

UDC 001.8

Semantic Information Units as L. Floridi's Ideas Development

Viktor Ya. Tsvetkov

State Institute of Information Technologies and Telecommunications, Russia
21, a structure 2, Bryusov a lane, Moscow, 125009
Dr. (technical), professor, leading research scientist
E-mail: cvj2@mail.ru

Abstract. This article describes the semantic information units. They are additional to the ideas of L. Floridi, concerning the creation of semantic theory of information. The work features semantic information units, shows semantic information units structure and their distinguishing features, describes the properties of semantic information units, defined by their relative position, justifies that the proposals develop the ideas of L. Floridi.

Keywords: information; semantic information theory; information units; semantic information units.

Введение. Особенность современных исследований в области информации.

Особенность современных исследований в области информации состоит в углубленном изучении смысловой стороны информации [1, 2, 3, 4] и информационных сообщений [5, 6]. Информационное сообщение не просто передача символов (Шеннон), а отражение и реализации накопленного опыта человечества и мировой культуры. Сущностные характеристики информации и информационного сообщения - это диалектическая связь с мышлением, методами коммуникации, которое предполагает не одностороннее (как у Шеннона) а двустороннее сотрудничество и отправляющего (источника, сервера) и принимающего (приемника, клиента)

Хартли и К.Э. Шеннон [7, 8] безуспешно пытались описать семантику сообщений с помощью Булевой алгебры. В конце концов, Шеннон пришел к выводу о том, что «семантические проблемы информации не релевантны проблемам математической теории связи». То, что он называл количеством информации, в настоящее время называют информационной емкостью [9, 10].

Основные концепции и идеи Л. Флориди

Л. Флориди рассматривает информацию не только как инструмент снятия неопределенности, что делает Шеннон. Л. Флориди разделяет и развивает точку зрения Н. Винера о том, что главное в информации это содержательность и знания. Это он обозначает термином семантика.

Семантическая теория информации по Л. Флориди это теория, в которой главным является содержательность информации, а не информационная емкость как у Шеннона.

Главным критерием наличия семантики (содержательности) в информации Л. Флориди определяет истинность информации с позиций эпистемологии [2]. Необходимо отметить работы древнегреческих философов, которые поставили проблему поиска истины (эпистеме). Это определяет научное направление, которое называется эпистемология.

В тоже время, уже тогда существовала проблема отделения, или различия, между «эпистеме» (истинным знанием) и «докса» (правдоподобным знанием). Более 2000 лет человечество решает эту проблему. Правдоподобное знание может быть заблуждением, может быть мнением и может перейти в разряд эпистеме.

Различительных критериев между правдоподобным и истинным знанием Л. Флориди не вводит. Это ослабляет его теорию. Главная слабость теории Л. Флориди – это рассмотрение информации как некой однородной совокупности.

В отличии от Хартли и К.Э. Шеннона, пытавшихся применить для описания информации простые модели Булевой алгебры, Л. Флориди применяет более сложные математические модели с использованием теории множеств и логики первого порядка [1, 2, 3].

Как Хартли и К.Э. Шеннон, он рассматривает информацию как однородную сущность без выделения в ней информационных и семантических единиц. Л. Флориди говорит о смысловом содержании информации как обязательном факторе теории информации. Но вместе с тем, он не учитывает то, что смысловое содержание информационного сообщения зависит от информационных единиц и отношений между ними.

Именно подход Винера, а не Шеннона, развивает Лучиано Флориди [11], когда делает попытку ввести «общее определение информации» (The General Definition of Information - GDI). Для этого он выдвигает следующие положения:

σ является экземпляром информации, понимаемая как смысловое содержание, если и только если:

(GDI.1) σ состоит из одного или нескольких *данных*;

(GDI.2) данные в σ *хорошо сформированы*;

(GDI.3) хорошо сформированные данные в σ являются *значимым* [11].

Л Флориди отмечает полисемию и многозначность информации. Он говорит о том, что значение смысла информация зависит от окружения и контекста. Автор данной работы полностью согласен с этим тезисом и развивает его, вводя понятия информационных единиц и информационных семантических единиц.

Информационные семантические единицы

Исходным положением является то, что смысловое содержание информации можно представить в виде некой системы элементов (информационной семантической системы), между которыми существуют отношения. Смысл информационного сообщения определяют значения элементов и конкретные отношения между ними.

Системный анализ информационных семантических единиц [12, 13] обнаруживает их функционирование как открытой системы. Совокупность информационных единиц обладает всеми системными свойствами. Это дает основание ввести термин «информационная семантическая система».

