

ОБЩИЕ ВОПРОСЫ GENERAL QUESTIONS

УДК 004.6; 528; 004.8

В. Я. Цветков, д-р техн. наук, проф., советник, e-mail: cvj2@mail.ru,
Московский государственный технический университет радиотехники,
электроники и автоматики (МГТУ МИРЭА)

Информационная неопределенность и информационная определенность в науках об информации

Статья описывает информационную неопределенность и информационную определенность с позиций теории информации К. Э. Шеннона и Н. Винера. Раскрывается сущность информационной неопределенности как разновидности "не знания". Дается различие между понятиями "не знание" и "незнание". Показано, что неопределенность и "не знание" описываются интервальным рядом. Показано условие эквивалентности сообщений по Шеннону сообщениям по Винеру—Флориди. Вводится понятие информационной конструкции как альтернатива понятию информационное сообщение. Раскрывается понятие измеримости информационной конструкции.

Ключевые слова: науки об информации, теория информации, знание, не знание, информационная определенность, информационная неопределенность, информационные конструкции, информационная измеримость, информационные модели, содержательность сообщения, информационные единицы, информация

Введение

Термин "науки об информации" является более общим, чем термин "теория информации". С одной стороны, он включает информатику и различные науки об информационных технологиях. С другой стороны, теория информации до настоящего времени не сформирована и существует два разных ее направления.

Первое направление основано на работе К. Э. Шеннона "математическая теория связи" [1]. Под "информацией" в этой теории понимают "нечто", что уменьшает неинформированность и неопределенность, а именно, уменьшает "незнание". Количество информации в этой теории называют информационной емкостью носителя информации, сообщения или иной информационной конструкции безотносительно к содержанию. Ее измеряют в байтах и битах безотносительно к тому, что в этих байтах содержится или не содержится ничего.

Винер называет теорию К. Э. Шеннона "статистической теорией количества информации" [2, 3]. "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" [2]. "При этом Фишер исходил из статистической теории, а Шеннон — из проблемы кодирования информации" [2].

Первоначально К. Э. Шеннон не говорил о своей работе как о теории информации и отмечал что "*семантические аспекты информации не релевантны техническим проблемам связи*" [1]. Однако позже, в 1964 г., появилась совместная работа Шеннона и Уивера [4], в которой точка зрения "теории информации на основе математической теории связи" развивалась как основная. В ней рассматривались "три уровня" проблемы коммуникации: техническая проблема передачи сообщений, семантическая проблема, проблема эффективности передачи сообщений. Следует подчеркнуть, что речь идет, как и в первой работе, о передаче сообщений, а не 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 на этот вопрос не был полным и касался точности полученных символов в сравнении с исходными.

Такая точка зрения на теорию информации не была однозначно принята научным сообществом, и стали появляться работы, отражающие другие точки зрения. Вторым подходом развит в работах Н. Виннера [2, 3], Р. Карнапа [5], Л. Флориди [6], А. Н. Колмогорова [7], Ю. А. Шрейдера [8], А. А. Харкевича [9] и др. В этом подходе под информацией понимают содержание информационной конструкции безотносительно к ее объему и часто рассматривают это в аспекте семантики. Эту теорию часто называют

семантической теорией информации. Под "информацией" в этой теории понимают "нечто", что содержит знание и увеличивает в итоге знание получателя.

Оценка по Шеннону и по Винеру

По Шеннону за единицу количества информации принимается количество информации, передаваемое при одном выборе между двумя равновероятными альтернативами. По Винеру "Информация — обозначение содержания, полученного от внешнего мира в процессе приспособления к нему" [3]. Для пояснения различия в оценке количества информации по Шеннону и по Винеру возьмем пример из работы [10]. Рассмотрим три фразы:

"казнить нельзя, помиловать";
"казнить, нельзя помиловать";
"нить, ватнепомльзя казило".

По Шеннону все три фразы содержат одинаковое количество информации, поскольку имеют одинаковое число символов, содержащих равное число битов.

