

ФОРМАЛИЗАЦИЯ СЕМАНТИЧЕСКОЙ ИНФОРМАЦИИ

Дубовиков Н.М.,
Зеленко Н.Н.
ЧВУЗ «ВШБ-ИЭМ», Киев

FORMALIZATION OF SEMANTIC INFORMATION

Dubovykov M.M.
Zelienko N.N.
Higher school of business, Kiev

Аннотация

Формализована семантическая информация при помощи тензоров совокупности фундаментальных дефиниций, разработан синтаксис получения нового знания из потока информации при помощи этих тензоров посредством аналитической и синтетической процедур с возможностью верификации получаемой информации.

Abstract

Semantic information is analyzed by the tensors of mass data of fundamental definitions. New knowledge is extracted from mass data by these tensors and analytical and synthetic procedures. This method allows verifying of mass data.

Ключевые слова: информация, дефиниция, семантическая информация, анализ потока данных, верификация.

Keywords: information, definition, semantic information, data mining, verification.

Постановка вопроса. Под формализацией понимают отображение объектов некоторой предметной области с помощью символов, какого - либо языка, при этом формализация включает в себя четыре следующих элемента: 1) введение терминов исходных понятий, а также терминов основных отношений между этими понятиями, 2) введение переменных и правил построения на их основе соответствующих формул, 3) введение исходных доказуемых формул (аксиом), 4) введение правил логического вывода, позволяющих из аксиом получать производные от них доказуемые формулы (теоремы) [1].

Под информацией обычно понимают совокупность сведений об окружающем мире, о всевозможных протекающих в нем процессах, которые могут быть восприняты живыми организмами, электронными машинами и другими информационными системами являющихся объектом хранения, передачи и преобразования и измеряют в битах и производных от них единицах [2].

Вслед за Шенноном, принято считать, что понятие информации складывается из трех аспектов: синтаксического, семантического и прагматического [3].

Под семантической информацией понимают смысловой аспект информации, отражающий отношение между формой сообщения и его смысловым содержанием [1].

Синтаксический аспект информации характеризует информацию с точки зрения количества, структуры, построения передаваемых сообщений безотносительно к их смысловому содержанию [1].

Прагматический аспект информации это характеристика информации с точки зрения полезности, пригодности для решения задачи [1].

В связи с тем, что в настоящее время отсутствует адекватная и непротиворечивая теория семантической информации и ее синтаксис, разработаем формализм обеспечивающий реализацию возможности получения нового знания посредством синтаксического преобразования семантической информации. Актуальность решения этого вопро-

са еще обусловлена необходимостью повышением эффективности обработки больших баз данных и проблемами при создании искусственного интеллекта.

Анализ последних достижений и публикаций. Существуют разные подходы к измерению семантической информации. В частности в концепции И. Бар-Хиллела и Р. Карнапа сообщение понимается как пропозициональная формула, а семантическая информация измеряется числом состояний универсума, при которых эта формула ложна [4]. Это реализация идеи

Г. В. Лейбница о том, что логически истинные предложения, верные во всех возможных мирах, не могут нести фактической информации. Такой подход приводит к абсурдным выводам, на которые указал Флориди [5]. Он заметил, что согласно теории Бар - Хиллела и Карнапа: «„Треугольник имеет четыре стороны“ потому что согласно их теории семантической информации в этом противоречии заключено больше смыслового содержания, чем в условно истинном утверждении „Земля имеет только одну Луну“». Флориди назвал это «парадоксом Бар-Хиллела-Карнапа» [5].

Л. Флориди рассматривает информацию не как инструмент снятия неопределенности, что делает Шеннон [3], а разделяет и развивает точку зрения Н. Винера [6] о том, что главное в информации это содержательность и знания.

Флориди ввел понятие условно ложного предложения (contingently false sentence), представляющего собой конъюнкцию двух его составных частей, одна из которых истинная, а вторая — ложная. С точки зрения классической логики условно ложное предложение является просто ложным и несет только дезинформацию. В его собственной интерпретации количество семантической информации в сообщении определяется степенью соответствия этого сообщения ситуации (то есть тому, что происходит в данном месте и в данное время). Несоответствие возникает либо в результате бессодержательности сообщения, либо в результате его неточности. В своей теории Флориди непосредственно не использует понятие дезинформации, вместо этого он вводит понятие степени неточности условно ложных предложений. Для определения истинности атомарных предложений Флориди принимает принцип априорного всезнания [5], что само является парадоксом.

К сожалению не одна из теорий не оперирует с конкретной информацией и не переводит ее в прагматическую информацию обеспечивающую оптимизацию функционирования субъекта, человека или автомата, информационного потока данных и не имеет синтаксиса обеспечивающего получения нового знания посредством преобразования конкретной семантической информации потока данных.

