

УДК 004.912
DOI 10.17513/snt.40815



НЕЧЕТКАЯ МОДЕЛЬ СМЫСЛОВОГО ПРОСТРАНСТВА В ВЫЧИСЛИТЕЛЬНОЙ ТЕОРИИ СЕМАНТИЧЕСКОЙ ИНТЕРПРЕТАЦИИ

Вишняков Ю. М., Вишняков Р. Ю.

*Федеральное государственное бюджетное образовательное учреждение высшего образования
“Кубанский государственный университет”, Краснодар, Российская Федерация,
e-mail: renat.vishnyakov@mail.ru*

На социальное взаимодействие значительное влияние оказывают большие языковые модели и искусственный интеллект. Их эффективность обеспечивается большими вычислительными и финансовыми ресурсами, доступными лишь крупным корпорациям, но при этом модели требуют дообучения и обладают низкой конфиденциальностью. Однако ряд задач обработки естественного языка может решаться менее ресурсозатратным способом без обращения к большим языковым моделям – например, на основе вычислительной теории семантической интерпретации, где семантика моделируется функционалом смысла и вычислительными процедурами над функционалом смысла. Цель исследования – развитие данной теории путем дополнения ее нечеткой моделью семантического значения. В работе рассматриваются ключевые элементы теории: операция контекстного уточнения смысла, функционал смысла в нотации, подобной обратной польской записи. Теория дополняется геометрической интерпретацией в виде круга значимости и функцией значимости, которые полагаются в основу нечеткого множества. Вводится различие универсального и индивидуального множеств смыслов слова, где последнее представляет модель индивидуального смыслового пространства. Формулируется динамика обучения (расширение множества смыслов, изменение значимости с учетом забывания) и числовая мера взаимопонимания между субъектами на основе индекса сходства нечетких множеств. Использование нечеткости позволяет формализовать понятие смысла с учетом его множественности, субъективности и динамики усвоения. Предложенный подход ориентирован на создание формальных методов и алгоритмов обработки естественно-языковой информации, учитывающих субъективную природу смыслов.

Ключевые слова: естественно-языковая обработка, вычислительная теория семантической интерпретации, функционал смысла, нечеткая формальная модель, круг значимости, функция значимости, нечеткое множество, динамика обучения, взаимопонимание

A FUZZY MODEL OF THE SEMANTIC SPACE IN THE COMPUTATIONAL THEORY OF SEMANTIC INTERPRETATION

Vishnyakov Yu. M., Vishnyakov R. Yu.

*Federal State Budgetary Educational Institution of Higher Education “Kuban State University”,
Krasnodar, Russian Federation, e-mail: renat.vishnyakov@mail.ru*

Large language models and artificial intelligence have a significant impact on social interaction. Their effectiveness is ensured by large computational and financial resources available only to large corporations, while the models require fine-tuning and have low confidentiality. However, a number of natural language processing tasks can be solved in a less resource-intensive way without resorting to large language models, for example, on the basis of a computational theory of semantic interpretation, where semantics is modeled by a meaning functional and computational procedures over the meaning functional. The aim of the study is to develop this theory by supplementing it with a fuzzy model of semantic meaning. The paper examines the key elements of the theory: the operation of contextual clarification of meaning, the meaning functional in notation similar to reverse Polish notation. The theory is supplemented by a geometric interpretation in the form of a circle of significance and a significance function, which are taken as the basis of a fuzzy set. A distinction is introduced between the universal and individual sets of word meanings, the latter representing a model of the individual semantic space. The dynamics of learning (expansion of the set of meanings, change in significance taking into account forgetting) and a numerical measure of mutual understanding between subjects based on the similarity index of fuzzy sets are formulated. The use of fuzziness makes it possible to formalize the concept of meaning taking into account its multiplicity, subjectivity, and dynamics of acquisition. The proposed approach is aimed at creating formal methods and algorithms for natural language processing that take into account the subjective nature of meanings.

Keywords: natural language processing, computational theory of semantic interpretation, meaning functional, fuzzy formal model, circle of significance, function of significance, fuzzy set, learning dynamics, mutual understanding

Введение

Сегодня социальное взаимодействие формируют большие языковые модели (LLM) [1–3]. Они сопряжены с искусственным интеллектом (ИИ) [4–6], поддерживаются вычислительными и финансовыми

ресурсами, сильными крупным корпорациям. Пользователь общается с ИИ на естественном языке (ЕЯ) через интеллектуальные чат-боты типа ChatGPT [7], DeepSeek [8], LLaMA, Gemini и др., но дружелюбность общения не гарантирует конфиден-

циальность. Информация накапливается в дата-центрах, формируя цифровой портрет пользователя, и такой процесс является неотъемлемой частью глубокого машинного обучения (Deep Learning) [9]. Он явно проявляется в маркетплейсах и чат-ботах [9; 11]. LLM имеют корни в модели BERT [12], и на сегодня создано много инструментов и технологий по естественно-языковой обработке (NLP), называемых совместно Бертологией.

