Текст лекції

Визначення основних понять

Інформація (від лат. Informatio - роз'яснення, виклад) спочатку - відомості, передавані одними людьми іншим людям усним, письмовим або яким-небудь іншим способом (наприклад, за допомогою умовних сигналів, з використанням технічних засобів зв'язку і т.д.) (БСЕ - III изд., Т. 10, з. 353). Із збільшенням потоку інформації (відомостей) виникли проблеми опису передачі інформації між людьми, між людиною і автоматом, між автоматами. З'явилися деякі теорії кількісної оцінки окремих видів інформації (наприклад, в техніці зв'язку, кібернетиці). Проте повного і всеохватывающего визначення інформації на сьогодні не існує.

Класична теорія інформації не розглядає ні питання про зміст передаваних повідомлень, ні ефекту дії цих повідомлень на одержувача. Тому термін "інформація" трактує як приріст відомостей про джерело інформації, що утворюється у одержувача при отриманні інформації. Якась частка інформації була у нас апріорно (або повна її відсутність), решта відомостей про стан джерела нам не відома (є апріорна невизначеність джерела інформації). Отримання в результаті відповіді інформації про джерело збільшує у нас кількість інформації і знімає невизначеність у джерела.

Розглянемо зміну об'єму інформації в конкретних випадках. Хай у нас досвід має лише один результат і не містить ніякої невизначеності, тоді ми наперед знаємо результат цього досвіду. В результаті здійснення досвіду ми не отримаємо ніякої інформації (При передачі повідомлення "Волга впадає в Каспійське море" ми не одержуємо ніякого нового повідомлення, оскільки нам це було відоме наперед).

Хай досвід має два рівноімовірні результати (наприклад, прийом однієї посилки бінарного сигналу). Сигнал, що приймається, несе певну інформацію (вірогідність кожного сигналу $P = 1/2$).

Хай третій досвід пов'язаний з можливістю отримати один з 10 рівноімовірних результатів. В цьому випадку буде велика попередня невизначеність щодо джерела, а прийняте повідомлення дасть більш уточнену характеристику стану джерела. Вірогідність кожного результату $P(xi)= 1/10$ - менше ніж в другому досвіді.

Висновок: чим менше апріорна вірогідність події, тим більше інформації несе про джерело повідомлення (тобто тим більше несподіваний результат).

В третьому випадку невизначеність вище. Може показатися, що ступінь невизначеності визначається числом можливих станів системи. Проте, в загальному випадку це не так. Розглянемо РЕС, технічний стан якого може бути в двох станах: справно і несправно. Припустимо, що до отримання відомостей (апріорі) вірогідність справної системи

0,99, а відмова - 0,01. Така система має малий степінь невизначеності: майже напевно можна припустити, що РЕС справно. При киданні монети також два стани, але ступінь невизначеності набагато вище.

Висновок: ступінь невизначеності системи визначається не тільки числом можливих станів, але і вірогідністю станів.

Тому природно припустити, що кількісною мірою невизначеності окремого повідомлення, а також непередаваної їм інформації може бути величина, зворотна його апріорній вірогідності 1/P(xi) (що і запропонував Р.Хартлі в 1928 р.). Проте, така міра незручна (при P(xi)= 1 достовірна подія, кількість інформації виявляється не 0, а 1; крім того, немає властивості аддитивності, оскільки вірогідність двох і більш подій перемножуються). Клод Шеннон в 1948 р. ввів логарифмічну міру кількості інформації.

$$I(x_i) = \log_a \frac{1}{P(x_i)} = -\log_a P(x_i) \qquad (2.1)$$

При цьому кількість інформації, що міститься в складному повідомленні, що представляє сукупність подій xi і xj буде

$$I(x_i, x_j) = \log \frac{1}{P(x_i)P(x_j)} = \log \frac{1}{P(x_i)} + \log \frac{1}{P(x_j)} = I(x_i) + I(x_j). \qquad (2.2)$$

Властивості міри Шеннона:

1. Логарифмічна міра володіє властивостями аддитивності.
2. У разі події з одним результатом, детерміновані, тобто певні повідомлення I(x)= 0.
3. Величина інформації росте із зростанням несподіванки результату (оскільки обернено пропорційна вірогідності події).
4. Значення інформації ≥0 (позитивна).

