УДК 004.6, 004.9, 004.827, 004.824

## ИНФОРМАЦИОННЫЕ ЕДИНИЦЫ КАК СРЕДСТВО ПОСТРОЕНИЯ КАРТИНЫ МИРА

## Цветков В.Я.

Московский государственный технический университет радиотехники, электроники и автоматики, Москва, e-mail: cvj7@list.ru

Статья делает собой обзор применения информационных единиц в разных научных направлениях. Статья выполняет первичную систематизацию информационных единиц. Проводиться соотнесение информационных единиц с разными областями применения. Статья показывает, что информационные единицы являются разновидностью широко применяемых базовых элементов теории. Отмечено, что теория информации пока не является согласованной наукой. Отмечены различия в точках зрения К.Э. Шеннона и Н. Винера на количество информации. Описаны характеристики информационных единиц: структурная вложенность, информационное окружение, информационная определенность, эмерджентность. Показано различие между структурными и семантическими информационными единицами. Показаны примеры физических информационных единиц. Показано различие между субстанциональными, процессуальными и атрибутивными информационными единицами.

Ключевые слова: информатика, информационные единицы, структурная вложенность, информационное окружение, информационная определенность, субстанциональные информационные единицы, процессуальные информационные единицы

## INFORMATION UNITS AS A MEANS OF CONSTRUCTION WORLD PICTURES Tsvetkov V.Y.

Moscow State Technical University of Radio Engineering, Electronics and Automation, Moscow, e-mail: cvj7@list.ru

Article makes an overview of the application of information units in the various scientific fields . Article performs the initial systematization of information units. Article performs correlation information units with different areas of application . Article shows that the information units are a kind commonly used basic elements of the theory . Paper shows that information theory is not yet a coherent science. Article shows the differences in points of view KE Shannon and Wiener on the amount of information. The article reveals the characteristics of information units : structural nesting , information environment , information certainty emergence . Article shows the difference between structural and semantic information units . Article shows examples of physical units of information . Article shows the difference between the substantive, procedural and attributive information units

Keywords: computer science, information units, structural attachments, information environment, information certainty, the substantial information units, procedural information units

Во многих научных направлениях применяют: базовые элементарные частицы, атомарные объекты, элементы и т.п., для построения теорий или объяснения явлений и процессов окружающего мира. Обобщенно можно обозначить их термином базовые элементы теории. В науках об информации и науках, которые используют информацию (языкознание, лингвистика, когнитивная семантика, искусственный интеллект, управление и др.) такими базовыми элементами теории часто являются информационные единицы [1]. Следует отметить, что термин «науки об информации» является более адекватным, 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 можно условно назвать статистической теорией информации. Оно основано на работе К.Э. Шеннона «математическая теория связи» [2]. Под «ин-

формацией» в этой теории понимают «нечто», что уменьшает неинформированность и неопределенность, а именно уменьшает «незнание». Количеством информации в этой теории называют информационную емкость 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.

Второй подход развит в работах Н. Виннера [3], Р. Карнапа [4], Л. Флориди [5]. В этом подходе под информацией понимают содержание информационной конструкции безотносительно к ее объему и часто рассматривают это в аспекте семантики. Эта теория часто называется семантической теорией информации. Под «информацией» в этой теории понимают «нечто», что содержит знание и увеличивает в итоге знание получателя.

Однако общим для обеих теорий является использование информационных единиц, хотя в разных аспектах рассмотрения.

Поэтому исследование информационных единиц представляет интерес как для наук об информации, так и для наук, в которых осуществляется построение картины мира [6, 7].

Основная часть. Информационные единицы — это единицы, которые переносят порции информации безотносительно к содержанию или характеризуют содержание порции информации безотносительно к информационному объему. Как базовые элементы теории, информационные единицы (ИЕ) обладают свойством неделимости по какому-либо признаку [8]. Информационные единицы служат основой построения сложных: языковых описаний, информационных конструкций или информационных объектов [1].

Как многие информационные понятия, информационные единицы являются полисемическим, многоаспектным понятием. Поэтому для разграничения и уточнения видов информационных единиц необходимо указывать аспект их рассмотрения и область применения.

В аспекте формирования структуры ИЕ выделяют составные и простые информационные единицы. Простые ИЕ не включают в свой состав другие единицы. Составные информационные единицы включают в свой состав другие информационные единицы. Например, информационная единицы «предложение» включает информационные единицы «слово» включает информационные единицы «слово» включает информационные единицы «символы».

Для многих составных информационных единиц имеет место характеристика — структурная вложенность. Структурная вложенность информационных единиц — это не структура, а отношение иерархии компонент единицы и ее окружения.

