

УДК 004.6; 528; 004.8.

ИНФОРМАЦИЯ, УМЕНЬШАЮЩАЯ НЕОПРЕДЕЛЕННОСТЬ И ИНФОРМАЦИЯ, УВЕЛИЧИВАЮЩАЯ СОДЕРЖАТЕЛЬНОСТЬ

Елсуков Павел Юрьевич,
к.т.н., старший научный сотрудник,
e-mail: elsukov_p@bk.ru,

Федеральное государственное бюджетное учреждение науки Институт систем энергетики им. Л.А. Мелентьева Сибирского отделения Российской академии наук (ИСЭМ СО РАН),
http://isem.irk.ru,
Иркутск, Россия

Статья анализирует два вида информации. Первый вид информации связан с энтропией и работами К.Э. Шеннона. Эта информация уменьшает неопределенность. Второй вид информации связан с негэнтропией и работами Н. Виннера и Л. Флориди. Статья описывает различие между понятиями «незнание» и «знание». Приводится пример, показывающий различие между количеством информации по К.Э. Шеннону и по Н. Виннеру. Описан метод оценки негэнтропии на основе «не гаусовости».

Ключевые слова: науки об информации, теория информации, количество информации, знание, не знание, информационная неопределенность

Введение

В настоящее время в науке иногда применяют «догму одномерности». Она проявляется в ситуациях, когда сложное или полисемическое понятие пытаются интерпретировать с помощью одного определения. Такая ситуация возникает, когда сложное явление пытаются описать одним термином или одной дефиницией. Такая ситуация возникает, когда множество разных ситуаций пытаются заменить одной общей моделью, в реальности точно описывающую только одну из множества ситуаций, а другие ситуации – приближенно. Например, говорят, что поезд идет со средней скоростью 55 км/час, что является упрощением, так как в реальности скорость поезда меняется в зависимости от ситуации. Однако вариации отбрасывают и принимают упрощенную ситуацию. По существу подобные упрощения реальности исключают детали и различия. В некоторых случаях сведение множества ситуаций к одной с исключением остальных создает искаженную модель реальности. Подобная ситуация имеет место с теорией информации, когда всю информационную теорию связывают с математической теорией коммуникации К.Э. Шеннона, исключаяющей семантику.

Материалы и методы

В качестве материалов использовались работы по статистической теории информации К.Э. Шеннона. Использовались работы по семантической теории информации Н. Виннера, Р. Карнапа, Л. Флориди, А. Н. Колмогорова. В качестве методов использовался системный и логический анализ.

Два подхода к теории информации

Когда К.Э. Шеннон писал свою знаменитую работу «Математическая теория связи» [1], он не ставил целью создание теории информации. В соответствии с исследователями [2], описывающими его деятельность, к тому времени использовались следующие системы, характеризующие разные направления связи: телеграф (1830), телефон (1870), беспроводный телеграф (1890), радио (1900), телевидение (1930), телетайп (1930), системы частотной модуляции (1930), вокодер (1930) и др. Не имелось универсальной теории для объединения разных направлений. Поэтому коммуникация или передача информации была главной целью исследования. В своей работе [1] К.Э. Шеннон пишет замечательную фразу «*семантические аспекты информации не релевантны техническим проблемам связи*». Этим он делает четкое разграничение между содержательностью информации и ее передачей по каналам связи. Однако в последующие годы, вплоть до настоящего времени, многие авторы, особенно специали-

сты в области гуманитарных наук, трактуют теорию К.Э. Шеннон как всеобщую теорию информации, а информацию, которую рассматривал К.Э Шеннон как единственную универсальную информацию. Работа [1] и последующая его работа с Уивером [3] рассматривает определенный тип информации не равнозначный другим видам информации. Информацией в теории Шеннона-Уивера называют «нечто», что уменьшает неопределенность [4] и уменьшает «незнание».

Винер назвал теорию К.Э. Шеннона «статистической теорией количества информации» [5]. «В этой теории за единицу количества информации принимается количество информации, *передаваемое при одном выборе между двумя равновероятными альтернативами*. Такая мысль возникла одновременно у Р.А. Фишера, у доктора Шеннона из *Bell Labs* и у автора настоящей книги» [5]. «При этом Фишер исходил из статистической теории, а Шеннон из проблемы кодирования информации» [5].

