Математические основы теоретической информатики. Теория нечёткой семантической информации о точке. Семантика естественного языка

А. В. Чечкин[1]

Цель. Исследовать семантику естественного языка.

Методы. Математическое моделирование. Системный и ультрасистемный анализ и синтез.

Результаты. Строится математическая модель семантики языковых конструкций, понятий естественного языка, математическая основа теоретической информатики. Определяется количественная мера семантической информации о точке. Делаются общие семантические выводы.

Ключевые слова: естественный язык и нечёткая семантическая информация о точке, семантика естественного языка.

1. Введение

Информатика – наука об информации, методах и процессах её сбора, хранения, анализа, синтеза, передачи, и использования в различных целях. Термин "информация" происходит от латинского слова "informatio", что означает сведение или сообщение о чем-либо или о ком-либо. Для человека информация связана с его естественным языком и, следовательно, с его интеллектом. Кибернетика тоже - наука об информации. Разница здесь в предназначении информации. В ***кибернетике*** имеют дело с информацией для управления одним объектом, то есть с информацией, относящейся только к одной точке, которая в силу единственности не требует себе уникального семантического указателя. Кибернетика обслуживает системы автоматического регулирования. ***Информатика*** использует информацию принципиально о разных точках, что требует принципиально разных, автономных уникальных семантических указателей этих точек. У человека кибернетика отвечает за безусловные рефлексы первой сигнальной системы, за подсознательное регулирование биомеханизмами жизнеобеспечения человека. Информатика у человека отвечает за сознательное

[1] *Чечкин Александр Витальевич* — доктор физ.-мат. наук, профессор, Военная академия РВСН имени Петра Великого, Финансовый университет при правительстве Российской Федерации, e-mail: a.chechkin@mail.ru.

Chechkin Alexander Vitalievich — Doctor of Phys.-Math. Sci., Professor, Strategic Missile Forces Military Academy named after Peter the Great, Financial University under the Government of the Russian Federation.

когнитивное поведение человека, за условные (приобретённые) рефлексы, связанные в первую очередь с естественным языком, с обучением и со второй сигнальной системой человека (с языком), [1-2]. Информатика имеет дело с интеллектуальными системами, использующие информацию о различных точках.

Понятие “чёткой семантической информации о точке” было введено и изучено в [3-4]. Это понятие строится на базе теории фильтров и ультрафильтров А. Картана и только в рамках теории классических чётких множеств. Тогда как понятие “нечёткой семантической информации о точке” требует специальной формализации, о которой пойдет речь в настоящей статье.

2. Алгебра де Моргана нечётких множеств — основа семантики естественного языка

Определим *нечёткие множества* по L.A.Zade [5-6].

Определение 2.1. Нечёткое множество A определяется *функцией принадлежности*:

$$\mu_A : X \to [0; 1] \tag{1}$$

Здесь X *опорное* или *универсальное* множество, из элементов которого определяется нечёткое множество $A = \{x : \mu_A(x) > 0, x \in X\}$. Функция принадлежности задает экспертную оценку принадлежности к нечеткому множеству A точек X. На прикладном языке нечёткое множество A определяется некоторым нечётким свойством, которым все точки A обладают, но в разной степени. Точки с $\mu_A(x) = 0, x \in X$ не принадлежат нечеткому множеству. Они *полностью* (*единогласно с точки зрения экспертной группы)* не обладают свойством A. Точки считаются принадлежащими нечеткому множеству в максимальной степени, если для них $\mu_A(x) = 1, x \in X$. Про такие точки говорят, что они *полностью или безоговорочно принадлежат* A. В этом случае подразумевается единогласное оценивание виртуальной экспертной группой точек из X на наличие у них некоторого нечёткого свойства A. Еще про такие точки говорят, что они образуют *ядро* A и пишут $KerA = \{x : \mu_A(x) = 1, x \in X\}$ Наличие ядра нечёткого множества является важным свойством нечёткого множества. Если в A имеется непустое ядро $KerA \neq \emptyset$ то нечёткое множество A называют *нормальным.* В противном случае нечёткое множество называется *ненормальным.* Наконец, про точки с дробными значениями функции принадлежности, у которых $0 < \mu_A(x) < 1, x \in X$, говорят, что точки принадлежат нечёткому множеству A *частично* и *образуют размытую границу нечёткого множества.* Другими словами, эти точки обладают свойством A, но *с оговорками, неединогласно с точки зрения экспертной группы*. На практике размытая граница нечёткого множества часто вызвана объективным наличием некоторого *пограничного эффекта* нечёткого свойства A.