В работе [7] преодолены указанные недостатки: введено понятие фундаментальных дефиниций, остенсивных (ostensive definition) или квази остенсивных (определенных через посредство измерительного прибора), создан формально математический подход, позволивший отказаться от конвенционного подхода в дефинициях и перейти к приданию им семантического смысла, кодифицирующего явление, сущность или объективную реальность, показано, что такой подход позволяет однозначно связать абстрактный, кодифицирующий смысл всех дефиниций с их фундаментальным, остенсивным смыслом, получена формула связывающая количество битов информации о явлении, сущности или объективной реальности и количеством дефиниций, определяющим это явление, сущность или объективную реальность:

$$\log_2 \left(\sum_{m=0}^n D_m \right) = n \quad (1)$$

где $D_m = C_n^m$ m - дефиниция, представляющая собой сочетание из n битов группы по m битов, дающих полную информацию о данном явлении, сущности или объективной реальности.

Указанные подходы в теории семантической информации позволяют решить главную задачу: передачу фундаментальных свойств смысловой информации, передачу ее содержания и истинности этого содержания. Эти фундаментальные дефиниции, как указано в работе [7], кодируются сочетаниями битов информации, фиксируемыми органами чувств человека. То есть, объективная реальность, данная индивидуальному человеку в его ощущениях, есть совокупность этих ощущений, о чем еще догадывался

Кант [1]. Этим и решается вопрос об истинности фундаментальных дефиниций. Они настолько истинны, насколько точны физические ощущения, возникающие при отражении объективной реальности [7]. В этом смысле семантическая информация в понятиях фундаментальных дефиниций полностью содержательна и истинна.

Цель работы. Целью работы является возможность получения нового знания посредством синтаксического преобразования семантической информации потока данных.

Изложение основного материала. Как было показано в работе [8] семантическую информацию можно рассматривать только как совокупную семантическую информацию, как сумму человеческих знаний. За вычетом дублирующей информации вида $A \wedge B = A$ и тождественной информации вида $A = B$ количество фундаментальных дефиниций совокупной семантической информации будет равно [7]:

$$\sum_{m=0}^n C_n^m = 2^n \quad (2)$$

где n - количество уникальных битов информации; C_n^m - число их сочетаний из n битов по m .

Следует учитывать, что в данном случае биты информации выступают как уникальные, а не тождественные.

Множество фундаментальных дефиниций, в этом случае, будет описываться вектором фундаментальных дефиниций [7], элементы которого есть множество сочетаний C_n^m вида $n_1, n_2, n_3, n_1 n_2, n_1 n_3, n_2 n_3, n_1 n_2 n_3, \dots, n_1 \dots n_n$. Здесь фундаментальные дефиниции вида n_n являются однобитовыми, назовем их фундаментальными битами информации.

Информация органами чувств человека воспринимается посредством сопоставления с порогом чувствительности, как с эталоном – есть сигнал «1», нет сигнала «0», а количественно в количестве порогов чувствительности. Например, фоторецепторная клетка (палочка) может быть возбуждена одним квантом света, а обонятельная клетка одной молекулой вещества [9], то есть пороговым значением будет являться один квант света или одна молекула вещества, которые и определяют разницу между двумя различными по величине сигналами. Откуда факт воздействия, например света, можно зафиксировать конъюнкцией фундаментального бита информации света и его количества, кратного числу порогов (квантов) $n_1 \wedge n_2$, а некоторую освещенную область двумерной таблицей при помощи элементарного алгоритма [10], который позволяет элементарно сравнить поступающую информацию с тезаурусом в смысле [11] и в дальнейшем эту поступившую информацию, можно использовать как прагматическую информацию [1], что подтверждает продуктивность предложенного подхода.

Так как мы имеем дело с различными потоками информации: зрительной, слуховой, обонятельной и т.д. информацией - передаваемой в организме различными путями [9], то логично разделить биты этих потоков информации, введя понятие векторов этих битов информации, соответствующих каждому из потоков информации, представляющие из себя непересекающиеся массивы данных информации, что легко реализуется аппаратно в кибернетическом устройстве. То есть фундаментальные дефиниции будут описываться массивами векторов соответствующих битов информации вида:

$$D = n_1 i \wedge n_2 j \wedge n_3 k \wedge n_4 l \quad (3)$$

где D – фундаментальная дефиниция, n – биты информации, i, j, k, l – обозначения массивов векторов (потоков) бит информации.

Соответственно векторы полной группы фундаментальных дефиниций [7]:

$$D_i = \left| \begin{array}{c} n_1 \\ n_2 \\ \dots \\ n_i \\ n_1 \wedge n_2; \dots; n_1 \wedge n_i \\ \dots \\ n_1 \wedge n_2 \wedge \dots \wedge n_{i-1} \wedge n_i \end{array} \right| \quad (4)$$

где D_i – вектор полной группы фундаментальных дефиниций для i количества битов информации о явлении, сущности или объективной реальности содержащий, естественно, 2^i членов, в каждой строке которого находятся уникальные биты информации или их все сочетания из i битов последовательно от одного до сочетания из i битов, преобразуются в их тензоры.