И тем не менее в ряде задач использование LLM избыточно и затратно, что формирует предпосылки к поиску иных подходов. Так, в работах Ю. М. Вишнякова и Р. Ю. Вишнякова [13] развивается вычислительная теория семантической интерпретации, где семантика текстовых фрагментов моделируется вычислительными процедурами. В их же работах [14; 15] представлены формальная модель семантического объекта и алгоритмы поиска текстов заданной семантической направленности, а в [16] описан программный комплекс поиска семантических следов объектов. Решения восходят к формально-грамматическому подходу работы авторов А. Ахо, Дж. Ульман [17, с. 103–241], допускают математическую оценку результатов и не требуют значительных вычислительных ресурсов.

Цель исследования – дополнить вычислительную теорию семантической интерпретации нечеткой моделью индивидуального семантического пространства, отражающей субъективный смысл слова.

Материал и методы исследования

Обсудим ключевые элементы вычислительной теории семантической интерпретации. Пусть задана цепочка q слов a_i , представляющая целостный по смыслу текстовый фрагмент (предложение или его часть):

$$q = a_1 a_2 \dots a_n. \quad (1)$$

Целостность означает, что все слова входят в словосочетания и образуют отношение непосредственного подчинения на множестве слов фрагмента.

Обозначим через $S(a_i)$ множество смысловых значений слова a_i , тогда множество $S(q)$ смысловых значений q (1) в общем виде представляется функционалом Φ вида

$$S(q) = \Phi(S(a_1), S(a_2), \dots, S(a_n)), \quad (2)$$

который назовем функционалом смысла.

Выделим в q какое-либо словосочетание ab (a – главное, b – зависимое слова) и представим смысловую зависимость (стрелка показывает направление) записью вида

$$\overline{a:b}, \quad (3)$$

Смысловое значение словосочетания есть подмножество смысловых значений его главного слова, формируемое смыслом зависимого слова (контекстом). Оно по форме напоминает условную вероятность и может быть представлено в подобной нотации:

$$S(\overline{a:b}) = S(a)|_{S(b)}. \quad (4)$$

$$S(\overline{a:b}) \subset S(a), S(\overline{a:b}) \neq S(\overline{b:a}), S(\overline{a:a}) = S(a). \quad (5)$$

Введем для словосочетания операцию контекстного уточнения смысла вида

$$S(\overline{a:b}) = S(a) \overline{\cap} S(b), \quad (6)$$

где $\overline{\cap}$ – операция контекстного уточнения смысла (стрелка задает направление выполнения операции) и второе тождество (5) указывает на ее несимметричность.

Если слово a – главное слово в нескольких словосочетаниях и b_1, b_2, \dots, b_p – зависимые слова, то контекстной связкой слова a называется запись:

$$\overline{a:\{b_1, b_2 \dots b_p\}}, \quad (7)$$

Распространив (6) на контекстную связку (7), получим пересечение множеств вида

$$S(\overline{a:\{b_1, b_2 \dots b_p\}}) = S(\overline{a:b_1}) \cap S(\overline{a:b_2}) \cap \dots \cap S(\overline{a:b_p}), \quad (8)$$

$$S(\overline{a:(b_1, \dots, b_p)}) = \cap_1^p S(\overline{a:b_i}), \quad (9)$$

Если слово a_i главное слово в q (1), то из (6) и (9) следует

$$S(q) \subset S(a_i), \quad (10)$$

Раскрывая выражение $S(q)$ (6) с учетом (6), (8) и (9), получим формульное представление функционала смысла фрагмента q в виде

$$q = \underbrace{\text{«международное»}}_a \underbrace{\text{«признание»}}_b \underbrace{\text{«образовательных программ»}}_c \underbrace{\text{«российских вузов»}}_d \underbrace{\text{«»}}_e \underbrace{\text{«»}}_f \quad (11)$$

Закодируем слова латинскими буквами, тогда смысловое значение $S(q)$ принимает вид

$$S(q) = (S(b) \bar{\cap} S(a)) \cap S(b) \bar{\cap} ((S(d) \bar{\cap} S(c)) \cap ((S(d) \bar{\cap} (S(f) \bar{\cap} S(e))))). \quad (12)$$

В построенных по правилам [17, с. 55–67] синтаксических деревьях (рис. 1) для (6), (9) и (12) из-за несимметричности операций контекстного уточнения смысла в их узлах стрелка поменяла направление на обратное, а ветвь главного слова в кусте стала последней.

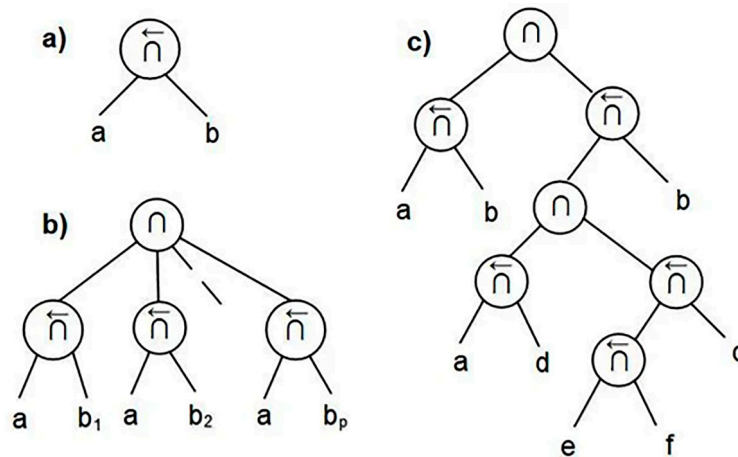


Рис. 1. Синтаксические деревья: а) операции контекстного уточнения смысла, б) контекстной связки и с) функционала смысла
Примечание: составлен авторами по результатам данного исследования