Дані властивості відносяться до дискретної системи. Оскільки інформація випадкова, то потрібна середня міра оцінки інформації (середнє на одне повідомлення).

I(x)= M[1][2][-loga(P(xi)]= $-\sum_{i=1}^{n} P(ix_i)\log_a P(ix_i) = H(X).$ (2.3)

[1] Понятие усреднения (определение среднего значения), процедура которого обозначается символом M[ ] (см в лекции 4)

[2] Понятие усреднения (определение среднего значения), процедура которого обозначается символом M[ ] (см в лекции 4)

В основі кількості інформації лежить <u>апріорна невизначеність</u> повідомлення, тому отриманий вираз називають ще "<u>ентропією</u>" (термін запозичений з термодинаміки, де аналогічний вираз характеризує середню невизначеність стану системи молекул речовини).

Не дивлячись на збіг виразів для I(x) і H(x) ентропія і кількість інформації принципово різна. Інформація розглядається у зв'язку з своєю протилежністю – ентропією.

Ентропія визначає середню невизначеність джерела (можливий об'єм інформації у джерела), інформація зв'язується у нас з отриманням повідомлення.

Одиниця вимірювання інформації залежить від вибору підстави loga

log2 – binary digit = битий (двійкова одиниця). В основному використовується <u>біт</u>.

log10 – decimal digit = дит.

loge – natural digit = нат.

<u>Властивості ентропії:</u>

1. <u>H – речовинна, позитивна >0, обмежена оскільки P()<1.</u>
2. H = 0 для детермінованих повідомлень (з визначення).
3. H – max, якщо всі події рівноімовірні.

Для доказу скористаємося методом невизначеного множника Лагранжа (λ).

Складемо допоміжну функцію:

$$F = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i) - \lambda .$$

Оскільки $\sum_{i=1}^{n} P(x_i) = 1$, то на таку величину можна помножити ?:

$$F = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i) - \sum_{i=1}^{n} P(x_i)\lambda = -\sum_{i=1}^{n} \left[P(x_i) \log_2 P(x_i) + P(x_i)\lambda\right] = \sum_{i=1}^{n} F_i \; .$$

Необхідно знайти <u>max значення Fi</u> по змінній $P(x)_i$, для цього продиференціюємо і прирівняємо = 0:

$$\frac{\partial F_i}{\partial P(x_i)} = -P(x_i)\frac{\partial}{\partial P(x_i)}[\log_2 P(x_i)] - 1 \cdot \log_2 P(x_i) - 1 \cdot \lambda =$$

$$= -P(x_i)\frac{\log_2 \ell}{P(x_i)} - \log_2 P(x_i) - \lambda = -\log_2 \ell - \log_2 P(x_i) - \lambda = 0.$$

Для знаходження максимального значення знайдемо екстремум функції:

∂Fi/∂ P(xi)= 0

log2 P(xi)= -log2? - ? = const (i =1, 2, 3 ., n) звідки

$$P(x_i) = const = 1/n$$

Що і вимагалося довести (не залежать від номера i), що тільки тоді, коли все $P(x_i)$ однакові $1/n$.