Все точки нечёткого множества образуют *носитель нечёткого множества* и пишут $SuppA = \{x : \mu_A(x) > 0, x \in X\}$

Согласно определению 2.1. обычные классические чёткие подмножества X так же относятся к нечётким подмножествам и составляют *подсемейство чётких подмножеств* в семействе всех нечётких подмножеств X. Для классических чётких подмножеств выполняется определяющее их равенство $KerA = SuppA$, еще они не имеют размытой границы.

Определение 2.2. В определении 2.1 нечёткого множества можно вместо экспертной оценки (1) перейти к вероятностной оценке выполнения у точек нечёткого свойства A, согласованной с данной функцией принадлежности (1), т.е. перейти к нормированной вероятностной мере некоторой воображаемой случайной величины, отражающей субъективную статистику экспертной группы оценивания свойства A. Заметим, что в математической статистике особо выделяют гауссовские процессы с высокими (большими) выбросами [7].

Для случая конечного или счетного опорного множества X исходов дискретное распределение плотности вероятности, соответствующее данной функции принадлежности Заде (1), определяется формулой

$$p_A(x) = \frac{\mu_A(x)}{\sum_{x \in X} \mu_A(x)}, x \in X, \sum_{x \in X} p_A(x) = 1 \tag{2}$$

Для случая континуального измеримого опорного множества исходов X вероятностная мера, соответствующая измеримой функции принадлежности (1), определяется формулой

$$p_A(x) = \frac{\mu_A(x)}{\int_X \mu_A(x)dx}, x \in X, \int_X p_A(x)dx = 1 \tag{3}$$

Пример 2.1. Рассмотрим устойчивое словосочетание естественного языка A = *"Начало недели"*. ***Семантика*** этого лингвистического понятия является нечёткое подмножество A точек опорного множества X семи дней недели

$$X = \{\text{Пн}, \text{Вт}, \text{Ср}, \text{Чт}, \text{Пт}, \text{Сб}, \text{Вс}\}$$

Нечёткое множество A определяет объём лингвистического понятия *"Начало недели"*, который, в свою очередь определяется функцией принадлежности $\mu_A : X \to [0; 1]$. Пусть, например, функция принадлежности задана в табличном виде

$$\mu_A(x) = \{\frac{\mu_A(x)}{x} : x \in X\} = \{\frac{1}{\text{Пн}}, \frac{0.8}{\text{Вт}}, \frac{0}{\text{Ср}}, \frac{0}{\text{Чт}}, \frac{0}{\text{Пт}}, \frac{0}{\text{Сб}}, \frac{0}{\text{Вс}}\} \qquad (4)$$

Для конечного или счётного нечёткого множества (4) принята *польская запись* в форме суммы формальных дробе, отражающей табличное задание ее функции принадлежности (4). При этом в сумме члены с нулевыми значениями функции принадлежности принято опускать (их явно не писать, считая по умолчанию их присутствие)

$$A = \frac{1}{\text{Пн}} + \frac{0.8}{\text{Вт}} + \frac{0}{\text{Ср}} + \frac{0}{\text{Чт}} + \frac{0}{\text{Пт}} + \frac{0}{\text{Сб}} + \frac{0}{\text{Вс}} = \frac{1}{\text{Пн}} + \frac{0.8}{\text{Вт}} \qquad (5)$$

Конкретное значение функции принадлежности $\mu_A(\text{Вт}) = 0.8$ для вторника здесь отражает, например, некоторую частную статистику опроса: “Является ли вторник началом недели?”. В данном примере из каждых десяти опрошенных восемь считают вторник началом недели, а двое нет. Перейдём от функции принадлежности (5) к согласованной с (5) вероятностной (нормированной) мере (2). Тогда от экспертной оценки принадлежности точек к нечёткому множеству перейдём к случайному процессу отбора дней, относящихся к началу недели членами экспертной группы. Получим распределение вероятностей частной статистики, согласованной с (5)

$$p_{\text{Пн}} = \frac{1}{1+0.8} = \frac{5}{9}, p_{\text{Вт}} = \frac{0.8}{1+0.8} = \frac{4}{9}, p_{\text{Пн}} + p_{\text{Вт}} = 1$$