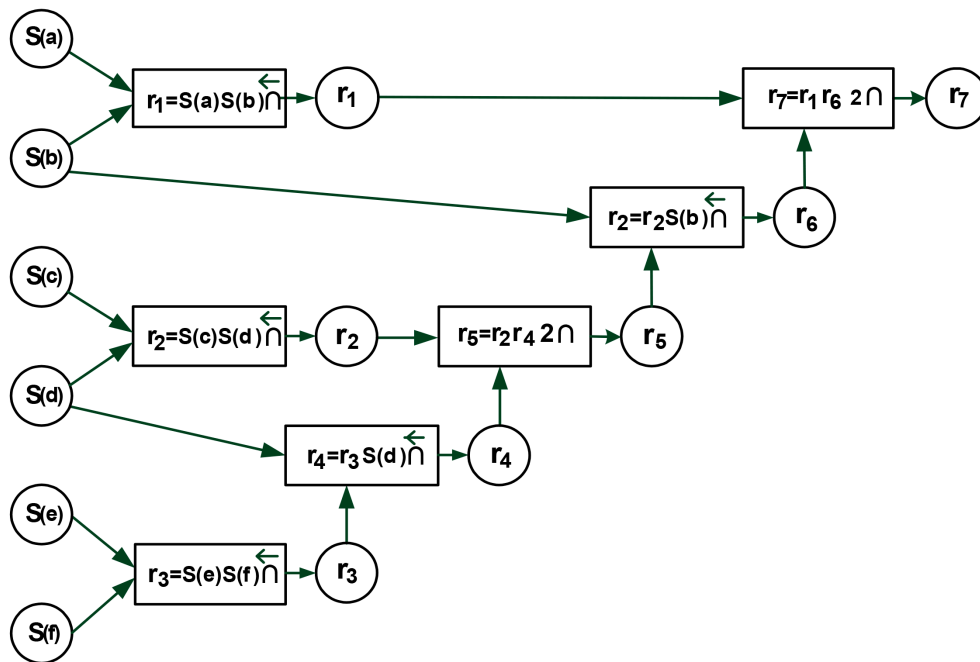


Рис. 2. Семантическая схема функционала смысла q
Примечание: составлен авторами по результатам данного исследования

Результаты обходов деревьев (рис. 1) представляют ОПЗ-подобные нотации вида

$$S(\overline{a:b}) = S(b)S(a)\bar{\cap} \quad (13)$$

$$S(\overline{a:(b_1, \dots, b_p)}) = \underbrace{S(b_1)S(a)\bar{\cap}}_1 \underbrace{S(b_2)S(a)\bar{\cap}}_2 \dots \underbrace{S(b_p)S(a)\bar{\cap}}_p \underbrace{p}_{\cap}. \quad (14)$$

$$S(q) = S(a)S(b)\bar{\cap} S(c)S(d)\bar{\cap} S(e)S(f)\bar{\cap} S(d)\bar{\cap} \cap S(b)\bar{\cap} \cap. \quad (15)$$

При вычислении (15) выполняется самая левая операция, результат сохраняется в переменной r ; замещая в строке операцию с операндами, далее процесс повторяется. Графическое представление вычислительной процедуры (рис. 2) называется семантической схемой.

В семантической схеме прямоугольные элементы (элементы смысла) представляют операции, круглые – их входы и выходы. Если текстовый фрагмент q представлен n элементами смысла, m из которых содержатся в схеме фрагмента t , то критерий семантической близости (q, t) равен m/n . Таким образом сравнение фрагментов сводится к сравнению вычислительных процедур, в которых слова представляют переменные.

В естественном языке слово и смысл образуют неразрывную пару $(x, S(x))$, слово как знак объективно (написано, произнесено), смыслом владеет субъект. В речевых актах (чтение, письмо, разговор) субъект интерпретирует слова и придает им определенные смысловые значения, возможно и новые. Например, во фрагменте (11) слово $x = \text{«признание»}$ относится к образовательным программам вуза, отвечающим международным требованиям, это смысловое значение раскрывается функционалом смысла (12), (15) или семантической схемой (рис. 2). В иных случаях слову «признание» можно придать другие смыслы (признание заслуг, признание в любви и пр.), актуализируемые другими контекстными окружениями.

Например, для слова «яблоко» множество $S(\text{яблоко})$ можно представить в виде

$$S(\text{яблоко}) = \{S(\text{яблоко}, y) | y \in Y\},$$

где $Y = \{\text{кислое, кисло-сладкое, сладкое, терпкое, ...}\}$ – вкусовая гамма.