По Виннеру и Флориди первая и вторая фразы противоположны по смыслу, а третья бессмысленна, поскольку не содержит полезной информации. Критерий полезности можно использовать по Харкевичу [9].

По Шеннону информация уменьшает неопределенность, по Виннеру и Флориди — информация несет полезное содержание или содержит знание.

Дихотомия пары "не знание — знание"

Дихотомический анализ является одним из основных подходов [11], применяемых в системном анализе, в проектировании и при декомпозиции проектов. Он применим и при анализе содержания информационных сообщений.

По Шеннону информация уменьшает "наше незнание", по Виннеру и Флориди — информация увеличивает "наше знание". Увеличение знания и уменьшение незнания — не всегда одно и то же, что будет показано ниже. При этом термин "не знание" отличается от "незнания". Незнание обозначает одну сущность или элемент множества. "Не знание" обозначает совокупность сущностей или множество.

Дихотомическая пара "не знание — знание" представляет собой оппозицию [12, 13] и задает шкалу анализа для исследования содержательности информационных сообщений или других информационных конструкций.

Говоря о "не знании", следует остановиться на работе Нариньяни о "НЕ-факторах" [14], которая берет корни в его более ранних работах 1982 г. При этом мы упрощаем постановку задачи, поскольку Нариньяни говорит о знаниях в аспекте области

искусственного интеллекта, а мы переносим его рассуждения в область теории информации.

"НЕ-факторы отнюдь не нечто диковинное, — они "встроены" в нашу Картину мира как основная составляющая знаний, являясь их материей и строительным материалом" [14]. Следует отметить, что "картина мира" или "модель картины мира" может служить критерием оценки результата научных исследований или образования [15, 16].

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

В ряде работ Нариньяни формирует идею *недоопределенности* как общего явления. При переносе этого подхода в область теории информации мы приходим к 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" [17]. Это отражает различие между "не знанием" и "незнанием", а также факт того, что информационная неопределенность может уменьшаться.

Здесь следует напомнить, что в "четкой логике" и в "четкой" физике или математике одно качественное явление (фактор) характеризуется одним количественным значением. Например говорят, что поезд прошел 400 км; в текущий момент времени он идет со скоростью 60 км/ч; он тормозит с ускорением $0,1 \text{ м/с}^2$. Это означает парадигму "один фактор — одно числовое значение", т. е. человек с вероятностью 1 определяет это значение.

При отсутствии уверенности человек задает интервал значений. Например, говорят, что поезд прошел не менее 250 км и не более 400 км; в текущий момент времени он идет со скоростью от 45 до 65 км/ч; он тормозит с ускорением $0,1 \pm 0,01 \text{ м/с}^2$. Это соответствует парадигме "один фактор — ряд числовых значений" или "один фактор — множество числовых значений". Такая ситуация характерна для "НЕ-факторов".

Работы в области "НЕ-факторов" относят к области плохо формализуемых явлений. Это направление в последнее время подвергается исследованию и изучению. Однако в неявной форме в науке придерживаются неформального правила "равенства категорий", которое восходит еще к Аристотелю. Согласно этому правилу сопоставлять можно величины, относящиеся к одной категориальной группе. Пространственные точечные величины сопоставляют с точечными и на этой основе решают задачи. Площадные величины сопоставляют с площадными, линейные с линейными и т. д.

При введении "не-фактора" или "не знания" оказывается, что оппозиционные факторы и "не знания" могут относиться к разным категориям. Например,

"человек — не человек". К области "не человек" можно отнести камень, автомобиль, бактерию, дерево и даже компьютерную программу, т. е. совокупности из разных несопоставимых или мало сопоставимых между собой категориальных групп. Это еще более усугубляет "малоформализуемость" таких явлений.

Модели неопределенности

Для пояснения различия между подходами Шеннона и Винера рассмотрим еще один пример. Допустим, точка M лежит в плоскости и содержится в первом квадранте.