Количеством информации в этой теории называют информационную емкость носителя информации, или информационную емкость сообщения. Ее измеряют в байтах и битах, безотносительно к тому, что в этих байтах содержится или не содержится ничего. По К.Э. Шеннону совокупность символов содержит информацию, определяемую количеством информации в символах, безотносительно к тому, в какой последовательности эти символы объединены. Они могут образовать слово из словаря или бессмысленную совокупность. На количество информации по Шеннону это не влияет.

Первоначально К.Э. Шеннон не относил свою работу к теории информации, но затем появилась совместная работа Шеннона и Уивера [3], в которой возникла точка зрения «теории информации на основе математической теории связи». В работе [3] рассматривается «три уровня» проблемы, опять же коммуникации: техническая проблема передачи сообщений, семантическая проблема, проблема эффективности передачи сообщений. В этой работе [3] как и в первой [1] речь идет о передаче сообщений, а не о передаче информации. Проблема семантики сводится к вопросу «**How** precisely do the transmitted symbols convey the desired meaning?» «Как точно передаваемые символы передают желаемое значение?». Однако ответ и на этот вопрос семантики не касался, а касался точности полученных символов сообщения в сравнении с символами в исходном сообщении.

Основоположником информационной теории следует считать Р. Хартли [6], который занимался исследованием информации и ее количества. Именно Р. Хартли ввел понятие количества информации в сообщении. Именно его идеи использовал К.Э. Шеннон. Еще она идея, которую использовал К.Э. Шеннон, это идея энтропии. Останавливаемся на этом факте потому, что существует экономико-математический словарь, в котором утверждается, что впервые понятие энтропии ввел К.Э. Шеннон.

На самом деле почти за 200 лет до работы [1] К.Э. Шеннона это понятие ввел Клаузиус (1865). Независимо от него это понятие ввел Виллард Гибс (1873). Затем (так считается) независимо от Клаузиуса энтропию ввел Л. Больцман (1897):

$$S = H = k \ln W, \quad (1)$$

где W – вероятность состояния; k – постоянная Больцмана. Единица измерения величины S определялась как (джоуль/ (k х микросостояние)). Таким образом, *постоянная Больцмана* задавала единицу измерения энтропии, которая оценивалась на одно микросостояние. Энтропия у К.Э. Шеннона рассматривалась как мера неопределенности и информация, определяемая на ее основе, рассматривалась как информация, снимающая неопределенность, то есть уменьшающая незнание.

Такая точка зрения на теорию информации была не однозначно принята научным сообществом и стали появляться работы отражающие другие точки зрения на информацию. Другое направление было изложено в работах Н. Виннера [5; 7], Р. Карнапа [8; 9], Л. Флориди [10], А. Н. Колмогорова [11], Ю.А. Шрейдера [12] и др. [13; 14]. В этом подходе под информацией понимают семантику содержания информации безотносительно к ее объему. Эта теория часто называется семантической теорией информации. Под «информацией» в этой теории понимают «нечто», что содержит знание и увеличивает в итоге знание получателя. Количество информации по К.Э. Шеннону и Н. Виннеру различается [15] и это является предметом регулярного обсуждения.

Анализ оппозиционной пары «незнание – знание»

По Шеннону информация уменьшает «незнание», по Виннеру и Флориди – информация увеличивает «знание». «Незнание» относится к «не-факторам» [16], то есть к оппозиционной величине [17], образованной отрицанием первой величины.

Для раскрытия содержания «незнания» можно воспользоваться работой Нариньяни о «НЕ-факторах» [16]. Нариньяни утверждает «НЕ-факторы отнюдь не нечто диковинное, – они «встроены» в нашу Картину мира, как основная составляющая знаний, являясь их материей и строительным материалом» [16]. Картина мира как модель служит универсальной моделью оценки результатов научных исследований [18].

Исследование НЕ-факторов Нариньяни связывает с идеей *недоопределенных множеств*, которую затем стали связывать с теорией нечетких множеств (*fuzzy sets*).