При дополнении вкусовых качеств размерами элементы $S(\text{яблоко}, Y)$ могут быть сами множествами, например,

$$\{S(\text{яблоко сладкое маленькое}), S(\text{яблоко сладкое среднее}), \dots\}.$$

Подобным образом вводятся цвет, форма и пр., включая стилизованный логотип Apple. Но и тогда многообразие смыслов слова «яблоко» не исчерпывается. В предделе множество смыслов «человека собирающего», включающее существовавшие, существующие и будущие смыслы слова x , определим как потенциально бесконечный универсум $S_{\text{univ}}(x)$. Ограниченное множество смыслов (смысловое пространство) $S_{\text{ind}}^p(x)$ конкретного человека p формируется жизненным опытом, поэтому оно ограничено. Также смысловые пространства людей одной культурной-профессиональной группы схожи, что подметил профессор В. В. Налимов в работе [18, с. 65–109]. В дальнейшем бесконечность универсума $S_{\text{univ}}(x)$ и конечность $S_{\text{ind}}^p(x)$ считаем модельным допущением.

Поставим в соответствие множеству $S_{\text{univ}}(x)$ круг с центром в слове x . Элементы $S_{\text{univ}}(x)$ упорядочим по концентрическим окружностям: на меньшем радиусе – наиболее

привычные смысловые значения с равными значимостями, далее – менее привычные и т. д. Данный круг назовем кругом значимости смысловых значений слова.

Определим для радиусов окружностей круга значимости меру длины, для чего представим множество радиусов $R(x)$ упорядоченным множеством в виде

$$R(x) = \{r_i | i = 1, 2, \dots, n; \text{ где } r_i < r_{i+1} \text{ и } n \rightarrow \infty\}. \quad (16)$$

Введем функцию значимости смыслового значения, ставящую в соответствие каждому i -му частному смыслу слова x некоторую величину μ_i вида

$$\mu: S_{\text{univ}}(x) \rightarrow [0, 1], \quad (17)$$

$$0 \leq \mu_i \leq 1. \quad (18)$$

При $\mu_i \rightarrow 1$ значимость смысла повышается (окружность сдвигается к центру круга), при $\mu_i \rightarrow 0$ – понижается (окружность удаляется от центра). График $\mu(R)$ представлен на рис. 3.

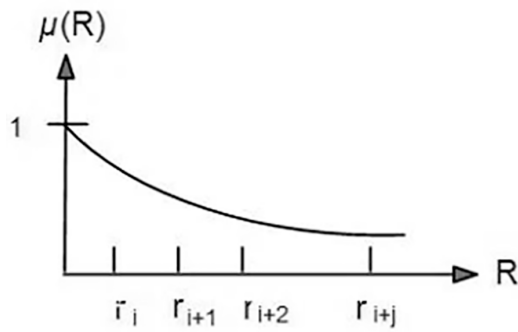


Рис. 3. Функция значимости смыслового значения

Примечание: составлен авторами по результатам данного исследования

Для $\mu_i = f(S_{\text{univ}}(x))$ можно выбрать разные формульные представления, например:

$$\mu = e^{-kR}, k > 0. \quad (19)$$

Но более привычна и понятна функция, обратная радиусу окружности круга значимости:

$$\mu = \frac{1}{1+R}. \quad (20)$$

Если для слова «яблоко» привычен смысл «сладкое яблоко», менее привычен – «кислое яблоко» и совсем не привычен «яблоко – логотип Apple» и если расположить их на концентрических окружностях $r_1 = 0,5, r_2 = 1,0, r_3 = 2,0$, то из формулы (20) получим их смысловые значения $\mu_1 = 0,67, \mu_2 = 0,5, \mu_3 = 0,33$, убывающие с удалением от центра. Круг значимости и монотонность убывания функции значимости считаются модельными допущениями.

Пусть в паре $(s_i(x), \mu_i)$ $s_i(x)$ – смысловое значение слова x и μ_i – мера его значимости. Интерпретируем μ_i смысл как степень принадлежности нечеткой логики Лютфи Заде [19]. Тогда совокупность пар образует нечеткое множество $\tilde{S}_{\text{univ}}(x)$ универсума $S_{\text{univ}}(x)$:

$$\tilde{S}_{\text{univ}}(x) = \{(s_j(x), \mu_j) | j = 1, 2, \dots, n, n \rightarrow \infty\}. \quad (21)$$

По аналогии с (21) индивидуальный круг значимости смысловых значений слова x конкретного человека p представим также нечетким множеством $\tilde{S}_{\text{ind}}^{(p)}(x)$ вида

$$\tilde{S}_{\text{ind}}^{(p)}(x) = \{(s_j(x), \mu_j) | j = 1, 2, \dots, m, m < \infty\}. \quad (22)$$

Пространства $\tilde{S}_{\text{univ}}(x)$ и $\tilde{S}_{\text{ind}}^{(p)}(x)$ различаются, поскольку субъект p не может вместить всего многообразия смыслов универсума, но, обучаясь, модифицирует $\tilde{S}_{\text{ind}}^{(p)}(x)$ добавлением новых смыслов или измене-

нием значимости существующих. При этом источником обучения является контекстное окружение слова в речевом акте. Обсудим теперь модели обучения.

Если слово x употребляется в ранее неизвестном субъекту значении S_{new} , то оно включается в $\tilde{S}_{\text{ind}}^{(p)}(x)$:

$$\tilde{S}_{\text{ind}}^{(p)}(x) \leftarrow \tilde{S}_{\text{ind}}^{(p)}(x) \cup \{(s_{\text{new}}, \mu_{\text{new}})\}. \quad (23)$$

Так, речевой акт (11) вводит в смысловое пространство $\tilde{S}_{\text{ind}}^{(p)}$ (признание) новое значение

S_{new} = "международное признание образовательных программ российских вузов".