Здесь вероятности удовлетворяют полной группе событий и соответствуют некоторой частной статистике. Например, телефонному опросу людей, от которых просили: “Позвоните обязательно в начале недели”. Конечно, эта статистика частная и отражает особенность группы опрашиваемых. Семантика устойчивого слова сочетания “Начало недели” является нечёткой и представляет собой нечёткое множество (5). Здесь имеем ядро и носитель нечеткого множества

$KerA = \{\text{Пн}\}$, $SuppA = \{\text{Пн}, \text{Вт}\}$ Размытой границей этого нечеткого множества является множество $\{\text{Вт}\}$

Заметим, что в примере 2.1. опорное множество X является чётким и задаётся списком дней. Оно может рассматриваться как нечёткое множество с функцией принадлежности, совпадающей с характеристической функцией множества X, т.е. равной 1 во всех точках чёткого множества (6) или равномерным распределением вероятностей соответствующего случайного процесса $p_X(x) \equiv \frac{1}{7}$

$$A = \frac{1}{\text{Пн}} + \frac{1}{\text{Вт}} + \frac{1}{\text{Ср}} + \frac{1}{\text{Чт}} + \frac{1}{\text{Пт}} + \frac{1}{\text{Сб}} + \frac{1}{\text{Вс}} \qquad (6)$$

У чётких множеств справедливо равенство $KerX = SuppX$ и нет размытых границ.

Утверждение 2.1. Если некоторое свойство A проявляется для точек опорного множества исходов X с заранее известной плотностью распределения вероятности $p_A(x) > 0, x \in X$ задающей вероятностную меру на X то существует экспертная оценка для виртуальной экспертной группы, согласованная с данным распределением плотности вероятности, в форме функции принадлежности (1) $\mu_A : X \to [0;1]$ следующего вида:

$$\mu_A(x) = \frac{p_A(x)}{\sup_A p_A(x)}, x \in X \tag{7}$$

Доказательство. Функция принадлежности (7) безусловно (очевидно) отражает экспертную оценку нечеткого свойства A соответствующую данной статистике с плотностью распределения $p_A(x) > 0, x \in X$

Определение 2.3. Пусть имеются три нечётких множестваA, B и C заданные функциями принадлежности $\mu_A : X \to [0;1]$, $\mu_B : X \to [0;1]$ и $\mu_C : X \to [0;1]$. Тогда будем говорить и писать следующее.

1. *Отрицанием* или *дополнением* нечёткого множества A до X называется нечёткое множество $\bar{A}$ с функцией принадлежности вида

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x), x \in X \tag{8}$$

2. Нечёткое множество C называется *объединением* нечётких множеств

A, B и пишут $A \cup B = C$ если

$$\mu_C(x) = \sup_{x \in X}\{\mu_A(x); \mu_B(x)\} \tag{9}$$

3. Нечёткое множество C называется *пересечением* нечётких множеств A, Bи пишут $A \cap B = C$ если

$$\mu_C(x) = \inf_{x \in X}\{\mu_A(x); \mu_B(x)\} \tag{10}$$

4. Нечёткое множество A называется *надмножеством* для нечёткого множества B или нечёткое множество B называется *подмножеством* нечёткого множества A и пишут $A \supseteq B$ если

$$\mu_A(x) \geq \mu_B(x), x \in X \tag{11}$$

Замечание. В случае строгого включения нечётких подмножеств, когда $A \supset B$, $\mu_A > \mu_B$ и $SuppA = SuppB$, выполняется $A \cup B = A$, $A \cap B = B$.