В ряде работ Нариньяни формирует идею *недоопределенности* как общего явления. При переносе этого метода в область информации возникает понятие информационной неопределенности. «Выяснилось, что интервальное представление может служить примером как недоопределенного числа, так и неточного с тем принципиальным различием, что первое может уточняться, а второе нет» [19]. Развивая этот подход, следует констатировать наличие незнания с одной стороны, знания, с другой стороны, и нечетких множеств между этими понятиями.

Для пояснения различия между подходами Шеннона и Винера рассмотрим один пример. Полагаем, что небольшой кубик (*K*) лежит на одной из 81-й клетки квадрата 9 x 9 (рисунок 1). Неопределенность состоит в вопросе: «На какой клетке лежит кубик?».

Сообщение, уменьшающее неопределенность и незнание звучит так: «Кубик *не лежит* на клетке *a*».

Сообщение, содержащее определенность и знание звучит так: «Кубик *лежит* на клетке *b*».

1	2	3	4	5	6	7	8	9
10								18
19								27
28								36
37								45
46					<i>K</i>			54
55								63
64								72
73	74	75	76	77	78	79	80	81

Рисунок 1 – Условное положение кубика (*K*) в квадрате

По Шеннону все сообщения, уменьшающие неопределенность, следующего характера: «Кубик не содержится в первой клетке»; «Кубик не содержится в четвертой клетке»; «Кубик не содержится в десятой клетке»; «Кубик не содержится в сороковой клетке» и так далее.

Каждое из перечисленных сообщений информирует получателя и уменьшает его незнание (информационную неопределенность). Однако, только при получении восьмидесяти сообщений, получатель методом исключения может прийти к выводу, что кубик находится в клетке 50.

По Винеру правильным и содержательным будет только одно сообщение «Кубик содержится в 50 клетке». Все остальные сообщения типа: «кубик может содержаться в 10 и/или в 20 квадрате», «кубик не содержится...» являются бессодержательными по Виннеру, но содержательными по К.Э. Шеннону. По рассмотренному примеру следует, что 80 сообщений по Шеннону, уменьшающие неопределенность, равнозначны одному содержательному сообщению по Винеру.

В общем случае для *n* состояний совокупность *n-1* сообщений, уменьшающих «незнание», эквивалентны одному сообщению, содержащему знание.

Негэнтропия и энтропия как отражение отношений между определенностью и неопределенностью

Понятие знания и незнания связаны с понятием системности и хаоса. Энтропия, с помощью которой определяют информацию снимающую неопределенность, характеризует неопределенность и хаос. В противовес ей введено понятие негэнтропии, которая характеризует системность.

Негэнтропия имеет множество значений: отрицательная энтропия, синтропия, экстропия, эстропия и entaxy. Термин «отрицательная энтропия» впервые был введен физиком Эрвином Шредингером в 1944 г. в книге «Что такое жизнь?» [20]. Леон Брилюэн сократил этот термин до «негэнтропии» [21]. При этом его идея состояла в переносе понятия в природу и живые системы. В силу этого понятие негэнтропии используется биологами в качестве основы исследования цели. В 2009 году Махуликар и Хервиг дали новое определение негэнтропии как динамически упорядоченной подсистемы или как специфический дефицит энтропии в упорядоченной подсистеме по отношению к окружающему хаосу [12].

В силу разнообразия применения негэнтропия имеет больше единиц измерения в системе единиц СИ. В частности, ее определяют на основе удельной энтропии на единицу массы, или на основе удельной энтропии на единицу энергии.

Термин «негэнтропия» применяется в информационной области. Ее связывают с понятием «негауссовости». Распределение Гаусса приводят, когда говорят о случайной величине и неопределенности. Рост энтропии можно рассматривать как стремление элементов системы к случайным величинам, то есть к распределению Гаусса.

Это определило один из подходов нахождения негэнтропии на основе выделения разницы между распределением случайной величины и реальным распределением величины p_x , характеризующей статистический ансамбль или систему. Согласно этому подходу негэнтропию $J(p_x)$ определяют как разницу между случайным Гауссовским распределением $S(\phi_x)$ и распределением в реальной ситуации $S(p_x)$ [23; 24].

$$J(p_x) = S(\phi_x) - S(p_x).$$

Функция $S(\phi_x)$ есть дифференциальная энтропия гауссовой плотности с тем же средним и дисперсией, как p_x . Функция $S(p_x)$ есть дифференциальная энтропия распределения величины p_x с тем же средним и дисперсией, как p_x . Такой подход привел к появлению термина «негауссовость» (Non-Gaussianity). С этих позиций метод оценки $J(p_x)$ зависит от метода измерения негауссовости (Measures of Non-Gaussianity) (рисунок 2) [24].