На значимость смыслов слова влияют многие факторы: частота употребления, авторитетность источника, эмоциональная окраска и пр. Если $\mu_{\text{new}}(s)$ значимость нового понимания, то простая модель обучения с обновлением имеет вид

$$\mu_{t+1}(s) = \max(\mu_t(s), \mu_{\text{new}}(s)). \quad (24)$$

Такая усложненная модель, учитывающая частоту, представляется в виде

$$\mu_{t+1}(s) = \frac{n_t(s) \cdot \mu_t(s) + \mu_{\text{new}}(s)}{n_t(s) + 1}, \quad (25)$$

где $n_t(s)$ – количество предшествующих актов, в которых данный смысл актуализировался.

В моделях с накоплением отсутствует забывание неактуализируемых смыслов, но его можно ввести коэффициентом забывания $\lambda \in (0, 1)$, после чего получим модель с забыванием:

$$\mu_{t+1}(s) = \max(\lambda \mu_t(s), \mu_{\text{new}}(s)). \quad (26)$$

При $\lambda = 1$ забывания нет. Если $\lambda < 1$, отсутствие актуализации смысла $\mu_{\text{new}}(s) = 0$ снижает значимость $\mu_{t+1}(s) = \lambda \mu_t(s)$, а актуализация $\mu_{\text{new}}(s) > \lambda \mu_t(s)$ ее восстанавливает.

Пусть текущая значимость смысла «международное признание программ» $\mu_t = 0,5$ и в новом контексте она актуализируется $\mu_{\text{new}} = 0,9$. Тогда в модели с накоплением (24) $\mu_{t+1} = \max(0,5; 0,9) = 0,9$, а в модели с забыванием (26) при $\lambda = 0,8$

$\mu_{t+1} = \max(0,8 \times 0,5; 0,9) = \max(0,4; 0,9) = 0,9$ и перекрывает забывание. Когда смысл не актуализируется ($\mu_{\text{new}} = 0$), в модели с накоплением $\mu_{t+1} = 0,5$, а с забыванием $\mu_{t+1} = 0,8 \times 0,5 = 0,4$ (снижается).

Таким образом, в динамическом обучении индивидуальное смысловое пространство $\tilde{S}_{\text{ind}}^{(p)}(x)$ обновляясь, стремится к универсуму $\tilde{S}_{\text{univ}}(x)$, но никогда его не достигает.

Для оценивания взаимопонимания слова x субъектами p и q введем меру *сходства* $\text{Sim}(p, q, x) \in [0, 1]$ (*sim* – сокращение от англ. *similarity* – сходство). Если $\text{Sim} = 1$, то смысловые пространства субъектов $\tilde{S}_p(x)$ и $\tilde{S}_q(x)$ полностью совпадают (идеальное взаимопонимание), при $\text{Sim} = 0$ общих смыслов нет (полное непонимание). Другие значения Sim соответствуют частичному взаимопониманию. Поскольку индивидуальные смысловые пространства заданы нечеткими множествами, то степень их сходства $\text{Sim}(p, q, x)$ можно выразить мерой близости нечетких множеств коэффициентом Жаккара:

$$\text{Sim}(p, q, x) = \frac{|\tilde{S}_p(x) \cap \tilde{S}_q(x)|}{|\tilde{S}_p(x) \cup \tilde{S}_q(x)|}. \quad (27)$$

В (27) нечеткие операции пересечения и объединения определяются выражениями

$$\mu_{\tilde{S}_p \cap \tilde{S}_q}(s) = \min(\mu_p(s), \mu_q(s)), \quad (28)$$

$$\mu_{\tilde{S}_p \cup \tilde{S}_q}(s) = \max(\mu_p(s), \mu_q(s)), \quad (29)$$

а мощность нечеткого множества представляется суммой значимостей:

$$|\tilde{S}| = \sum_{s \in \text{supp}(\tilde{S})} \mu(s). \quad (30)$$

Пусть слово «признание» для субъектов A и B представлено нечеткими множествами:

$$\tilde{S}_A = \{(\text{«международное признание программ»}, 0,9), (\text{«признание заслуг»}, 0,2)\};$$

$$\tilde{S}_B = \{(\text{«международное признание программ»}, 0,7), (\text{«признание в любви»}, 0,5)\},$$

Нечеткое пересечение

$$\tilde{S}_A \cap \tilde{S}_B = \{(\text{«международное признание программ»}, \mu = \min(0,9, 0,7) = 0,7)\}$$

имеет мощность $|\tilde{S}_A \cap \tilde{S}_B| = 0,7$.

Объединение включает три смысла со значимостями 0,9, 0,2, 0,5; и его мощность

$$|\tilde{S}_A \cup \tilde{S}_B| = 0,9 + 0,2 + 0,5 = 1,6.$$

Степень взаимопонимания

$$\text{Sim}(A, B, \text{«признание»}) = 0,7 / 1,6 = 0,4375.$$

Сконструируем и проиллюстрируем на моделях динамического обучения общий алгоритм отбора претендентов в команду, который применим к большому классу задач.