Пример 2.2. Пусть нечёткое множество A и опорное чёткое множество X будут те же самые, что в примере 2.1. Тогда имеем, согласно (11)

$A \subseteq X$ что соответстует семантике понимания смысла: "*неделя*" $= X$ является надмножеством для $A =$ "*начало недели*". Найдем отрицание $\bar{A} =$ "*не начало недели*". Согласно (5) и (8) имеем нечёткое (ненормальное) множество

$$\bar{A} = \frac{0}{\text{Пн}} + \frac{0.2}{\text{Вт}} + \frac{1}{\text{Ср}} + \frac{1}{\text{Чт}} + \frac{1}{\text{Пт}} + \frac{1}{\text{Сб}} + \frac{1}{\text{Вс}} \tag{12}$$

Здесь нечёткое множество является семантикой нечёткого лингвистического устойчивого словосочетания "*Не начало недели*". Найдем теперь объединение двух нечётких множеств $\bar{A}, A$. Согласно (5), (9) и (12) будем иметь семантику устойчивого словосочетания естественного языка "Начало или не начало недели":

$$A \cup \bar{A} = \frac{1}{\text{Пн}} + \frac{0.8}{\text{Вт}} + \frac{1}{\text{Ср}} + \frac{1}{\text{Чт}} + \frac{1}{\text{Пт}} + \frac{1}{\text{Сб}} + \frac{1}{\text{Вс}}; \text{Здесь} A \cup \bar{A} \neq X \tag{13}$$

где $\mu_{A \cup \bar{A}}(\text{Вт}) = \sup\{\mu_A(\text{Вт}); \mu_A(\text{Вт})\} = \sup\{0,8; 0,2\} = 0,8$.

Найдем пересечение нечётких множеств $\bar{A}; A$ Согласно (5), (10) и (12) будем иметь семантику устойчивого словосочетания естественного языка "Начало и не начало недели":

$$A \cap \bar{A} = \frac{0}{\text{Пн}} + \frac{0.2}{\text{Вт}} + \frac{0}{\text{Ср}} + \frac{0}{\text{Чт}} + \frac{0}{\text{Пт}} + \frac{0}{\text{Сб}} + \frac{0}{\text{Вс}}; \text{Здесь} A \cap \bar{A} \neq \emptyset \tag{14}$$

Вывод. Формулы (13) и (14) иллюстрируют нарушение "Закона исключенного третьего", который в логике Моргана нечётких множеств не имеет место. Тогда как в булевой алгебре чётких множеств "Закон исключенного третьего" справедлив, там имеют место равенства

$$A \cup \bar{A} = X \text{и} A \cap \bar{A} = \emptyset \tag{15}$$

означающие: "Третьего не дано". В алгебре де Моргана нечётких множеств этот закон не выполняется. Интересно, что закон нарушается в точках размытой границы нечётких множеств $\bar{A}; A$, (5), (12), (13), (14).

Утверждение 2.2. Об алгебре и решётке нечётких множеств.

Пусть X — чёткое опорное множество. Рассмотрим семейство всех нечётких подмножеств $A \subseteq X$, которое определяется семейством всех функций принадлежности $\mu_A : X \to [0; 1]$. В семействе всех нечётких подмножеств рассмотрим три алгебраические операции. Две бинарные операции, *объединение* подмножеств (9) и *пересечение* подмножеств (10), а так же одну унарную операцию, *отрицание* или *дополнение* нечёткого подмножества до X (8). В этом семействе всех нечётких подмножеств рассмотрим частичный порядок (11).

Тогда семейство нечётких подмножеств будет образовывать *алгебру Моргана*, в которой выполняются все законы алгебры Буля чётких подмножеств кроме одного, “Закона исключенного третьего”, т.е. будет

$$A \cup \bar{A} \neq X \text{и} A \cap \bar{A} \neq \emptyset \tag{16}$$

Семейство нечётких подмножеств является дистрибутивной решёткой частичного порядка (11).

Доказательство. Докажем первый закон де Моргана для любых двух нечётких подмножеств $A\bar{\cup}B = \bar{A}\cap\bar{B}$. Перейдем к фукциям принадлежности. Для любых двух действительных чисел, $\mu_A(x); \mu_B(x), x \in X$ имеем очевидно

$$\begin{aligned} \mu_{A\bar{\cup}B}(x) = 1 - \mu_{A\cup B} = 1 - \sup\{\mu_A(x); \mu_B(x)\} = \\ = \inf\{1 - \mu_A(x); 1 - \mu_B(x)\} = \mu_{\bar{A}\cap\bar{B}} \end{aligned} \tag{17}$$

