perez:* Notions généralisées d'incertitude, d'entropie et d'information du point de vue de la théorie de martingales (Transactions of the First Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Praha 1957, str. 183—208).
[12] *A. Perez:* Sur la Théorie de l'information dans le cas d'un alphabet abstrait (tamtéž str. 209—243).
[13] *A. Perez:* Sur la convergence des incertitudes, entropies et informations échantillon (sample) vers leurs valeurs vraies (tamtéž str. 245—252).
[14] *N. Wiener:* Cybernetics, New York, 1948.
[15] *C. Shannon:* A Mathematical Theory of Communication (BSTJ, *27*, 1948, str. 379—423).
[16] *L. Szilard:* Über die Entropieverminderung in einem thermodynamischen System bei Eingreifen intelligenter Wesen (Zeits. für Phys., *53*, 1929, str. 840—856).
[17] *L. Brillouin:* Maxwell's Demon Cannot Operate (Jour. Appl. Phys., *22*, 1951, str. 334).
[18] *А. Н. Колмогоров:* О некоторых асимптотических характеристиках вполне ограниченных метрических пространств (ДАН СССР, *108*, 1956, № 3, стр. 385—388).
[19] *P. Halmos:* Measure Theory, New York, 1950.
[20] *А. Я. Хинчин:* Метод произвольных функций и борьба против идеализма в теории вероятностей (Философские вопросы современной физики, Москва, 1952, стр. 522 až 538).
[21] *Colin Cherry:* Information Theory, Third London Symposium, 1955, London, 1956.
[22] *А. Я. Хинчин:* Понятие энтропии в теории вероятностей (УМН, VIII, 1953, стр. 3—20).
[23] *J. Doob:* Stochastic Processes, New York, 1953.
[24] *M. Loève:* Probability Theory, New York, 1955.

(Pokračování.)