Пусть некий вуз планирует выход на международный уровень через процедуру международной профессиональной аккредитации [20]. Вуз стремится пройти аккредитацию с минимальными издержками, поскольку его ресурсы ограничены. Сотрудники вуза имеют поверхностное представление о процедурах аккредитации, и в этой связи один из них (далее – «Учитель») направляется на обучения в признанный аккредитационный центр. По возвращению Учитель должен сформировать компетентную команду для решения поставленной задачи. Учитель разрабатывает общий план – обучение, оценка компетентности, отбор и реализует следующий алгоритм.

Алгоритм

Шаг 1. Выделение ключевых понятий для обучения. Выделены ключевые понятия для понимания аккредитации: «признание образовательных программ», «компетентностный подход», «внутренняя система качества» (использован функционал смысла (2), контент-анализ в сфере профессиональной аккредитации). Для каждого понятия x сконструировано универсальное смысловое пространство $S_{\text{univ}}(x)$. Например, круг значимости понятия «признание»: на ближайшей окружности ($r = 0,5$) размещены наиболее привычные для аккредитологов смыслы: «оценка соответствия стандартам», «международная валидация». На средней окружности ($r = 1,0$) – «подготовка самоотчета», «выездная экспертиза». На дальней окружности ($r = 2,0$) – «логотип агентства» (менее

значимый). По формуле (20) $\mu = 1 / (1 + R)$ вычислены значимости: $\mu_{\text{привычный}} = 0,67$, $\mu_{\text{средний}} = 0,5$, $\mu_{\text{периферийный}} = 0,33$.

Результат шага: множество смыслов S_{univ} (признание), S_{univ} (компетентностный подход), S_{univ} (внутренняя система качества) и их функции значимости μ .

Шаг 2. Начальное тестирование кандидатов. Проводится путем опроса и интервью с кандидатом p по каждому ключевому понятию x

Результат шага: начальные индивидуальные смысловые пространства в виде нечетких множеств $\tilde{S}_p(x)$ кандидатов.

Анна:

$$\tilde{S}_{\text{Анна}}(\text{«признание»}) = \{(\text{«оценка соответствия стандартам»}, 0,9), (\text{«выездная экспертиза»}, 0,2)\}.$$

Борис:

$$\tilde{S}_{\text{Борис}}(\text{«признание»}) = \left\{ \begin{array}{l} (\text{«оценка соответствия стандартам»}, 0,7), \\ (\text{«подготовка самоотчета»}, 0,4), \\ (\text{«международная валидация»}, 0,3) \end{array} \right\}.$$

Максим:

$$\tilde{S}_{\text{Максим}}(\text{«признание»}) = \{(\text{«оценка соответствия стандартам»}, 0,6)\}.$$

Шаг 3. Динамическое обучение. В течение четырех циклов обучения Учитель актуализирует ключевые смыслы, используя контексты, сформированные по функционалу смысла. Экспертные оценки забывания: для Анны и Бориса $\lambda = 0,8$; для Максима – $\lambda = 0,6$ (ускоренное забывание).

Шаг 4. Итоговое тестирование и оценка взаимопонимания. По завершению обучения для каждой пары вычисляется индекс сходства $\text{Sim}(p, q, x)$ (27)–(30). Порог включения в команду устанавливается равным 0,7.

Шаг 5. Принятие решения. Кандидаты, достигшие порога с хотя бы одним коллегой, рекомендуются в команду.

Результат работы алгоритма представлен табл. 1–3.

Итак, Анна и Борис прошли отбор, а Максим, несмотря на начальное присутствие ключевого смысла, был отсеян из-за быстрого забывания ($\lambda = 0,6$).

Таблица 1

Итоговые смысловые пространства кандидатов

Кандидат	«Оценка соответствия»	«Подготовка отчета»	«Выездная экспертиза»	«Международная валидация»	Примечание
Анна	0,90	0,80	0,90	0,80	$\lambda = 0,8$
Борис	0,90	–	0,80	0,70	$\lambda = 0,8$
Максим	0,14	0,25	0,54	0,42	$\lambda = 0,6$

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 2

Попарные индексы сходства Sim

Пара	Сумма пересечения	Сумма объединения	Sim	Достижение порога 0,7
Анна – Борис	2,40	3,40	0,706	Да
Анна – Максим	1,35	3,40	0,397	Нет
Борис – Максим	1,10	2,65	0,415	Нет

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 3

Решение о включении в команду

Кандидат	Лучший Sim (с кем)	Решение	Обоснование
Анна	0,706 (с Борисом)	Включена	Достигает порога с Борисом
Борис	0,706 (с Анной)	Включен	Достигает порога с Анной
Максим	0,415 (с Борисом)	Не включен	Ниже порога со всеми

Примечание: составлена авторами на основе полученных данных в ходе исследования

Таким образом модель динамического обучения с забыванием способна выявлять пробелы в знаниях и способность субъекта к долгосрочному удержанию понятийного аппарата предметной области и позволяет:

1. Формализовать начальные индивидуальные смысловые пространства кандидатов.
2. Смоделировать динамику обучения с учетом забывания.
3. Количественно оценить взаимопонимание с помощью индекса сходства Sim.
4. Принять объективное решение о формировании команды на основе заданного порога.