Термин «информационная семантическая система» вводится в альтернативу термину информационная система, под которым понимают системы сбора обработки и представления информации. Информационная система реализована на *физическом уровне обработки информации*.

Информационная семантическая система представляет собой структурную систему связанных семантических единиц. Информационная семантическая система реализована на *логическом уровне описания информации*.

Существует понятие сложной системы (в общей теории систем), которая рассматривает систему как совокупность однородных элементов связанных между собой, образующих подсистемы и части системы. Информационная семантическая система отличается от сложной системы принципиально тем, что состоит из качественно различных элементов. Этими элементами являются семантические информационные единицы.

Напомним, что элементом называют часть системы, которая не делима на более мелкие составляющие части. Неделимость является признаком элемента системы. Информационная семантическая система имеет несколько элементов, каждый из которых обладает своим признаком неделимости по разным критериям.

Формой существования коммуникативной информации является коммуникативная деятельность, речь, то есть оно всегда функционально, а главное процессуально. Формой существования описательной информации являются информационные единицы, хранения, записи, обмена, обработки и представления.

Семантические информационные единицы — это единицы, рассматриваемые в аспекте семантической содержательности. Выделяют следующие семантические информационные единицы [9, 12, 13]: символ, слово, предложение, фраза.

Символ — информационная единица, обладающая *неделимостью по структурному признаку*. Символ, это атомарный объект на который может быть поделен фрагмент текста. 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.

Слово — информационная единица, обладающая *неделимостью по смысловому признаку*. В тексте слово — предельная смысловая составляющая сообщения (предложения),

способная непосредственно соотноситься с предметом отражения и указывать на него; вследствие этого слово приобретает определенные смысловые свойства. Слово минимальная единица информации, имеющая смысловое значение.

Однако слово характеризуется сигнификативным смысловым значением и возможностью изменения, что влечет изменение смысла в определенных пределах. Поэтому смысл слова как независимого информационного объекта и смысл слова в предложении может различаться. Слово – сигнификативно неделимая информационная единица.

Предложение – информационная единица, обладающая неделимостью по смысловой совокупности связанных слов и выражающая законченную мысль. Предложение – представляет собой предикативное словосочетание, поэтому имеет предикативное смысловое значение. Предложение – предикативно неделимая информационная единица.

В отличие от слова для ряда предложений представляется возможность проверки их на истинность. Это возможно за счет использования свойства предикации.

Предикация – отношение содержания сообщения к действительности, осуществляемая в предложении (в отличие от словосочетания).

Фраза (фразеологическая единица) – информационная единица, обладающая неделимостью по связанности предложений и выражающая законченную мысль. *Фраза* – информационная единица сообщения, обладающая максимальной смысловой содержательностью.

Для фразы и предложения смысловое содержание дополняется контекстом. Предложение в составе фразы может обладать свойством антецедентности (ассоциативности), т.е. соотносится по смыслу с предыдущим (другим) предложением. Последнее приводит к появлению в предложении ассоциативных связей с другими предложениями или мыслимыми объектами. Это дополнительно наполняет данную информационную единицу ассоциативным смысловым содержанием. Сообщение, содержащее фразы, может нести максимальную смысловую нагрузку

В таблице 1 приведены основные семантические информационные единицы (СИЕ) с указанием критерия их делимости и основного значения. Термин основное значение обусловлен внутренней неоднозначностью информационных объектов и полисемией понятий, вытекающей из этого [4].

Обобщенной информационной единицей является знак. Это многозначное понятие, поскольку знаком можно обозначать символ или закодировать слово, предложение, фразу, ситуацию. Поэтому при анализе используют более конкретное понятие – символ, понимая под этим структурную составляющую слова.

Таблица 1.

Делимость информационных единиц

Часть сообщения	Критерий делимости	Основное значение
Знак	Структура	Собственное
Слово	Смысл свободная форма	Сигнификативное
Предложение	Смысл связанное словосочетание	Предикативное
Фраза	Фразеологическая связь	Фразеологически связанное

Важным является распределение информационных единиц и по смысловым характеристикам. Именно они являются одним из главных критериев формального определения информационных единиц.

В таблице 2 приведены основные информационные единицы, разделенные по смысловым характеристикам.

Таблица 2.

Смысловые характеристики информационных единиц

Часть сообщения	Смысловый уровень	Морфология	Контекст 1	Контекст 2
Знак	0	-	-	-
Слово	1	+	-	-
Предложение	2	+	+	-
Фраза	3	+	+	+
			Семантика	Ассоциации

Целесообразно сравнение или сопоставление различных единиц в информационном окружении. В таблице 3 приведены основные информационные единицы, выраженные через информационную единицу слово

Таблица 3.