Для многих составных информационных единиц имеет место характеристика – окружение информационной единицы. Окружение информационной единицы – это другие, связанные с ней информационные единицы и характеристики, необходимые для однозначной интерпретации информационной единицы и ее информационной определенности. Информационное окружение единицы проявляется при ее непосредственном использовании. Например, информационным окружением информационной единицы «слово» в предложении или во фразе, будут все связанные с этим словом символы и другие слова, а также такие информационные характеристики как позиция слова и вид его написания.

Для многих составных информационных единиц имеет место характери-

стика — информационная определенность [10]. Информационная определенность — это совокупность параметров ИЕ, полностью определенных на основе измерений или сбора информации. Информационная единица вместе с окружением является уникальной и обладает свойством эмерджентности. Эмерджентность составной информационной единицы означает несводимость свойств ИЕ к свойствам ее частей.

В аспекте применения выделяют: формальные, логические, физические – информационные единицы. Примером физической информационной единицы являются бит, байт, кластер. Бит (байт) характеризует информационную емкость носителя информации, а не количество информации. Эти единицы являются аналогами физических единиц измерения емкости литр, кубический сантиметр и др. Кластер характеризует физическую единицу дисковой памяти [11]. Эти характеристики применяют в компьютерных технологиях.

В аспекте отражения свойств внешнего мира, для любой предметной области, выделяют: субстанциональные, процессуальные, атрибутивные и комбинированные — информационные единицы. Субстанциональные информационные единицы характеризуют сущности, процессуальные — процессы, а атрибутивные единицы описывают свойства.

Субстанциональные информационные единицы можно рассматривать как элементарные описания объектов, фактов, явлений – окружающего мира.

Примером атрибутивной информационной единицы является реквизит в базах данных. Реквизиты – логически неделимые элементы, соотносимые с определением свойств отображаемого объекта или процесса [10].

Процессуальная информационная единица — это единица, которая описывает динамику: процесс или его часть. Примером процессуальных информационных единиц являются транзакция и бод, битрейт. Транзакцией называют неделимую последовательность операций манипулирования данными в БД. Это свойство информационной единицы.

Бод (boud) — единица скорости передачи сигнала, количество изменений информационного параметра несущего периодического сигнала в секунду. Названа по имени Эмиля Бодо, изобретателя кода Бодо — кодировки символов для телетайпов. Зачастую, ошибочно считают что бод это количество бит (переданное в секунду). В действительности же, это верно лишь для двоичного кодирования, которое используется не всегда. Бодами выражают полную емкость канала,

включая служебные символы (биты), если они есть. Эффективная же скорость канала выражается другими единицами, например, битрейтами.

Битрейт (bitrate) –скорость прохождения битов информации. Битрейт принято использовать при измерении эффективной скорости передачи информации по каналу, то есть скорости передачи «полезной информации». Онтологии [12] также можно рассматривать как сложные процессуальные информационные единицы

В аспекте информационного объема и структуры информационных объектов выделяют структурные информационные единицы. Структурные информационные единицы – это единицы, рассматриваемые в аспекте структуры информационного объекта и обладающие свойством неделимости по структурному признаку безотносительно к смысловому содержанию единицы.

Структура информационного объекта, в отличие от структурной вложенности, представляет собой схему со связями между компонентами информационной конструкции или информационного объекта, которая включает информационные единицы как компоненты. Характеристика «структура» является более полным описанием по сравнению с характеристикой «структурная вложенность».

Примером структурной единицы является символ. Он не носит смысл, а является элементом структуры текста. Символ - информационная единица, обладающая неделимостью по структурному признаку. Символ, это атомарный объект на который может быть поделен фрагмент теста или иная информационная конструкция. Как самостоятельный объект символ - формальное обозначение, которое выступает либо как представитель другого предмета, явления, действия, либо отражает самого себя. Символ не имеет смыслового значения, а является носителем информации. Главным назначением символов в информационных технологиях является определение информационного объема той информационной конструкции или информационного объекта, которые они образуют.

В аспекте содержательности выделяют семантические информационные единицы. Семантические информационные единицы — это единицы, рассматриваемые в аспекте семантической содержательности [8, 9] безотносительно к физическому объему и обладающие свойством неделимости по какому-либо смысловому признаку. Примерами таких единиц являются: слово, предложение, фраза. Семантические информационные единицы характеризуются топологией семантических значений [13].

Слово – информационная единица, обладающая неделимостью по сигнификативному смысловому признаку, способная непосредственно соотноситься с предметом отражения и указывать на него. В информационной конструкции слово - предельная смысловая составляющая, имеющая семантическое окружение, вследствие этого слово приобретает дополнительные смысловые свойства. Слово минимальная единица информации, имеющая смысловое значение. Согласно Поланьи «слова несут в себе только ранее вложенное в них значение, и, и несмотря на то, что в предложении или во фразе их значение модифицированно, оно, как правило, не открывается впервые» [14]. Слово характеризуется возможностью изменения словоформы. Это влечет изменение смысла слова в предложении в определенных пределах. Слово – сигнификативно неделимая информационная единица.