Результаты исследования и их обсуждение

Полученные результаты дополняют вычислительную теорию семантической интерпретации субъективной природой смысла слова x и позволяют конструировать модели субъективного понимания и взаимопонимания введением универсального $S_{\text{univ}}(x)$ и индивидуального $S_{\text{ind}}^{(p)}(x)$ смысловых пространств. Первое представляет потенциально бесконечный универсум всевозможных смыслов слова, второе – смысловое пространство субъекта, конечное и динамически изменяемое жизненным опытом.

Геометрическая интерпретация кругом значимости с монотонно убывающей функцией μ позволяет представить смысловые пространства нечеткими множествами и воспользоваться аппаратом нечеткой логики.

Для моделирования динамики индивидуального смыслового пространства $\tilde{S}_p(x)$ предложены модели обучения с накоплением и забыванием. Обновление происходит включением ранее неизвестных смыслов или изменением значимостей известных. В моделях с забыванием значимости неактуализируемых смыслов понижаются вплоть до полного исключения. Модель можно уточнять, учитывая частоту актуализаций, время между ними, конкуренцию между смыслами и пр.

Предложена мера взаимопонимания между субъектами p и q в виде степени сходства $\text{Sim}(p, q, x) \in [0, 1]$ и способ ее

вычисления. Здесь 1 соответствует идеальному взаимопониманию (пространства совпадают), 0 – полному непониманию (совпадения нет).

Сконструирован алгоритм создания компетентной команды, формализующий построение начальных смысловых пространств субъектов, моделирующий динамику обучения с забыванием и оценивающий взаимопонимание кандидатов. Применимость продемонстрирована примером профессиональной аккредитации образовательных программ.

Эмпирическая верификация модели путем проведения психолингвистических экспериментов, анализа корпусов текстов с экспертной разметкой, сравнения с векторным представлением слов (word embeddings) и LLM-моделями предполагает отдельное исследование ближайшей перспективы.

Заключение

В работе предложена нечеткая модель смыслового пространства, дополняющая вычислительную теорию семантической интерпретации субъективной природой смысла слова. Она включает геометрическую интерпретацию пространства (круг значимости с функцией значимости), представление пространств нечеткими множествами, динамику обучения с накоплением и забыванием, а также количественную меру взаимопонимания. На основе модели построен общий алгоритм формирования компетентных команд для широкого круга задач.

Подход допускает точную математическую оценку, не требует сравнимых с LLM-моделями вычислительных ресурсов и создает основу для построения эффективных алгоритмов и программных систем для обработки естественно-языковой информации.

Список литературы

1. Che Ting Chien, Chao Heng Chien Hybrid AI-Driven Computer-Aided Engineering Optimization: Large Language Models Versus Regression-Based Models Validated Through Finite-Element Analysis // Applied Sciences. 2025. Vol. 15. Is. 18. P. 10123. URL: <https://www.mdpi.com/2076-3417/15/18/10123> (дата обращения: 28.05.2026). DOI: 10.3390/app151810123.