Смысловые характеристики информационных единиц, выраженные через информационную единицу слово

Часть сообщения	Выражение
Символ	Словесный знак
Слово	Морфема
Предложение	Словесное выражение
Фраза	Словесное окружение

Следует отметить различие между смыслом и значением смысла информационных единиц. Смысл является более общим понятием, а значение смысла его реализацией.

Например, символ, как информационная единица, смысла не имеет, а служит строительным материалом для создания более сложных единиц. Он является элементарным носителем информации. По существу символ является информационной единицей без семантики. Элементарной семантической информационной единицей является слово.

Совокупность символов образует более сложную СИЕ - слово и приобретает новое качество. Слово может иметь множество смысловых значений, что определяется его формой, отношением к различным областям и связью с другими словами в предложении.

Слово имеет основное сигнификативное значение и ряд других значений в зависимости от сочетания с другими словами и формой изменения слова. Для слова может иметь место полисемия и неоднозначность смысловых значений. Это создает неопределенность, т.е. есть ситуацию которую рассматривал Шеннон. Эта неопределенность устраняется путем формирования более сложных информационных единиц.

Предложение как совокупность слов приобретает новое качество. Эта СИЕ обладает свойством предикативности. Т.е. появляется возможность проверки содержательности данной СИЕ с помощью методов логики или других видов анализа.

Совокупность предложений образует фразу и приобретает новое смысловое качество контекст. Это создает возможность замены некоторых слов их «заместителями» местоимениями, аббревиатурой, условными символами и др. Делается это для уменьшения информационной емкости содержательной части сообщения и для облегчения процесса интерпретации сообщения.

Полисемия информационных единиц возникает как результат обозначения информационной единицей, например, словом разных объектов или отнесение слова к разным областям. Так для термина «Информация» словарь Вебстер дает 17 разных определений [14]. Полисемия создает неопределенность, которая устраняется не только как у Шеннона количеством передаваемых символов и слов, но и качественным анализом смысла в сложных СИЕ: предложениях и фразах.

Предикативность предложения позволяет осуществлять его проверку на истинность. Контекст фраз позволяет выявлять *объекты* и *среду*, в которой они находятся, осуществлять

связь между ними и интерпретировать эту связь (гештальт [9]). Таким образом, содержательность сообщений определяется набором семантических информационных единиц и наличием в них нескольких смысловых уровней.

Одно из замечательных свойств СИЕ – изменение смысла от взаимного расположения информационных единиц. Выражая различные значения, информационные семантические единицы организуют ядро категории предикатных отношений. Семантическая сущность информационных единиц связана с наличием признаков релятивности, референциальности и ситуативности [9].

Например, изменение написания семантической информационной единицы (*курсив*) означает референцию, то есть ссылку на повторение этой единицы или на ее подробной разъяснение.

В отличие от «безликих символов» Шеннона, для информационных единиц существует понятие позиции, которое определяет их положение и отношение к друг другу. Это препозиция, интерпозиция и постпозиция. Позиция информационных единиц влияет на смысл сообщения и соотносит некоторые информационные единицы с определенной позицией в предложении.

Например, «флажковый знак» в кодах переменной длины [15] 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.

Слово, стоящее в начале предложения, начинается с заглавного символа. Это признак препозиции.

Каждая семантическая информационная единица находится в определенном «окружении»: предложение, фраза. В смысловом аспекте информационные единицы обладают рядом свойств, среди которых наиболее значимы: полисемия, антонимия, синонимия, омонимия.

Полисемия семантической информационной единицы – единица обладает рядом смысловых значений, реализуемых единицей в условиях реального информационного окружения.

Антонимия семантической информационной единицы – для семантической информационной единицы существуют другая единица с противоположным смыслом (опозиционная единица) [16].

Синонимия семантической информационной единицы – для семантической информационной единицы существуют другая единица с таким же смыслом.

Омонимия семантической информационной единицы – для семантической информационной единицы существуют другая единица с таким же обозначением, но другим смыслом

Важным понятием является «поля информационных единиц». Об особенности информационного поля говорится в [17]. Поля информационных единиц представляют собой группировки слов по близости значений, которые покрывают определенную тематическую или предметную область. Особенностью этих полей является дискретность или «разрывность». Это очень важное свойство, которое отличает такие поля от известных в физике непрерывных полей.