Аналогично можно доказать все другие законы алгебры Моргана для нечётких множеств. Докажем дистрибутивность решетки нечётких множеств. Для любых трех действительных чисел $\mu_A(x); \mu_B(x); \mu_C(x)$, $x \in X$ легко проверяется перебором равенство

$$\begin{gathered} \mu_{A\cup(B\cap C}(x) = \sup\{\mu_A(x); \inf\{\mu_B(x); \mu_C(x)\}\} = \\ = \inf\{\sup\{\mu_A(x); \mu_B(x)\}; \sup\{\mu_A(x); \mu_C(x)\}\} = \\ = \inf\{\mu_{A\cup B}(x); \mu_{A\cup C}(x)\} = \mu_{(A\cup B)\cap(A\cup C)}(x), x \in X\Box \end{gathered}$$

3. Нечёткие сведения и нечёткая информация о точке

Сформулируем с информационной точки зрения факт принадлежности точки $x_0 \in X$ нечёткому множеству $A \subseteq X$, т.е. определим нечёткое сведение о точке $x_0 \in A \subseteq X$.

Определение 3.1. Пусть задано нечеткое множество $A \subseteq X$ опорного множества X с функцией принадлежности $\mu_A : X \to [0; 1]$. Рассмотрим факт принадлежности $x_0 \in A \subseteq X$. Будем этот факт называть *нечётким сведением о точке* $x_0 \in X$ и записывать любой из двух равносильных формул

$$(\mu_A)A(x_0) = (p_A)A(x_0), x_0 \in X \tag{19}$$

Здесь функция принадлежности $\mu_A(x_0)$ выражает экспертную оценку сведения, а функция плотности вероятности $p_A(x_0)$ вероятность (достоверность) сведения, предикатный символ A в (19) совпадает с обозначением нечёткого множества A и, следовательно, совпадает с обозначением нечёткого свойства A точек этого множества.

Утверждение 3.1. Корректности определения нечёткого сведения о точке и его нечёткой семантики.

Определение 3.1 и формула (17) корректно определяют понятие нечёткого сведения о точке $x_0 \in X$ и семантики словосочетания "точка $x_0 \in X$ обладает нечётким свойством A"

Доказательство. Сопоставим (19) с формализацией обычного чёткого сведния о точке $x_0 \in \delta = A \subseteq X$ вида $(1)\delta(x) = \delta(x_0)$, где $\delta = A$ — чёткое подмножество [3-4]. Так как чёткое сведение является частным случаем нечёткого сведения с функцией принадлежности, совпадающей с характеристической функцией чёткого множества A, то имеем

$$\mu_A(x) = \chi_A(x) = \chi_\delta(x); (\mu_A)A(x_0) = (1)A(x_0) = \delta(x_0), x_0 \in X \qquad (20)$$

Это полностью согласует определение 2.4 и формулу (19) с определением чёткого сведения [3-4]. □

Определение 3.2. Семейство нечётких сведений *о точке* $x_0 \in X$ опорного множества X назовём нечёткой информацией о точке $x_0 \in X$ и будем это записывать в любой из двух равносильных формул

$$\{(\mu_A)A(x_0) = \{(p_A)A\}(x_0) = \mathfrak{I}(x_0), x \in X, \text{если} \qquad (21)$$

$\mathfrak{I} = \{(\mu_A)A\} = \{(p_A)A\}$ – фильтр нечётких подмножеств, содержащих точку $x_0 \in X$. При этом базис фильтра, подсемейство $\mathfrak{B} \subseteq \mathfrak{I}$, определяет носитель $\mathfrak{B}(x_0)$ информации $\mathfrak{I}(x_0)$.

Пример 3.1. Рассмотрим устойчивые словосочетания естественного языка, относящиеся к бинарной иерархической классификации живых систем шведского естествоиспытателя Карла Линнея или же устойчивые словосочетания языка, относящиеся к периодическому закону химических элементов русского химика Дмитрия Ивановича Менделеева. Например, такие устойчивые словосочетания как "медведь бурый", "человек разумный Homo Sapiens)" или "берёза белая" имеют общепризнанную мировую научную семантику (описательные биологические особенности строения или научно подробные ДНК-коды). Аналогично, устойчивые словосочетания "Закон Менделеева о зависимости свойств веществ от их атомного веса", "Благородные газы". Семантика этих общепринятых во всем мире научных словосочетаний известна, доступна и всегда очень интересна.