По Шеннону все сообщения, уменьшающие неопределенность, содержат информацию, т. е. информационно измеримы. Информационно измеримой назовем такую информационную конструкцию, которая кроме качественной оценки может иметь одну или ряд количественных оценок, позволяющих сравнивать ее с другими информационными конструкциями.

Это сообщения (информационные конструкции) следующего характера: "Точка M не содержится во втором квадранте"; "Точка M не содержится в третьем квадранте"; "Точка M не содержится в четвертом квадранте". Каждое из перечисленных сообщений информирует получателя сообщения и уменьшает его информационную неопределенность. Однако получив только три сообщения, получатель методом исключения (так как квадрантов всего четыре) определит, что точка находится в первом квадранте.

По Винеру правильным и содержательным будет одно сообщение "Точка M содержится в первом квадранте". Все остальные сообщения типа: "точка может содержаться в первом или во втором", "точка не содержится..." являются бессодержательными. Но следует отметить, что в этом примере три сообщения по Шеннону, уменьшающие неопределенность, равнозначны одному содержательному сообщению по Винеру. Это означает, что интервал (интервальное значение) сообщений по Шеннону эквивалентен одному (точечное значение) сообщению по Винеру.

Можно усложнить эксперимент. Разобьем плоскость на 360 секторов. Пусть точка находится условно в 360-м секторе. В этом случае 359 сообщений по Шеннону "точка не находится в секторе I ", $I = 1...359$, будут равнозначны одному сообщению по Винеру "точка находится в секторе 360". Этим подчеркивается разница между "не знанием" и знанием. То есть для N состояний интервал из $N - 1$ сообщений, уменьшающих информационную неопределенность "не знание", эквивалентны одному сообщению, содержащему знание. Интервал из $N - 1$ сообщений назовем полным, так как он снимает информационную неопределенность.

Это дает возможность ввести условие эквивалентности сообщений по Шеннону сообщению по Винеру. Совокупность сообщений по Шеннону будет эквивалентна содержательному сообщению по Винеру, если они образуют полный интервальный ряд сообщений. Термин "сообщение" можно заменить на термин "информационная конструкция".

Следующий пример. Можно от дискретной модели перейти к непрерывной. Пусть точка имеет координаты $X = 3$, $Y = 4$. По Винеру одно высказывание, назовем его условно (A), будет истинным, т. е. информативным " $X = 3$, $Y = 4$ ", все остальные будут неинформативными.

По Шеннону бесконечное число высказываний будут содержать количество информации, уменьшающее неопределенность, т. е. быть информативными:

Первая группа $X = 3$, $Y =$ любое, кроме 4;

Вторая группа $X =$ любое, кроме 3, $Y = 4$;

Третья группа $X =$ любое, кроме 3, $Y =$ любое, кроме 4.

Но в этих группах не будет ни одного высказывания, равнозначного высказыванию (A).

Информационные единицы как составляющие сообщений

Общим для подходов Шеннона и Винера является использование информационных единиц [18, 19], хотя в разных аспектах рассмотрения. Поэтому исследование информационных единиц представляет интерес как для наук об информации, так и для наук, в которых понятие информации применяется [20, 21].

Простейшее информационное сообщение может включать одно слово. Такое слово можно рассматривать также как информационную модель [22] или как информационную единицу [23]. Слово есть обозначение. Смысл слова или его значение раскрывается через интерпретацию данного слова, через совокупность интерпретирующих его предложений [24].

Слово (как информационная единица) по Шеннону содержит такое количество информации, которое определяется количеством информации в символах, входящих в это слово. Если переставить символы в слове, то смысл его исчезнет, но по Шеннону количество информации не изменится. Информационный объем символов определяет количество информации. Следовательно, мера Шеннона определяет информационный объем сообщения.

По Винеру информация, которую содержит слово, определяется числом интерпретируемых его предложений. Если переставить символы в слове, то смысл его исчезнет, и это влечет исчезновение интерпретируемых предложений, т. е. количество информации становится равным нулю. Следовательно, мера Винера и Флориди определяет семантику сообщения.