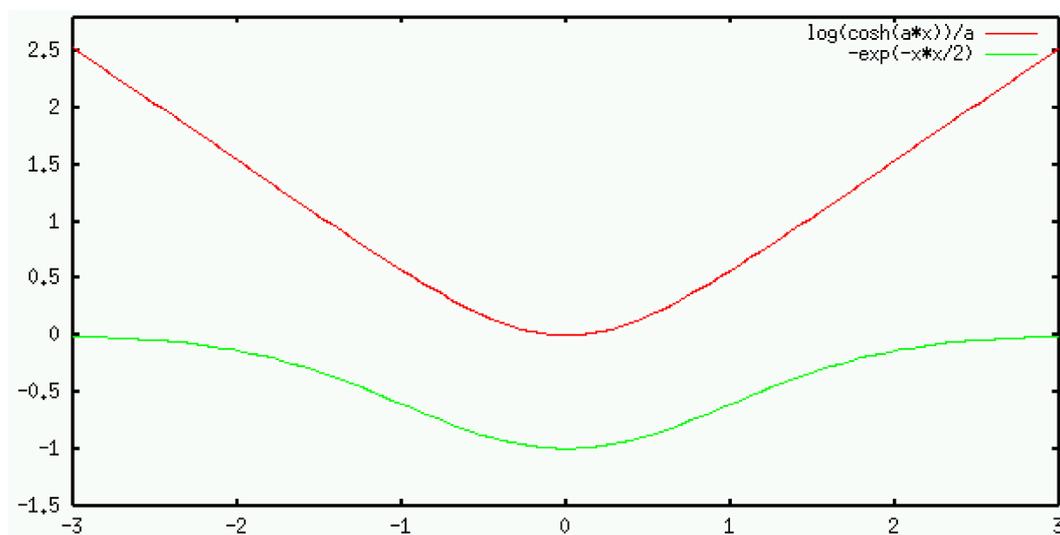


Рисунок 2 – Негэнтропия

Сравнивая рисунок 2 с графиком энтропии [13] можно отметить, что максимум негэнтропии соответствует минимуму энтропии и наоборот.

В молекулярной биологии и термодинамике негэнтропия J состояния определяется как разница между максимальной энтропией S_{\max} и энтропией этого состояния S .

$$J = S_{\max} - S.$$

Содержательное сообщение в приведенном примере – это пример применения негэнтропии. Сообщения, снимающие неопределенность это пример применения энтропии. Это разные виды информации, имеющие разные качества и приравнять их нельзя. Этот пример подчеркивает различие между негэнтропией и энтропией в качественном и количественном смысле.

Заключение

Информация, снимающая неопределенность, имеет другое качество в сравнении с информацией, создающей знание. Следует отметить наличие различных видов информации, для каждого из которых существуют свои условия применения. Информация, снимающая неопределенность, связана с энтропией. Информация, создающая знание, связана с негэнтропией. Теория информации, связанная с работами К.Э. Шеннона, согласно Н. Виннеру, должна называться статистической теорией информации и это должно определить ее сущность и границы применения. Теория информации, связанная с работами Н. Виннера и Л. Флориди, относится к семантической теории информации. Семантическая информация исследует задачи передачи семантики и развивает теории создания знания. Обе информации дополняют друг друга. Использование информации, снимающей неопределенность, более допустимо для случаев, когда нельзя получить информацию создающей знание.

Список литературы.