4. Заключение

Уточним базовые для теоретической информатики определения “фильтра” и “базиса фильтра” для нечётких подмножеств X. Фильтром нечётких подмножеств называется такое семейство нечётких подмножеств, для которого выполняются три аксиомы:

1) *Каждое подмножество фильтра непустое* $A \neq \emptyset$

2) *Пересечение любых двух подмножеств фильтра принадлежит фильтру* $A \cap B \in \mathfrak{I}$

3) *Каждое надмножество* $B \supset A$ любого нечёткого подмножества $A \in \mathfrak{I}$ принадлежит фильтру $B \in \mathfrak{I}$

Наконец, базисом $\mathfrak{B}$ *фильтра* $\mathfrak{I}$ называется часть фильтра $\mathfrak{B} \subseteq \mathfrak{I}$, в которой по любому подмножеству $A \in \mathfrak{I}$ найдётся подмножество $B \in \mathfrak{B}$ такое, что $B \subseteq A$. *Любой базис фильтра однозначно определяет свой фильтр.*

Сделаем общие семантические выводы.

Уточним понятие “семантика естественного языка”. В философии, в семиотике, в рамках концепции “треугольника Фреге” семантикой слова (знака) называют контент (сигнификат, смысл) и противопоставляют её денотанту (референту). В теоретической информатике под семантикой сведения о точке понимают принесённую этим сведением информацию о точке. Выделим два методологических принципа описания семантики словарных слов-понятий. Принцип неединственности описания (Семантика, взятая из разных словарей) и принцип дополнительных описаний (Семантика дополнений из разных словарей). Термины “устойчивое словосочетание или устойчивое слово-понятие” подчёркивают только широкую, именно всеми признанную их семантику. Заметим, что у человека в центральной нервной системе имеются две сигнальные системы сенсориума. Первая система — это когда его рецепторы органов чувств воспринимают сигналы от естественных объектов или процессов природы и далее происходит подсознательное реактивное жизнеобеспечивающее поведение человека. Вторая система — это когда его рецепторы органов чувств воспринимают сигналы от искусственных символов, языковых и происходит восприятие семантики этих языковых символов как подсознательного сенсорного образа точки первичной сигнальной системы и далее когнитивное целенаправленное поведение человека, требующее использования разнообразных сведений и знаний о разных точках (объектах и процессах). При этом ведущей в семантике слов и словосочетаний естественного языка выступают подсознательные образы первой сигнальной системы (первичного сенсориума), для описания которой используется общепризнанная чёткая или нечёткая семантика устойчивых слов-понятий и устойчивых словосочетаний естественного языка, например, частная относительно экспертной группы (явной или воображаемой). В обеих сигнальных системах человека ведущую роль играют специфические нейроны — аттракторы (сборщики) сведений только об одной точке. Например, так называемые, нейроны “Моей бабушки”, “Моего дома” и т.п. [1-2]. Отметим важную особую роль в когнитивных свойствах естественного языка семантического указателя точки или уникального собственного имени точки (x_0). Именно, благодаря указателю точки, происходит процесс выделения целостного восприятия той или иной сущности (некоторого объекта или некоторого процесса). Происходит сборка частностей в единое целое. Это подчёркивает объективную системную структуру мира природы.

Список литературы

[1] И.П. Павлов, *Лекции о работе больших полушарий головного мозга. Полное собрание трудов в 5-и т.: Т.1, 2, 5.*, изд. “Культура”, М., 1973.

[2] Г.С. Воронков, А.В. Чечки, “Проблемы моделирования сенсориума и языковой системы естественного интеллекта индивидуума”, *Интеллектуальные системы*, **2**:1–4 (1997), 35–54.

[3] А.В. Чечкин, *Математическая информатика*, Наука, М., 1991.

[4] А.В. Чечкин, “Математические основы теоретической информатики. Теория чёткой сематической информации о точке”, Труды XI Международной научно-практической конференции: “Современная математика и концепции инновационного математического образования”, **11**:1 (2024), 249–271.

[5] L. Zade, “Fuzzy Sets”, *Information and Control*, **8** (1965), 338–353.

[6] Л.А. Заде, *Понятие лингвистической переменной и его применение к принятию приближённых решений*, Мир, М., 1976.

[7] И.А. Козик, В.И. Питербарг, “Большие выбросы гауссовских нестационарных процессов в дискретном времени”, *Фундаментальная и прикладная математика*, **22**:2 (2018), 159–169.