Максимальне значення ентропії

$$H(X)_{max} = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{n} = \log_2 n . \qquad (2.4)$$

4. Ентропія бінарної системи (2-х альтернативної) змінюється від 0 до 1.

$$P(x1) + P(x2) = 1.$$

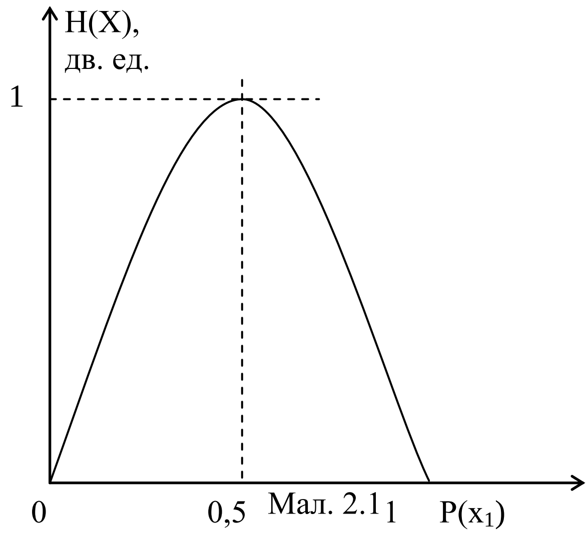
$$H(x) = -P(x1)\log_2 P(x1) - P(x2)\log_2 P(x2) = -P(x1)\log_2 P(x1) - [1-P(x1)]\log_2[1-P(x1)].$$

Якщо $P(x1)= 0, P(x2)= 1 \quad H(x)= 0$

$P(x1)= 1, P(x2)= 0 \quad H(x)= 0.$

Максимум буде, якщо $P(x1)= P(x2)= 0,5.$

$$H(x) = -\log_2(1/2) = 1 \text{ дв. ед.}$$



Мал. 2.1

2.2. Ентропія складних повідомлень

Розглянемо ентропію з'єднаної системи.

Під об'єднанням двох систем з можливими станами $x_1, x_2 ., x_n, y_1, y_2 ., y_n$ розуміється складна система (X, У), стан якої (x_i, y_i) є всіма можливими комбінаціями станів

xi, yi $_{\text{систем X}}$ і У. Очевидно, число можливих станів системи (Х, У) рівно m·n.

Позначимо через P(yy/xi) умовну вірогідність того, що система У приймає стан yj за умови, що система Х знаходиться в стані xi.

Визначимо тепер ентропію системи У за умови, що система Х знаходиться в стані xi (приватна умовна ентропія).

$$H(Y/xi)= MY[-logP(yj/xi)]= -\sum_{j=1}^{m} P(y_j / x_i) \log P(y_j / x_i) .$$

Середня по безлічі всіх можливих станів системи Х умовна ентропія (повна умовна ентропія)

$$H(Y/X)= MX[H(Y/xi)]= -\sum_{i=1}^{n}\sum_{j=1}^{m} P(y_j / x_i) \log P(y_j / x_i) \qquad (2.5)$$

Умовна ентропія H(Y/X) характеризує ступінь невизначеності системи У за умови, що стан системи Х повністю визначено.

Неважко переконатися, що

$$H(Y/X)= H(Y) \qquad (2.6)$$

при незалежності вірогідності систем Х і У, а також

$$H(Y/X)= 0 \qquad (2.7)$$

при однозначній (функціонального зв'язку) між системами.

З (2.6) і (2.7) очевидно, що умовна ентропія досягає максимуму при незалежності вірогідності систем. Це твердження можна строго довести методами варіаційного числення, але і так представляється достатньо очевидним, що невизначеність однієї системи не може збільшитися від того, що невизначеність якоїсь іншої системи зменшилася.

Доведемо наступну теорему.

Якщо дві системи Х і У об'єднуються в одну, то ентропія з'єднаної системи рівна ентропії однієї з складових частин плюс умовна ентропія другої частини щодо першої.

$$H(X,Y)= M[-logP(X,Y)]= M[-log(P(X)\cdot P(Y/X))] =$$
$$= M[-logP(X)]+ M[-logP(Y/X)]= H(X)+ H(Y/X). \qquad (2.8)$$

В окремому випадку, коли системи Х і У незалежні

$$H(Y/X)= H(Y) \quad і \quad H(X,Y)= H(X)+ H(Y).$$

Оскільки H(Y/X)? H(Y), то H(X,Y)? H(X)+ H(Y), тобто ентропія складної системи досягає максимуму у разі, коли її складові частини незалежні.