Заключение

Информационная неопределенность описывается интервальным рядом сообщений и определяет "не знание". Информационная определенность описывается одним содержательным сообщением или полным рядом, снимающим неопределенность, и определяет "знание".

Подход Шеннона решает две задачи. Он исследует информационную неопределенность и позволяет оценить информационную емкость сообщения (информационной конструкции). Подход Шеннона задает интервальный ряд информационных конструкций для описания некоего события. Подход Шеннона задает информационные конструкции, определяющие "не знание". Подход Шеннона можно использовать для получения истинного знания, но при этом надо использовать полный интервальный ряд сообщений, уменьшающих неопределенность. Теория информации, задаваемая подходом Шеннона, является "статистической" и не является "семантической".

Подход Виннера — Флориды исследует задачу передачи смысла в информационной конструкции. Подход Виннера — Флориды задает информационные конструкции, определяющие "знание" и фактически является основой семантической теории информации.

Список литературы

1. Shannon C. E. A Mathematical Theory of Communication // Bell System Technical Journal. 1948. July & October. Vol. 27. P. 379—423, 623—656.
2. Winner N. Cybernetics or Control and Communication in the Animal and the Machine. New York — Herman et Cie, Paris: The technology Press and John Wiley & Sons Inc. 1948. 194 p.
3. Винер Н. Кибернетика, или управление и связь в животном и машине. М.: Сов. радио, 1958.
4. Shannon C. E. and Weaver W. The Mathematical Theory of Communication. The University of Illinois Press URBANA, 1964. 132 p.

5. Carnap R. et al. An outline of a theory of semantic information. — Research Laboratory of Electronics, Technical Report. N. 247. Massachusetts: MIT, 1952. С. 221—274.
6. Floridi L. Semantic Conceptions of Information. URL: <http://plato.stanford.edu/entries/information-semantic>.
7. Колмогоров А. Н. Три подхода к определению понятия "количество информации" // Проблемы передачи информации. 1965. Т. 1, № 1. С. 3—11.
8. Шрейдер Ю. А. О семантических аспектах теории информации // Информация и кибернетика. М.: Сов. радио. 1967. С. 15—47.
9. Харкевич А. А. Избранные труды. В 3-х т. Том 3. Теория информации. Оpozнание образов. М.: Наука, 1973. 524 с.
10. Иванников А. Д., Тихонов А. Н., Цветков В. Я. Основы теории информации. М.: МаксПресс, 2007. 356 с.
11. Tsvetkov V. Y. Dichotomous Systemic Analysis // Life Science Journal 2014. N. 11 (6). P. 586—590.
12. Tsvetkov V. Y. Opposition Variables as a Tool of Qualitative Analysis // World Applied Sciences Journal. 2014. N. 30 (11). P. 1703—1706.
13. Tsvetkov V. Y. Correlative analysis and opposition variables // European Journal Of Natural History. 2014. N. 1. P. 48—52.
14. Нариньяни А. С. НЕ-факторы: краткое введение // Новости искусственного интеллекта. М.: КОМКНИГА, 2006. Вып. 2. С. 52—63.
15. Цветков В. Я. Картина мира как образовательная парадигма // European Social Science Journal. 2013. N. 10-1 (37). P. 28—34.
16. Tsvetkov V. Y. Worldview Model as the Result of Education // World Applied Sciences Journal. 2014. N. 31 (2). P. 211—215.
17. Нариньяни А. С. Недоопределенные модели и операции с недоопределенными значениями. Препринт ВЦ СО АН СССР, № 400, 1982.
18. Цветков В. Я. Семантика информационных единиц // Успехи современного естествознания. 2007. № 10. С. 103—104.
19. Цветков В. Я. Язык информатики // Успехи современного естествознания. 2014. № 7. С. 129—133.
20. Соловьёв И. В. Идеальное, формальное, материальное — в информационных сообщениях // Перспективы науки и образования. 2014. № 1. С. 51—55.
21. Лакофф Д. Когнитивная семантика // Язык и интеллект. М.: Прогресс. 1996. С. 143—184.
22. Цветков В. Я. Модели в информационных технологиях. М.: Макс Пресс, 2006. 104 с.
23. Tsvetkov V. Ya. Information objects and information Units // European Journal of Natural History. 2009. N. 2. P. 99.
24. Фреге Г. Логика и логическая семантика. М.: Аспект пресс, 2000.