2. Kang R. N., Tanaka Y., Sato T., Maeda S., Shimada H. The Development of a Japanese Version of the State Cognitive Fusion Questionnaire // *Japanese Psychological Research*. 2025. Vol. 67. Is. 3. P. 311–324. URL: <https://onlinelibrary.wiley.com/doi/10.1111/jpr.12453> (дата обращения: 28.05.2026). DOI: 10.1111/jpr.12453.
3. Guo F., Li T., Cunningham C. J. L. One year in the classroom with ChatGPT: empirical insights and transformative impacts // *Frontiers in Education*. 2025. Vol. 10. P. 1–17. URL: <https://www.frontiersin.org/journals/education/articles/10.3389/educ.2025.1574477/full> (дата обращения: 28.05.2026). DOI: 10.3389/educ.2025.1574477.
4. Ju B., Stewart J. B. Empowering Users with ChatGPT and Similar Large Language Models (LLMs): Everyday Information Needs, Uses, and Gratification // *Proceedings of the Association for Information Science and Technology*. 2024. Vol. 61. Is. 1. P. 172–182. URL: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/pr2.2.1018> (дата обращения: 28.05.2026). DOI: 10.1002/pr2.2.1018.
5. Rakovics Z., Rakovics M. Exploring the potential and limitations of large language models as virtual respondents for social science research // *Intersections: East European Journal of Society and Politics*. 2025. Vol. 10. Is. 4. P. 126–147. URL: <https://intersections.tk.hu/index.php/intersections/article/view/1326> (дата обращения: 28.05.2026). DOI: 10.17356/ieejsp.v10i4.1326.
6. Kumar H., Yoo S., Bernuy A., Shi J., Luo H., Williams J. J., Kuzminykh A., Anderson A., Kornfeld R. Large Language Model Agents for Improving Engagement with Behavior Change Interventions: Application to Digital Mindfulness // *Proceedings of the ACM on Human-Computer Interaction (CSCW)*. 2025. Vol. 9. Is. 7. P. 1–44. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12945337/> (дата обращения: 28.05.2026). DOI: 10.1145/3757619.
7. Lund B. D., Wang T., Mannuru N. R., Nie B., Shimray S., Wang Z. ChatGPT and a new academic reality: artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing // *Journal of the Association for Information Science and Technology*. 2023. Vol. 74. Is. 5. P. 570–581. URL: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.24750> (дата обращения: 28.05.2026). DOI: 10.1002/asi.24750.
8. Liu S., Chen L., Yan J., Jiang Y., Wang X., Li X., Yang Q. When DeepSeek-R1 meets financial applications: benchmarking, opportunities, and limitations // *Frontiers of Information Technology & Electronic Engineering*. 2025. Vol. 26. Is. 10. P. 1862–1870. URL: <https://link.springer.com/article/10.1631/FITEE.2500227> (дата обращения: 28.05.2026). DOI: 10.1631/FITEE.2500227.
9. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // *Proceedings – of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. June 2019. Vol. 1. P. 4171–4186. URL: <https://aclanthology.org/N19-1423/> (дата обращения: 28.05.2026). DOI: 10.18653/v1/N19-1423.
10. Зиновьева Е. С., Трапезников В. П. Международно-политическая предвзятость больших языковых моделей: критический дискурс-анализ нарративов ChatGPT, LLaMA, Gemini и DeepSeek // *Полис. Политические исследования*. 2026. № 1. С. 157–177. URL: <https://www.politstudies.ru/en/article/6368> (дата обращения: 28.05.2026). DOI: 10.17976/jpps/2026.01.11.
11. King J., Klyman K., Capstick E., Saade T., Hsieh V. User Privacy and Large Language Models: An Analysis of Frontier Developers' Privacy Policies // *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. 2025. Vol. 8. Is. 2. P. 1465–1477. URL: <https://ojs.aaai.org/index.php/AIES/article/view/36646> (дата обращения: 28.05.2026). DOI: 10.1609/aies.v8i2.36646.
12. Rogers A., Kovaleva O., Rumshisky A. A Primer in BERTology: What We Know About How BERT Works // *Transactions of the Association for Computational Linguistics*. 2020. Vol. 8. P. 842–866. URL: https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00349/96482/A-Primer-in-BERTology-What-We-Know-About-How-BERT (дата обращения: 28.05.2026). DOI: 10.1162/tacl_a_00349.
13. Вишняков Ю. М., Вишняков Р. Ю. Вычислительная семантическая интерпретация текстов научно-технического стиля // *Современные наукоемкие технологии*. 2016. № 12–2. С. 236–242. URL: <https://top-technologies.ru/article/view?id=36428&ysclid=m77urlmrjv786427462> (дата обращения: 28.05.2026).
14. Вишняков Ю. М., Вишняков Р. Ю. Формализация распознавания и идентификации семантических объектов в естественно-языковых текстовых потоках // *Известия ЮФУ. Технические науки*. 2024. № 4. С. 110–122. URL: https://izv-ti.ti.sfedu.ru/index.php/izv_tn/en/article/download/985/1172/3098 (дата обращения: 28.05.2026). DOI: 10.18522/2311-3103-2024-4-110-122.
15. Вишняков Ю. М., Вишняков Р. Ю. Поиск и идентификация текстов определенной семантической направленности в естественно-языковых потоках // *Современные наукоемкие технологии*. 2025. № 5. С. 32–40. URL: <https://top-technologies.ru/article/view?id=40387> (дата обращения: 28.05.2026). DOI: 10.17513/snt.40387.
16. Вишняков Ю. М., Вишняков Р. Ю. Программный комплекс идентификации текстов определенной семантической направленности в естественно-языковых потоках // *Современные наукоемкие технологии*. 2025. № 9. С. 29–38. URL: <https://top-technologies.ru/article/view?id=40482> (дата обращения: 28.05.2026). DOI: 10.17513/snt.40482.
17. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции / Пер. с англ. В. Н. Агафонова; Под ред. В. М. Курочкина. М.: Мир, 1978. Т. 1. 612 с. [Электронный ресурс]. URL: https://rusneb.ru/catalog/000199_000009_007597729/ (дата обращения: 28.05.2026).
18. Налимов В. В. Вероятностная модель языка. О соотношении естественных и искусственных языков. М.: Наука, 1979. 303 с. URL: <https://books.google.ru/books?id=Oob-AgAAQBAJ&printsec=frontcover&hl=ru#v=onepage&q&f=false> (дата обращения: 28.05.2026).
19. Zadeh L. A. Fuzzy sets // *Information and Control*. 1965. Vol. 8. Is. 3. P. 338–353. URL: <https://www.sciencedirect.com/science/article/pii/S00199586590241X?via%3Dihub>. DOI: 10.1016/S0019-9958(65)90241-X (дата обращения: 28.05.2026).
20. Похолков Ю. П. Развитие системы независимой профессионально-общественной аккредитации инженерных образовательных программ в России в период с 2000 по 2013 год // *Инженерное образование*. 2013. № 12. С. 50–57. URL: https://ojs.aeer.ru/public/archive/12/art_6.pdf. (дата обращения: 28.05.2026).

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest: The authors declare that there is no conflict of interest.

Финансирование: Исследование выполнено при финансовой поддержке Кубанского научного фонда в рамках научно-инновационного проекта «НИП-20.1.4».

Financing: The study was carried out with the financial support of the Kuban Science Foundation within the framework of the scientific and innovative project “NIP-20.1.4”.