Теорему про ентропію складної системи легко розповсюдити на будь-яке число об'єднуваних систем.

Виходячи з висловленого, можна пояснити, чому кількісне представлення інформації через ентропію виявилося так широко застосовним. Таке уявлення володіє наступними достоїнствами:

1. Задовольняє вимозі, що пред'являється до будь-якої міри – аддитивності, по якому загальна від декількох джерел інформація знаходиться підсумовуванням.

2. Добре відображає значення поняття "інформація" - середня кількість інформації про систему, яке може бути отриманий, тобто ентропія досягає максимуму у разі, коли апріорні дані про систему відсутні (всі стани системи рівноімовірні) і рівно 0, якщо невизначеність системи відсутня.

Досить поширеним є випадок, коли що цікавить нас система подій (випадкова величина) вивчається не безпосередньо, а шляхом вивчення іншої системи, пов'язаної з першою вірогідність. Оцінимо взаємну інформацію систем.

Хай нас цікавить система X. Возможные її стану визначаються апріорною вірогідністю P(x1), P(x2) ., P(xn). Хай також є система У, вірогідність пов'язана з системою X (відома умовна вірогідність P(xi/yk). При отриманні повідомлення, що система У знаходиться в к-м стані, змінився розподіл вірогідності системи X, тобто ми отримали певну інформацію про систему X. Приращение інформації про i-том стан системи X

$$I_{y_k \to x_i} = \log \frac{P(x_i / y_k)}{P(x_i)}.$$

Ця інформація називається інформацією "від події до події".

В середньому по всіх можливих станах системи X приріст інформації

$$I_{y_k \to X} = M_x[I_{y_k \to X}] = \sum_{i=1}^{n} P(x_i / y_k) \log_2 \frac{P(x_i / y_k)}{P(x_i)}.$$

Ця величина називається середньою приватною інформацією.

Середня по всіх можливих станах системи У інформація про систему X

$$\begin{aligned} I_{Y \to X} = M_y[I_{y_k \to X}] &= \sum_{i=1}^{n} \sum_{k=1}^{m} P(y_k) P(x_i / y_k) \log_2 \frac{P(x_i / y_k)}{P(x_i)} = \\ &= \sum_{k=1}^{m} \sum_{i=1}^{n} P(x_i, y_k) \log_2 \frac{P(x_i, y_k)}{P(x_i) P(y_k)}. \end{aligned} \tag{2.9}$$

Симетричність запису виразу (2.9) щодо X і У означає, що

IY>X = IX>Y = YX-Y = I(X,Y)= H(X) – H(X/Y)= H(Y) – H(Y/X). (2.10)

Середня кількість інформації, одержувана при неповній достовірності повідомлень рівно різниці безумовної апріорної інформації H(X) і умовної апріорної інформації H(X/Y), H(X/Y) потрактує як втрата інформації (ненадійність зв'язку).

З виразу (2.8) виходить:

$$H(X/Y)= H(X,Y) - H(Y)$$

і підстановка в рівняння (2.10) дасть

$$YX\text{-}Y = H(X) – H(X,Y)+ H(Y)= H(X)+ H(Y) – H(X,Y)$$

де H(X,Y) – втрата інформації.

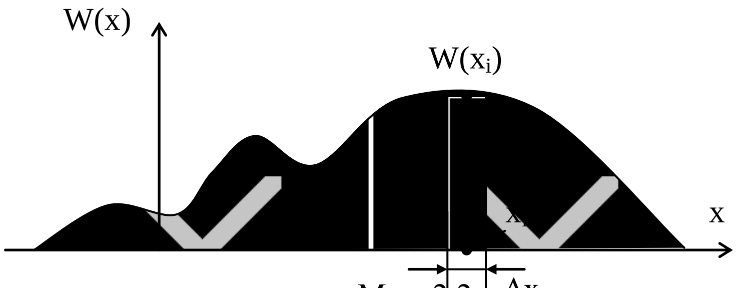
Підсумуємо властивості взаємної інформації.