Так как всякой фундаментальной дефиниции соответствует кодирующая ее абстрактная дефиниция (определение, theoretical definition [1]), то в дальнейшем не будем различать эти понятия без необходимости. В этом случае аналитическая процедура получения нового знания сведется к вычислению компонентов тензоров совокупности фундаментальных дефиниций по известным фундаментальным битам информации, подробнее в [7]. Синтез нового знания будет происходить за счет добавления новых фундаментальных битов информации и последующей аналитической процедуре описанной выше, подробнее в [7].

Верификация информации сведется к выполнению условия для компонентов тензора совокупности фундаментальных дефиниций:

$$\forall x P(x_i) = 1 \quad (5)$$

где X_i – компоненты тензора совокупности фундаментальных дефиниций.

При невыполнении условия, как для случая фундаментальных дефиниций, так и для случая абстрактных дефиниций будет означать ошибку наблюдений (измерений) либо ошибку абстрактных построений.

Выводы.

1. Создан формализм позволяющий однозначно описывать семантическую информацию: терминология, понятие фундаментальных дефиниций, фундаментальных векторизированных битов семантической информации и тензоров совокупности фундаментальных дефиниций.

2. Разработан синтаксис получения нового знания из потока информации при помощи тензоров совокупности фундаментальных дефиниций, посредством аналитической и синтетической процедур.

3. Доказана возможность получения новой семантической информации путем синтаксического преобразования уже имеющейся семантической информации.

4. Разработана формальная процедура верификации получаемой информации.

Список литературы

1. Новейший философский словарь. 2009. [Электронный ресурс] – Режим доступа: http://dic.academic.ru/dic.nsf/dic_new_philosophy/402/%D0%94%D0%95%D0%A4%D0%98%D0%9D%D0%98%D0%A6%D0%98%D0%AF
2. Лидовский В. В. Теория информации: Учебное пособие / Лидовский В. В.— М.: Наука, 2003. — 112 с.
3. Shannon C.E. The Mathematical Theory of Communication / Shannon C.E., Weaver W. - Urbana: University of Illinois Press, 1949. – (reprinted in 1998)
4. Bar-Hillel Y., Carnap R., Semantic information //The British Journal for the Philosophy Science, 1953, v. 4, № 14, p. 28 – 47.
5. Floridi L. Outline of a Theory of Strongly Semantic Information // Minds and Machines, 2004, 14(2), p. 197-222.
6. Винер Н. Я — математик / Винер Н. - М.: Наука, 1964, 354 с.
7. Дубовиков Н.М. Математическая формализация процедуры получения и накопления новой информации путем анализа и синтеза // Математическая модель социально – инновационной экономики: статьи/ Дубовиков Н.М.- LAP LAMBERT Academic Publishing, 2013.- с. 68- 83.
8. Дубовиков М. М. Визначення об'єктивного виду функції багатofакторної продуктивності праці//Часопис економічних реформ науково-виробничий журнал, 2012, № 4(8), с. 6-10.
9. А. Каламкар Г.М. Молекулярные механизмы зрительной рецепции / Г.Р. Каламкар, М. А. Островский. – М.: Наука, 2002. – 279 с.
10. Дубовиков Н.М. Алгоритм определения цвета. 2013. [Электронный ресурс] – Режим доступа: <http://vshb.org.ua/algorithm-opredeleniya-tsveta>.
11. Шрейдер Ю.А. Об одной модели семантической информации // Проблемы кибернетики: сборник/ М.: Наука, 1965, вып. 13.

МЕДИКО-БИОЛОГИЧЕСКИЕ ПРОБЛЕМЫ ЭЛЕКТРОМАГНИТНЫХ ПОЛЕЙ, ОКРУЖАЮЩИХ ЧЕЛОВЕКА

Мандругин А.А.

доктор химических наук, профессор,

Московский Государственный Университет имени М.В. Ломоносова, Москва

Баранов Н.Н.

доктор технических наук, профессор,

*Объединенный Институт Высоких Температур РАН,
Москва*

MEDICAL-BIOLOGICAL PROBLEMS ELECTROMAGNETIC FIELDS THAT SURROUND A PERSON

Mandrugin A.A.

doctor of chemical Sciences, Professor,

Moscow State University M. V. Lomonosov, Moscow, Russia

Baranov N.N.

doctor of technical Sciences, Professor,

Institute for High Temperatures RAS, Moscow, Russia

Аннотация

Человек в современном мире живет в окружении электромагнитных полей, охватывающих диапазон от ультрадлинных волн с частотой близкой к нулю до ультрадлин-