V. Ya. Tsvetkov, Advisor to the Rectorat, e-mail: cvj2@mail.ru, MSTU MIREA

Information Uncertainty and Information Certainty in Information Science

This article describes the information uncertainty and certainty of information. These concepts are revealed from the standpoint of information theory KE Shannon and Wiener. The article reveals the essence of information uncertainty as a kind of "not knowing." article gives the difference between the concepts of "not knowing" and "ignorance." It is shown that the uncertainty and "not knowing" describes interval number of information structures. Displaying posts equivalence condition Shannon reportedly N. Wiener — L. Floridi. Article introduces the concept of information design as an alternative to the concept of an information message. The notion of measurability of information construction

Keywords: information science, information theory, knowledge, not knowledge, information certainty, information uncertainty, information construction, information measurability, information model, the content of messages, information units, information

References

1. Shannon C. E. A Mathematical Theory of Communication". *Bell. System Technical Journal*. 1948. July & October. V. 27. P. 379—423 & 623—656.
2. Winner N. *Cybernetics or Control and Communication in the Animal and the Mashine*. New York — Herman et Cie, Paris: The Technology Press and John Wiley & Soris Inc., 1948. 194 p.
3. Winner N. *Cibernetika, ili upravlenie i svyaz v givotnom i mashine*. M.: Sovetskoe radio, 1958.
4. Shannon C. E., Weaver W. *The Mathematical Theory of Communication*. The University of Illinois Press — URBANA. 1964. 132 p.
5. Carnap R. et al. *An outline of a theory of semantic information*. Research Laboratory of Electronics, Technical Report. N. 247. Massachusetts: MIT, 1952. P. 221—274.
6. Floridi L. *Semantic Conceptions of Information*. URL: <http://plato.stanford.edu/entries/information-semantic>.
7. Kolmogorov A. N. Tri podhoda k opredeleniyu ponyatiya "kolichestvo inforvatsii". *Problemy peredachi informatsii*. 1965. V. 1, N. 1. P. 3—11.
8. Shreyder Yu. A. *O semanticheskikh aspektakh teorii informatsii*. M.: Sovetskoe radio. 1967. P. 15—47.
9. Kharkevich A. A. *Izbranye trudy v 3-h t. T. 3. Teoriya informatsii. Opoznanie obrazov*. M.: Nauka, 1973. 524 p.
10. Ivannikov A. P., Tichonov A. N., Tsvetkov V. Ya. *Osnovy teorii informatsii*. M.: MaksPress, 2007. 356 p.
11. Tsvetkov V. Y. Dichotomous Systemic Analysis. *Life Science Journal*. 2014. N. 11 (6). P. 586—590.
12. Tsvetkov V. Y. Opposition Variables as a Tool of Qualitative Analysis. *World Applied Sciences Journal*. 2014. N. 30 (11). P. 1703—1706.
13. Tsvetkov V. Y. Correlative analysis and opposition variables. *European Journal of Natural History*. 2014. N. 1. P. 48—52.
14. Narinyani A. S. Ne-factory: kratkoe vvedenie. *Novosti iskusstvennogo intellekta*. Vyp. 2. M.: KOMKNIGA, 2006. P. 52—63.
15. Tsvetkov V. Y. Kartina mira kak obrazovatel'naya paradigma. *European Social Science Journal*. 2013. N. 10-1 (37). P. 28—34.
16. Tsvetkov V. Y. Worldview Model as the Result of Education. *World Applied Sciences Journal*. 2014. N. 31 (2). P. 211—215.
17. Narinyani A. S. *Nedoopredelennyye modeli i operatsii s nedoopredelennymi znateniyami*. Preprint VTs SO AN SSSR, N. 400, 1982.
18. Tsvetkov V. Y. Semantika informatsonnykh edinits. *Uspechi sovremennogo estestvoznaniya*. 2007. N. 10. P. 103—104.
19. Tsvetkov V. Y. Yazyk informatiki. *Uspechi sovremennogo estestvoznaniya*. 2014. N. 7. P. 129—133.
20. Soloviev I. V. Idealnoe, formalnoe, materialnoe — v informatsonnykh soobcheniyach. *Perspektivy nauki i obrazovaniya*. 2014. N. 1. P. 51—55.
21. Lakoff D. *Cognitivnaya semantika. Yazyk i intellekt*. M.: Progress. 1996. P. 143—184.
22. Tsvetkov V. Y. *Modeli v informatsonnykh technologiyyach*. M.: MaksPress. 2006. 104 p.
23. Tsvetkov V. Y. Information objects and information Units. *European Journal of Natural History*. 2009. N. 2. P. 99.
24. Frege G. *Logika i logicheskaya semantika*. M.: Aspekt hress. 2000.