1. I(X,Y)? 0;   I(X,Y)= 0, коли X і У – незалежні.
2. I(X,Y)= I(Y,X).
3. I(X,Y)? H(X);   I(X,Y)= H(X), коли H(X/Y)= 0 при однозначному зв'язку.
4. I(X,X)= H(X) – власна інформація про себе.

Ентропія безперервної випадкової величини

В каналах передачі інформації часто використовуються сигнали, миттєві значення яких можуть приймати будь-які значення на деякому інтервалі (мовні, музичні, телевізійні сигнали і т.д.)

Розповсюдимо поняття ентропії на випадок безперервної випадкової величини (див. мал. 2.2)



Мал. 2.2

Вірогідність попадання сигналу в проміжок x.x+dx рівна

$$dp = \lim_{\Delta x \to 0} W(x)dx$$

Таким чином:

$$H(x)=\lim_{\Delta x\to 0}\left\{-\sum_i W(x_i)\Delta x \log[W(x_i)\Delta x]\right\}=$$
$$=-\int_{-\infty}^{\infty} W(x)\log W(x)dx-\lim_{\Delta x\to 0}\left[\log \Delta x \sum_i W(x_i)\Delta x\right]= \qquad (2.11)$$
$$=-\int_{-\infty}^{\infty} W(x)\log W(x)dx-\lim_{\Delta x\to 0}\log \Delta x.$$

Друга складова (2.11) при переході до межі перетворюється в ?. Таким чином, ентропія безперервної випадкової величини рівна ?. У зв'язку з цим може виникнути сумнів в доцільності энтропийного принципу вимірювання інформації стосовно безперервно розподілених сигналів. Проте, з теоретичної точки зору ця трудність не є принциповою. Річ у тому, що другий доданок виразу (2.11) не залежить від характеристик вірогідності випадкової величини, іншими словами, другий доданок з точністю до нескінченно малої величини однаково для всіх випадкових величин. Оскільки в реальних каналах завжди мають місце шуми і обчислення інформаційних характеристик каналів зводиться до визначення різниці ентропії сигналу і шуму, в результаті віднімання складові ентропії вигляду $\lim_{\Delta x\to 0}\log \Delta x$ і $\lim_{\Delta y\to 0}\log \Delta y$ взаємно знищуються.

При розгляді експериментальних даних справа спрощується ще і тим, що елементи ?x залишаються кінцевими, оскільки ці величини визначаються роздільною здатністю вимірювальних приладів, яка не може бути нескінченною.

Оскільки при обчисленні різниці ентропій другий доданок виразу (2.11) не представляє інтересу, використовують тільки першу складову виразу (2.11).

$$H^*(x)=-\int_{-\infty}^{\infty} W(x)\log W(x)dx \qquad (2.12)$$

яке називається приведеною або відносною ентропією.

Оскільки вирази для приведеної ентропії безперервної випадкової величини і ентропії дискретної випадкової величини аналогічні, очевидно, що приведена ентропія досягає максимуму при рівноімовірному розподілі станів.

Відомо, що при фіксованій дисперсії ентропія максимальна при нормальному законі розподілу, тобто безперервна випадкова величина з фіксованою дисперсією і нормальним розподілом володіє максимальною інформативністю.