Mathematical Foundations of Theoretical Computer Science (Informatics). The Theory of Fuzzy Semantic Information About a Point. Semantics of Natural Language

Introduction.

In cybernetics, they deal with information for managing a single object, that is, with information related to only one point, which, due to its uniqueness, does not require a unique semantic pointer. Cybernetics serves automatic control systems. **Computer science**(**Informatics**) uses information about fundamentally different points, which requires fundamentally different, autonomous, unique semantic pointers for these points. In humans, cybernetics is responsible for the unconditioned reflexes of the first signaling system, for the subconscious regulation of human life support biomechanisms. Computer science in humans is responsible for conscious cognitive behavior of humans, for conditioned (acquired) reflexes associated primarily with natural language, with learning and with the second human signaling system (with language), [1-2]. Computer science deals with intelligent systems that use information about various points. The concept of "clear semantic information about a point" was introduced and studied in [3-4]. This concept is based on the theory of filters and ultrafilters by A. Cartan and only within the framework of the theory of classical distinct sets. Whereas the concept of "fuzzy semantic information about a point" requires special formalization, which will be discussed in this article.

Conclusion.

Let's draw general semantic conclusions. Let us clarify the concept of "natural language semantics". In philosophy, in semiotics, within the framework of the concept of the "Frege triangle", the semantics of a word (sign) is called content (signification, meaning) and is opposed to its denotation (the referent). In theoretical computer science, the semantics of information about a point is understood as the information about a point brought by this information. Let us single out two methodological principles for describing the semantics of dictionary words-concepts. The principle of non-uniqueness of descriptions (Semantics taken from different dictionaries) and the principle of additional descriptions (Semantics of additions from different dictionaries). The terms "stable phrase or stable word-concept" emphasize only their broad, universally recognized semantics. Note that humans have two sensory signaling systems in their central nervous system. The first system is when his sensory receptors perceive signals from natural objects or natural processes, and then subconscious reactive life–sustaining behavior occurs. The second system is when his sensory receptors perceive signals from artificial symbols, language symbols, and the semantics of these language symbols are perceived as a subconscious sensory image of a point in the primary signaling system, followed by cognitive purposeful human behavior that requires the use of a variety of information and knowledge about different points (objects and processes). At the same time, the subconscious images of the first signaling system (primary sensorium) are the leading ones in the semantics of words and phrases of natural language, for which the

generally recognized clear or fuzzy semantics of stable words-concepts and stable phrases of natural language is used, for example, a private relative to an expert group (explicit or imaginary). In both human signaling systems, specific neurons play a leading role - attractors (collectors) of information about only one point. For example, the so-called neurons of “My grandmother”, “My house”, etc. [1-2]. Note the important special role of the semantic pointer of a point or the unique proper name of a point (x_0) in the cognitive properties of natural language. It is thanks to the dot pointer that the process of highlighting the holistic perception of an entity (some object or some process) takes place. The details are assembled into a single whole. This highlights the objective systemic structure of the natural world.

Keywords: almost complete predicting, predicting machine, prediction of superwords by a machine, criterion for predicting.

References

[1] I. P. Pavlov, *Lectures on the work of the cerebral hemispheres. Complete works in 5 volumes: Vol. 1, 2, 5*, Publishing house Culture, Moscow, 1973 (In Russian).

[2] G. S. Voronkov, A. V. Chechkin, “Problems of modeling the sensorium and the language system of an Individual’s natural intelligence”, *Intellektual’nye sistemy*, **2**:1–4 (1997), 35–54 (In Russian).

[3] A. V. Chechkin, *Mathematical computer science*, Nauka, Moscow, 1991 (In Russian).

[4] A. V. Chechkin, “Mathematical foundations of theoretical computer science. The theory of clear semantic information about a point”, Proceedings of the XI International Scientific and Practical Conference: “Modern mathematics and concepts of innovative mathematical education”, **11**:1 (2024), 249–271 (In Russian).

[5] L. Zade, “Fuzzy Sets”, *Information and Control*, **8** (1965), 338–353.

[6] L. Zade, *The concept of a linguistic variable and its application to making approximate decisions*, Mir, Moscow, 1976 (In Russian).

[7] I. A. Kozik, V. I. Peterbarg, “Large outliers of Gaussian nonstationary processes in discrete time”, *Fundamental and applied mathematics*, **22**:2 (2018), 159–169 (In Russian).