Поля информационных единиц часто предстают в виде объединения терминов. Часто такие объединения представляют собой семантические сети. Семантические сети могут объединять термины по сходству или по смежности их значений.

Поля информационных единиц включают ряд отношений для объектов поля. Это отношения иерархии, отношения подобия, отношения смежности, отношения следования и др. Подводя итог, следует отметить, что применение семантических информационных единиц обогащает семантическую теорию информации. Оно дополняет идеи Л. Флориди.

Подход Л. Флориди можно определить как создание теории на основе ее построения по принципу «сверху - вниз». Применение семантических информационных единиц является подходом проектирования или построения теории «снизу – вверх». Совместно оба подхода реализуют концепцию «встречных потоков» [18] и дополняют друг друга.

Примечания:

1. Floridi L. Philosophy of information <http://www.philosophyofinformation.net/publications/pdf/este.pdf>
2. Luciano Floridi In defence of the veridical nature of semantic information // European Journal of Analytic Philosophy. 2007. Vol. 3. No 1. p. 31–41.
3. Luciano Floridi's Website. Last updated: 20 May 2012 <http://www.philosophyofinformation.net/Welcome.html>
4. Малинников В.А., Соловьев И.В., Цветков В.Я. О классификации информатики и информации Materiály VIII mezinárodní vědecko - praktická conference «Efektivní nástroje moderních věd - 2012». - Díl 15. Právní vědy. Filosofie. Administrativa: Praha. Publishing House «Education and Science» s.r.o 2012. S. 89-95.
5. Цветков В.Я. Информационные единицы сообщений // Фундаментальные исследования. 2007. №12. С. 123–124.
6. Tsvetkov V.Ya. Logic units of information systems // European Journal of Natural History. 2009. № 2. p. 99–100.
7. Hartley R.V.L. "Transmission of Information," Bell System Technical Journal, July 1928 P. 533–541.
8. С.Е. Shannon, (1948), "A Mathematical Theory of Communication", Bell System Technical Journal, vol. 27, pp. 379–423 & 623–656, July & October, 1948.
9. Иванников А.Д., Тихонов А.Н., Цветков В.Я. Основы теории информации. М.: МаксПресс, 2007. 356 с.
10. Хелд Г. Технологии передачи данных. 7-изд. СПб.: Питер, 2003. 720 с.
11. Floridi L., Semantic Conceptions of Information First published Wed Oct 5, 2005; substantive revision Fri Jan 28, 2011 <http://plato.stanford.edu/entries/information-semantic>
12. Цветков В.Я. Семантика информационных единиц // Успехи современного естествознания. 2007. № 10. С. 103–104.
13. Цветков В.Я. Семантика сообщений в телекоммуникационных системах /электронная публикация // 2008-07 29 Федеральный портал по научной и инновационной деятельности http://www.sci-innov.ru/articles/itcs/contest_its
14. Merriam-Webster's Collegiate(r) Dictionary. - 2006 Merriam-Webster, Incorporated
15. Поляков А.А., Цветков В.Я. Прикладная информатика: Учебно-методическое пособие: В 2-х частях: Часть 1 / Под общ. ред. А.Н. Тихонова. М.: МАКС Пресс. 2008. Часть 1. 788 с., Часть 2. 860 с.
16. Цветков В.Я. Использование оппозиционных переменных для анализа качества образовательных услуг // Современные наукоёмкие технологии. 2008. № 1. С. 62–64.
17. Иванников А.Д., Кулагин В.П., Мордвинов В.А. и др. Получение знаний для формирования информационных образовательных ресурсов. М.: ФГУ ГНИИ ИТТ «Информика», 2008. 440 с.
18. Цветков В.Я., Вознесенская М.Е. Метод встречных потоков при проектировании программных продуктов // Успехи современного естествознания. 2010. №3. С. 138–139.

УДК 001.8

Семантические информационные единицы как развитие идей Л. Флориди

Виктор Яковлевич Цветков

Государственный научно-исследовательский институт информационных технологий и телекоммуникаций, Россия
 Доктор технических наук, профессор, ведущий научный сотрудник
 E-mail: cvj2@mail.ru

Аннотация. В статье описаны семантические информационные единицы. Они являются дополнением к идеям Л. Флориди по созданию семантической теории информации. Описаны особенности семантических информационных единиц. Показана вложенность семантических информационных единиц и их качественные отличия. Описаны свойства семантических информационных единиц, определяемые их взаимным расположением. Показано, что данные предложения дополняют идеи Л. Флориди.

Ключевые слова: информация; семантическая теория информации; информационные единицы; семантические информационные единицы.