Приведена ентропія нормально розподіленої випадкової величини

$$H^*(x) = -\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} \log_2 \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \right] dx =$$

$$= \frac{\log_2(\sigma\sqrt{2\pi})}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx + \frac{\log e}{2\sigma^3\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx = \log_2(\sigma\sqrt{2\pi e}). \qquad (2.13)$$

Приватна і середня взаємні ентропії безперервного сигналу

Виконавши перетворення, аналогічні тим, які були виконані для дискретного сигналу, можна отримати наступні вирази для взаємної інформації безперервного сигналу:

1. інформація від події до події

$$I(y, x) = \log_2 [W(x/y)/W(x)],$$

2. приватна інформація

$$I(y, X) = -\int_{-\infty}^{\infty} W(x, y) \log_2 \frac{W(x/y)}{W(x)} dx,$$

3. повна (середня) інформація

$$I(X, Y) = \int_{-\infty}^{\infty}\int W(x, y) \log_2 [W(x, y)/W(x)W(y)] dxdy =$$

$$= H^*(X) - H^*(X/Y) = H^*(Y) - H^*(Y/X).$$

Джерела інформації

Джерелами повідомлень можуть бути об'єкти, стан яких визначається деяким фізичним процесом, що відбувається в часі або в просторі. До джерел повідомлень з просторовим розподілом носія інформації відносяться книги, картини, грамплатівки і д.т. При передачі інформації відбувається, як правило, перетворення просторового розподілу в тимчасовий.

Джерела інформації можуть бути дискретними і безперервними.

По характеру роботи джерела діляться на дві групи: з регульованою і з нерегульованою продуктивністю (швидкістю вироблення інформації). До першої групи відносяться джерела з пам'яттю, видаючі інформацію залежно від режиму роботи кодопреобразователя або за запитом. До другої групи відносяться джерела без пам'яті.

Хай дискретне джерело повідомлень виробляє деяку послідовність символів, причому порядок проходження цих символів випадковий і характеризується деякою сукупністю вірогідності.

В найпростішому випадку для опису процесів достатньо тільки безумовної вірогідності символів. В більш загальному випадку, коли вірогідність появи символу залежить від того, яким були попередні, необхідно знати умовну вірогідність.

Дискретна послідовність, в якій вірогідність появи символу залежить тільки від того, яким був попередній, називається простим ланцюгом Маркова. Якщо корелятивні зв'язки тягнуться на більше (але кінцеве) число символів, процес називається складним ланцюгом Маркова.

Для простого ланцюга Маркова

$$H(X) = -\sum_{i=1}^{n}\sum_{k=1}^{n} P(x_{\ell,k}, x_{\ell-1,i}) \log P(x_{\ell,k} / x_{\ell-1,i}).$$

Для послідовності не зв'язаних між собою вірогідністю символів

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

Оскільки безумовна ентропія при заданій безумовній вірогідності більше любимої умовної, кількість інформації повідомлення, що доводиться на один символ, досягає максимуму у разі відсутності кореляційних зв'язків в повідомленні.

Безумовна ентропія має максимальне значення при равновероятности всіх символів. Отже, максимальне значення ентропії на символ має місце у тому випадку, коли, по-перше, між символами відсутні зв'язки вірогідності, а, по-друге, коли всі символи алфавіту рівноімовірні. Певне таким чином максимальне значення ентропії джерела називається інформаційною місткістю джерела. Інформаційна місткість джерела, що використовує алфавіт з підставою L

$$C = H_{max} = -L\frac{1}{L}\log\frac{1}{L} = \log L.$$

Для характеристики використовування символів в повідомленні введений параметр, званий надмірністю.

$$R = \frac{H_{max} - H(x)}{H_{max}} = 1 - \frac{H(x)}{H_{max}} = 1 - M. \qquad (2.14)$$

Величину $\frac{H(X)}{H_{max}}$ називають коефіцієнтом стиснення = M, H(x) – ентропія на один символ повідомлення.

Надмірність приводить до збільшення часу передачі інформації, зайвому завантаженню каналу зв'язку. Є і певна надмірність в російській мові і в європейських мовах. Приведемо таблиці відносної частоти появи букв (вірогідність) в російській і англійській мовах.