ИНФОРМАЦИЯ



V Международная научно-техническая конференция
"Открытые семантические технологии проектирования
интеллектуальных систем"
Open Semantic Technologies for Intelligent Systems
OSTIS-2015

19 – 21 февраля 2015 г., Белорусский государственный университет
информатики и радиоэлектроники, Минск, Республика Беларусь

НАПРАВЛЕНИЯ РАБОТЫ КОНФЕРЕНЦИИ

- Принципы, лежащие в основе семантического представления знаний, и их унификация. Типология знаний и особенности семантического представления различного вида знаний и метазнаний. Связи между знаниями и отношения, заданные на множестве знаний. Семантическая структура глобальной базы знаний, интегрирующей различные накапливаемые знания
- Языки программирования, ориентированные на параллельную обработку семантического представления баз знаний
- Модели решения задач, в основе которых лежит обработка знаний, осуществляемая непосредственно на уровне семантического представления обрабатываемых знаний. Семантические модели информационного поиска, интеграции знаний, анализа корректности и качества баз знаний, сборки информационного мусора, оптимизации баз знаний, дедуктивного и индуктивного вывода в базах знаний, правдоподобных рассуждений, распознавания образов, интеллектуального управления. Интеграция различных моделей решения задач
- Семантические модели восприятия информации о внешней среде и отображения этой информации в базу знаний
- Семантические модели мультимодальных пользовательских интерфейсов интеллектуальных систем, в основе которых лежит семантическое представление используемых ими знаний, и унификация этих моделей
- Семантические модели естественно-языковых пользовательских интерфейсов интеллектуальных систем. Структура семантического представления лингвистических баз знаний, описывающих естественные языки и обеспечивающих решение задач понимания естественно-языковых текстов и речевых сообщений, а также задач синтеза естественно-языковых текстов и речевых сообщений, семантически эквивалентных заданным фрагментам баз знаний
- Интегрированные комплексные логико-семантические модели интеллектуальных систем, основанные на семантическом представлении знаний, и их унификация
- Различные технические платформы и варианты реализации интерпретаторов унифицированных логико-семантических моделей интеллектуальных систем, основанных на семантическом представлении знаний
- Средства и методы, основанные на семантическом представлении знаний и ориентированные на проектирование различных типовых компонентов интеллектуальных систем (баз знаний, программ, решателей задач, интерфейсов)
- Средства и методы, основанные на семантическом представлении знаний и ориентированные на комплексное проектирование различных классов интеллектуальных систем (интеллектуальных справочных систем, интеллектуальных обучающих систем, интеллектуальных систем управления, интеллектуальных робототехнических систем, интеллектуальных систем поддержки проектирования и др.)
- Прикладные интеллектуальные системы, основанные на семантическом представлении используемых ими знаний

Рабочие языки конференции: русский, белорусский, английский.

Вся необходимая информация о предстоящей и предыдущих конференциях OSTIS находится на сайте конференции <http://conf.ostis.net>