Предложение – информационная единица, обладающая неделимостью по предикативному смысловому признаку. Предложение – смысловая совокупность связанных слов, выражающая законченную мысль. В отличие от слова, для предложений представляется возможность проверки их на истинность. Это возможно за счет использования свойства предикации. Предикация – отношение содержания сообщения к действительности, осуществляемая в предложении. Предложение – предикативно неделимая информационная единица.

Фраза – информационная единица, обладающая неделимостью по контекстуальному смысловому признаку. Фраза смысловая совокупность предложений, выражающая законченную мысль. Фраза – ИЕ информационной конструкции, обладающая максимальной смысловой содержательностью. Для фразы смысловое содержание дополняется контекстом [8] и ассоциативностью. Предложение в составе фразы может обладать свойством ассоциативности, т.е. соотносится по смыслу с другим предложением. Такое свойство приводит к появлению в предложении ассоциативных связей с другими предложениями или объектами. Это свойство дополняет фразу ассоциативным смысловым содержанием. Информационная конструкция, содержащая фразы, несет максимальную смысловую нагрузку. Она включает сигнификативный смысл, предикативный смысл, ассоциативный смысл и контекстный смысл.

Информационные конструкции характеризуются также зависимостью смысла от позиции информационных единиц. Семантическая сущность информационных единиц связана с наличием признаков релятив-

ности, референциальности и ситуативности [15]. Например, изменение формы написания информационной единицы (курсив) означает референцию, то есть ссылку на ее подробное разъяснение.

Для информационных единиц существует характеристика «позиция», которая определяет их дополнительное смысловое значение. Выделяют следующие виды позиций: препозиция, интерпозиция и постпозиция. Позиция информационных единиц влияет на смысл информационной конструкции. Например, «флажковый знак» в кодах переменной длины [16] занимает постпозицию и символизирует о конце информационной единицы — предложения.

Характерным представителем интерпозиции является семантическая информационная единица «and». В текстовых выражениях она связывает другие семантические единицы (слова, предложения). В логике выполняет функции оператора связывающего логические переменные [17]. Слово, стоящее в начале предложения, начинается с заглавного символа. Это признак препозиции.

В лингвистическом аспекте выделяют: лингвистические, паралингвистические — информационные единицы. Определим традиционные (языковые) информационные единицы термином «лингвистические». К таким информационным единицам относятся: слово, предложение, фраза.

Паралингвистические информационные единицы не входят в систему лингвистических информационных единиц данного языка. Паралингвистические информационные единицы выполнять паралингвистические функции: дополнение, замещение, сочетание, потоковое воздействие [18]. Первые три функции изучают в традиционной паралингвистике, четвертая имеет место только в информационных технологиях, в частности в мультимедийных образовательных технологиях.

Функция «Дополнение» означает внесение дополнительной информации к лингвистическим информационным единицам. Например, голосовая окраска аудиосообщения.

Функция «Замещение» означает замещение лингвистических информационных единиц паралингвистическими. Например, использование мигающего восклицательного знака, окрашенного в красный цвет как символ сбоя или аварийной ситуации в компьютерной программе. Эта функция не языковая и одинаково понятна лицам, говорящим на разных языках для пользователей разных стран.

Функция «Сочетание» означает сочетание паралингвистических информаци-

онных единиц с лингвистическими информационным единицам в информационном сообщении для ускорения восприятия или усиления смысла информационного сообщения. Например, выделение в большеразмерной таблице значений цифр красным цветом, может служить дополнительным сигналом о качестве — критическом режиме или критической ситуации.

Рассмотренные три функции являются «внеязыковыми». Они не привязаны к конкретному языку, что дает основание использовать их как инструмент межьязыкового общения или сопровождения программных средств и информационных образовательных технологий.

Функция «потоковое воздействие» проявляется в мультимедийных технологиях и в первую очередь в интенсифицированных информационных потоках [19]. Например, можно показывать отдельные рисунки и дополнять показ вербальными средствами, но можно показать фильм без звукового сопровождения который будет намного понятней.

По аспекту прикладной области выделяют информационные единицы: сбора, хранения, обработки, представления, коммуникации. Например, при информационном поиске информационной единицей запроса является слово, логическая единица, микро информационная единица. В информационно-поисковых системах — словарная единица информационно-поискового языка, выраженная словом называется дескриптор. Единицей представления информационного поиска является не отдельное слово, документ произвольных форматов, то есть информационная конструкция.