Вірогідність появи букв в російському тексті

Буква	- (пропуск)	про	е, е	а, і	т, н	з	р	в	л
Вірогідність	0,175	0,090	0,072	0,062	0,053	0,045	0,040	0,038	0,035
Буква	до	м	д	п	у	я	ы, з	ь, ъ	би
Вірогідність	0,028	0,026	0,025	0,023	0,021	0,018	0,016	0,014	0,014
Буква	г	ч	й	х	же	ю, ш	ц	щ, э	ф
Вірогідність	0,013	0,012	0,010	0,009	0,007	0,006	0,004	0,003	0,002

Вірогідність появи букв в англійському тексті

Буква	- (пропуск)	e	t	про	a	n	i	r	s
Вірогідність	0,200	0,105	0,072	0,065	0,063	0,059	0,055	0,054	0,052
Буква	h	d	l	з	f, u	m	p	y, w	g
Вірогідність	0,047	0,035	0,029	0,023	0,022	0,021	0,018	0,012	0,011
Буква	b	v	до	x	j	q	z		
Вірогідність	0,010	0,008	0,003	0,002	0,001	0,001	0,001		

Російська мова містить 31 букву (е і е, ь і ъ – не розрізняємо). З урахуванням пропуску (-) між буквами – 32 символи.

За умови равновероятности і незалежності символів середня ентропія на символ буде максимальною

$$H(x)_{max} = \log 232 = 5 \ \frac{\text{дв.ед.}}{\text{симв.}} \ (\text{ в англійській мові Hmax} = 4{,}75 \ \frac{\text{дв.ед.}}{\text{симв.}}).$$

Якщо врахувати різну вірогідність символів, то

$$H1(x)= 4{,}39 \ \frac{\text{дв.ед.}}{\text{симв.}} \ (\text{в англійській мові H1} = 4{,}03 \ \frac{\text{дв.ед.}}{\text{симв.}}).$$

З урахуванням кореляції між двома символами ентропія зменшується

$$H2(x)= 3{,}52 \ \frac{\text{дв.ед.}}{\text{симв.}} \ (\text{в англійській мові H2} = 3{,}52 \ \frac{\text{дв.ед.}}{\text{симв.}}),$$

між трьома символами:

$$H3(x) = 3{,}00\ \frac{\text{дв.ед.}}{\text{симв.}}\ \text{(в англійській мові } H3 = 3{,}10\ \frac{\text{дв.ед.}}{\text{симв.}}\text{)},$$

між вісьма символами:

$$H8(x) = 2{,}00\ \frac{\text{дв.ед.}}{\text{симв.}}\ \text{(в англійській мові } H8 = 1{,}86\ \frac{\text{дв.ед.}}{\text{симв.}}\text{)}$$

і далі залишається незмінною, отже, надмірність російської мови:

$$R = 1 - \frac{H_8(x)}{H_{max}} = 1 - \frac{2}{5} = 0{,}6,$$

в англійській мові: $R = 1 - \frac{1{,}86}{4{,}75} = 0{,}61.$

У всіх європейських мовах надмірність приблизно однакова.

Надмірність розмовних мов сформувалася в результаті дуже тривалої суспільної практики і дозволяє відновлювати цілі слова і фрази при їх спотвореннях під впливом різних чинників, що заважають.

Ще джерела інформації оцінюються по кількості інформації, що виробляється в одиницю часу:

$$\overline{H} = \frac{H(X)}{\bar{\tau}_c} \tag{2.15}$$

де $\bar{\tau}_c$ - середня довжина символу.

Наприклад, для простого марківського джерела

$$\bar{\tau}_c = \sum_{i=1}^{L} \sum_{k=1}^{L} P(x_i) P(x_k / x_i) \tau_k ,$$

де ?$_k$ - тривалість к-го символу;

P(xk/xi) – вірогідність вироблення к-го символу за умови, що попереднім був і-тый символ.

Величину $\overline{H}(X)$ називають <u>швидкістю створення повідомлень</u>, <u>продуктивністю джерела</u>, а також <u>потоком повідомлень</u>.

Для отримання можливо більшій швидкості створення повідомлень, необхідно, по-перше, забезпечити можливо велику ентропію на символ, а, по-друге, зменшити до можливих меж середню тривалість символів.