1. *C.E. Shannon* (1948). «A Mathematical Theory of Communication» // *Bell System Technical Journal*. – Vol. 27. – Pp. 379–423 & 623–656, July & October, 1948.
2. *Thomas M. Cover, Joy A. Thomas*. *Elements of information theory*, 2nd Edition. – New York: Wiley–Interscience, 2006. ISBN 0471241954.
3. *Claude E. Shannon. Warren Weaver*. *The Mathematical Theory of Communication*. Univ of Illinois Press, 1963.
4. *Цветков В.Я.* Информационная неопределенность и определенность в науках об информации // *Информационные технологии*. – 2015. – №1. – С. 3–7.
5. *Winner N.* *Cybernetics or Control and Communication in the Animal and the Mashine*. The Technology Press and John Wiley & Soris Inc. – New York – Herman et Cie, Paris, 1948. – 194 p.
6. *Hartley R.V.L.* «Transmission of Information» // *Bell System Technical Journal*, July, 1928 – Pp. 535–541.
7. *Винер Н.* *Кибернетика, или управление и связь в животном и машине*. – М. Советское радио, 1958.
8. *Carnap R.* *Empiricism, semantics, and ontology* // *Revue internationale de philosophie*. – 1950. – С. 20–40.
9. *Bar-Hillel Y, Carnap R.* *An Outline of a Theory of Semantic Information* // *Bar-Hillel Y. Language and Information: Selected Essays on Their Theory and Application*. – London : Addison-Wesley, 1964. – P. 221 – 274.
10. *Floridi L.* (ed.). *The Blackwell guide to the philosophy of computing and information*. – John Wiley & Sons, 2008.
11. *Колмогоров А.Н.* Три подхода к определению количества информации // *Проблемы передачи информации*. – 1965. – Т. 1. Вып. 1. – С. 3–11.
12. *Шрейдер Ю.А.* О семантических аспектах теории информации // *Информация и кибернетика*. – М.: Советское радио, 1967. – С. 15–47.
13. *Иванников А.Д., Тихонов А.Н., Цветков В.Я.* *Основы теории информации*. – М.: Макс Пресс, 2007. – 356 с.
14. *Adriaans P., van Benthem J.* *Philosophy of information* // *Handbook of the Philosophy of Science*. – 2008. – № 8.
15. *Tsvetkov V.Ya.* The K.E. Shannon and L. Floridi’s amount of information // *Life Science Journal*. – 2014. – № 11 (11). Pp. 667–671.
16. *Нариньяни А.С.* НЕ-факторы: краткое введение // *Новости искусственного интеллекта*. – 2004. – Вып. 2. – М.: КОМКНИГА, 2006. – С. 52–63.
17. *Tsvetkov V.Ya.* Dichotomous Systemic Analysis // *Life Science Journal*. – 2014. – № 11(6). – Pp. 586–590.
18. *Tsvetkov V.Ya.* Worldview Model as the Result of Education // *World Applied Sciences Journal*. – 2014. – № 31 (2). – P. 211–215.
19. *Нариньяни А.С.* Недоопределенные модели и операции с недоопределенными значениями. – Препринт ВЦ СО АН СССР. – № 400. – 1982.
20. *Schrödinger Erwin.* *What is Life – the Physical Aspect of the Living Cell*. Cambridge University Press, 1944.
21. *Brillouin, Leon:* (1953) “Negentropy Principle of Information”, *J. of Applied Physics*. V. 24(9). – Pp. 1152–1163.

22. Mahulikar, S.P. & Herwig, H. (2009). «Exact thermodynamic principles for dynamic order existence and evolution in chaos», *Chaos, Solitons & Fractals*. V. № 41(4). Pp. 1939–1948.
23. URL: <https://en.wikipedia.org/wiki/Negentropy> (дата доступа: 20.05.2017).
24. Measures of Non-Gaussianity. URL: <http://fourier.eng.hmc.edu/e161/lectures/ica/node4.html> (дата доступа: 20.05.2017).