В образовании применяют информационные образовательные единицы. В психолингвистике, в частности, в контент-анализе, применяют информационные единицы анализа. Информационная единица анализа — лингвистическая единица речи или элемент содержания, служащие в тексте индикатором интересующих исследователя явления [20]. При многих видах анализа и декомпозиции применяют дихотомические или оппозиционные информационные единицы [21].

Информационная единица имеет определенный смысл, если существует ее интерпретация. Интерпретировать информационную единицу — это значит связать с ней семантическое окружение, т.е. конкретизировать локальную область, называемую также областью интерпретации. С точки зрения когнитивной семантики интерпретация может включать субъективные когнитивные процедуры.

Формальное описание информационных единиц позволяет осуществлять их морфологический анализ. Морфологический анализ включает идентификацию формы информационной единицы (в простейшем случае словоформы) и приписывание форме соответствующего комплекса морфологической информации.

Заключение. В настоящее время информационные единицы представляют собой совокупность групп единиц, применяемых в различных направлениях. Пока отсутствует общая теория информационных единиц и общие принципы их построения и сопоставления. Все информационные единицы можно рассматривать как элементы информационного поля. Общим для всех ИЕ является признак неделимости информационной единицы по какому-либо критерию. Общим является то, что все информационные единицы являются инструментом отображения внешнего мира и инструментом создания научной картины мира.

## Список литературы

- 1. Tsvetkov V.Y. Information objects and information Units // European Journal of Natural History.  $-2009. \cancel{N}_{2} 2. p.99.$
- 2. Shannon C.E. A Mathematical Theory of Communication, Bell System Technical Journal, vol. 27, pp. 379–423 & 623–656, July & October, 1948.
- 3. 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. 99p.
- 4. Carnap R. et al. An outline of a theory of semantic information. Research Laboratory of Electronics, Technical Report N247, MIT, 1952. 49 p.
- 5 Floridi, L., Semantic Conceptions of Information http://plato.stanford.edu/entries/information-semantic.
- 6. Степин В.С., Кузнецова Л.Ф. Научная картина мира в культуре техногенной цивилизации. М.: ИФ РАН, 1994. 274 с.

- 7. Tsvetkov V.Y. Worldview Model as the Result of Education // World Applied Sciences Journal. 2014. 31 (2). p.211-215.
- 8. Цветков В.Я. Информационные единицы сообщений // Фундаментальные исследования. 2007. №12. С. 123 124.
- 9. V. Ya. Tsvetkov. Semantic Information Units as L. Florodi's Ideas Development // European Researcher, 2012, Vol. (25), № 7, p.1036-1041.
- 10. Поляков А.А., Цветков В.Я. Прикладная информатика. В 2-х частях / Под общ.ред. А.Н. Тихонова М.: МАКС Пресс. 2008. Ч.1-788с., Ч2- 860с.
- 11. Кудж С.А., Соловьёв И.В., Цветков В.Я. Когнитивные модели и методы. Краткий словарь-справочник. М.: МГТУ МИРЭА, 2014. 95с., электронное издание, номер гос регистрации 0321400338 от 30 января 2014г ISBN 978-5-7339-1237-0.
- 12. Kudz S.A., Soloviev I.V., Tsvetkov V.Y. Spatial Knowledge Ontologies // World Applied Sciences Journal. 2014. 31 (2). p.216-221.
- 13. Chiang I. J. Discover the semantic topology in high-dimensional data //Expert Systems with Applications.  $-2007. T.33. N_{2}.1. C.256-262.$
- 14. Polanyi M. Personal knowledge: Towards a post-critical philosophy. University of Chicago Press, 2012.
- 15. Keith Donnelan Reference and Definite Descriptions // The Philosophy of Language (3 edition), A. P. Martinich (ed.), Oxford University Press, 1996.
- 16. Бауер Ф., Гооз Г. Информатика. М.: Мир, 1976. 486 с.
- 17. Готтлоб Фреге. Логика и логическая семантика: Сборник трудов. М.: Аспект пресс. 2000.
- 18. Цветков В.Я. Паралингвистические средства в дистанционном образовании // Дистанционное и виртуальное обучение. №10. 2013. с.4- 11.
- 19. Цветков В.Я., Тюрин А.Г. Управление потоками мультимедиа в образовательном пространстве // Информатизация образования и науки. -2014.- № 1.-c.170-178.
- 20. Денисенко В.Н., Чеботарева Е.Ю. Современные психолингвистические методы анализа речевой коммуникации М.: РУДН, 2008. 258 с.
- 21. Цветков В.Я. Использование оппозиционных переменных для анализа качества образовательных услуг // Современные наукоёмкие технологии. 2008. № 1 С. 62-64.