References

1. C.E. Shannon (1948). “A Mathematical Theory of Communication” // *Bell System Technical Journal*. – Vol. 27. – Pp. 379–423 & 623–656, July & October, 1948.
2. Thomas M. Cover, Joy A. Thomas. *Elements of information theory*, 2nd Edition. – New York: Wiley–Interscience, 2006. ISBN 0471241954.
3. Claude E. Shannon. *Warren Weaver. The Mathematical Theory of Communication*. Univ of Illinois Press, 1963.
4. Tsvetkov V.Ya. Informatsionnaya neopredelennost' i opredelennost' v nauках ob informatsii // *Informatsionnyye tekhnologii*. – 2015. – №1. – S. 3–7.
5. Winner N. *Cybernetics or Control and Communication in the Animal and the Mashine*. The Technology Press and John Wiley & Soris Inc. – New York – Herman et Cie, Paris, 1948. – 194 p.
6. Hartley R.V.L. “Transmission of Information” // *Bell System Technical Journal*, July, 1928 – Pp. 535–541.
7. Wiener N. *Cybernetics, or control and communication in an animal and a machine*. – M.: Soviet Radio, 1958.
8. Carnap R. Empiricism, semantics, and ontology // *Revue internationale de philosophie*. – 1950. – C. 20–40.
9. Bar-Hillel Y., Carnap R. *An Outline of a Theory of Semantic Information* // Bar-Hillel Y. *Language and Information: Selected Essays on Their Theory and Application*. – London : Addison-Wesley, 1964. –P. 221 – 274.
10. Floridi L. (ed.). *The Blackwell guide to the philosophy of computing and information*. – John Wiley & Sons, 2008.
11. Kolmogorov A.N. Tri podkhoda k opredeleniyu kolichestva informatsii // *Problemy peredachi informatsii*. – 1965. – T. 1. Vyp. 1. – S. 3–11.
12. Shreyder YU.A. O semanticheskikh aspektakh teorii informatsii // *Informatsiya i kibernetika*. – M.: Sovetskoye radio, 1967. – S. 15–47.
13. Ivannikov A.D., Tikhonov A.N., Tsvetkov V.YA. *Osnovy teorii informatsii*. – M.: Maks Press, 2007. – 356s.
14. Adriaans P., van Benthem J. *Philosophy of information* // *Handbook of the Philosophy of Science*. – 2008. – № 8.
15. Tsvetkov V.Ya. The K.E. Shannon and L. Floridi’s amount of information // *Life Science Journal*. – 2014. – № 11 (11). Pp. 667–671.
16. Narin'yani A.S. NE-factory: kratkoye vvedeniye // *Novosti iskusstvennogo intellekta*. – 2004. – Vyp. № 2. – M.: KOMKNIGA, 2006. – S. 52–63.
17. Tsvetkov V.Ya. Dichotomous Systemic Analysis // *Life Science Journal*. – 2014. – № 11(6). – Pp. 586–590.
18. Tsvetkov V.Ya. Worldview Model as the Result of Education // *World Applied Sciences Journal*. – 2014. – № 31 (2). – P. 211–215.
19. Narin'yani A.S. Nedoopredelennyye modeli i operatsii s nedoopredelennymi znacheniyami. – Preprint VTS SO AN SSSR. – 1982. – № 400.
20. Schrödinger Erwin. *What is Life – the Physical Aspect of the Living Cell*. Cambridge University Press, 1944.
21. Brillouin, Leon: (1953) “Negentropy Principle of Information”, *J. of Applied Physics*. V. 24(9). – Pp. 1152–1163.
22. Mahulikar, S.P. & Herwig, H. (2009). «Exact thermodynamic principles for dynamic order existence and evolution in chaos», *Chaos, Solitons & Fractals*. V. 41(4). Pp. 1939–1948.
23. URL: <https://en.wikipedia.org/wiki/Negentropy> (data view: 20.05.2017).
24. Measures of Non-Gaussianity. URL: <http://fourier.eng.hmc.edu/e161/lectures/ica/node4.html> (data view: 20.05.2017).

Elsukov P.Yu.,

Ph.D., Senior Researcher,

e-mail: elsukov_p@bk.ru,

Federal State Institution of Science Institute of Energy Systems. LA Melentyeva Siberian Branch of the Russian Academy of Sciences (ESI SB RAS), Irkutsk, Russia

The article analyzes two types of information. The first kind of information is related to entropy and the work of K. E. Shannon. This information reduces uncertainty. The second type of information is associated with negentropy and the works of N. Winner and L. Floridi. The article describes the difference between the concepts of "ignorance" and "knowledge". Paper describes an example showing the difference between the amount of information for KE Shannon and N. Winner. The article describes a method for estimating negentropy based on "non-Gaussian".

Keywords: information science, information theory, amount of information, knowledge, not knowledge, information uncertainty