

## **ПРИМЕНЕНИЕ СЕМАНТИЧЕСКОГО АНАЛИЗА ТЕРМИНОВ И ОПРЕДЕЛЕНИЙ ЦИФРОВОГО ТОЛКОВОГО СЛОВАРЯ ДЛЯ ПРЕОДОЛЕНИЯ ИХ НЕОДНОЗНАЧНОСТИ**

**Иванус А.И.**

*Федеральное государственное бюджетное образовательное учреждение высшего образования «Финансовый университет при Правительстве Российской Федерации» 125993, Москва, Ленинградский проспект, 49, e-mail: aiivanus@fa.ru*

---

**Решение задач управления все в большей и большей степени усложняется из-за наличия высокого уровня неопределённости. Одной из причин неопределённости является неоднозначность понимания одних и тех же терминов и определений разными людьми, формирующими управленческие решения. Для решения задачи преодоления этой неоднозначности предлагается концепция использования семантического анализа терминов, которая ориентирована на информационное обеспечение процессов решения широкого спектра задач управления социально-экономическими системами.**

---

Ключевые слова: термин, семантически анализ, истинность информации, новые знания, процессы мышления.

## **APPLICATION OF THE SEMANTIC ANALYSIS OF TERMS AND DEFINITIONS OF THE DIGITAL DICTIONARY DICTIONARY FOR THE OVERCOMING THEIR UNIFORMNESS**

**Ivanus A.I.**

*Federal State Educational Institution of Higher Education "Financial University under the Government of the Russian Federation», 125993., Russia, Moscow, the Leningrad prospectus e-mail: aiivanus@fa.ru*

---

**The task of management in the conditions of high level of uncertainty requires a more sophisticated mathematical methods. One reason for the uncertainty is the ambiguity of understanding of the same terms and definitions to different people, forming management decisions. To solve the problem of overcoming this ambiguity is the concept of using semantic analysis of terms, which is focused on information support of processes to resolve a wide range of tasks of socio-economic systems.**

---

Key words: term, semantic analysis, the validity of the information, new knowledge, processes of thinking.

### **Введение**

Известно, что в каждой сфере человеческой деятельности содержится огромное количество терминов и определений (далее – терминов), которые, несмотря на то, что они воспринимаются как вполне понятные и очевидные, тем не менее, допускают большой или малый разброс в их понимании. Это относится практически ко всем терминам, независимо от их частоты использования, степени общеупотребительности и уровня квалификации тех, кто их использует.

Само по себе это явление, как правило, на бытовом уровне не несёт каких-либо отрицательных последствий. Однако, есть отрасли, где степень ответственности использования терминов необычайно высока.

Так, это в полной мере относится к юриспруденции, где совершенно недопустимо разное понимание одних и тех же терминов такими участниками судебного процесса, например, как судья, обвинитель и адвокат. По отзывам юристов именно эта проблема является актуальной сейчас, тем более что количество законов и законодательных актов лавинообразно увеличивается. (Кстати, относительно термина законодательный акт в юридическом словаре сказано: «Законодательный акт – понятие, вызывающее в науке права различные подходы. По мнению одних авторов .....»).

Аналогичные проблемы характерны и для других видов научной и профессиональной деятельности.

Поэтому неспроста в медицине принята латынь как международный язык медиков. Видимо, медики одни из первых поняли опасность неоднозначного понимания терминов для лечения больных.

Образцом однозначности понимания терминов среди современных наук служит также и математика, где неоднозначность в толковании математических терминов доведена практически до нуля.

В связи с изложенным, задача создания системы семантического анализа терминов, представляется весьма актуальной. Это необходимо, прежде всего, потому, что проблема неоднозначности понимания терминов конвертируется в проблему неоднозначности управленческих решений, что часто приводит к большим рискам и потерям.

Современные методы математического моделирования социально-экономических систем достигли высокого уровня эффективности их использования. Но это относится, прежде всего, к методам, использующих численные количественные меры. Однако, есть огромное количество характеристик, которые численно не выражаются, а выражаются словами и утверждениями на вербальном уровне. То есть, их описание возможно на семантическом уровне. А вот семантические методы описания пока ещё недостаточно развиты по сравнению с классическими математическими методами.

Кроме того, есть ещё одно подтверждение в пользу необходимости создания толкового словаря нового, цифрового типа для более эффективного упорядочения информационного пространства [5].

Инструментальной основой для использования мягкой математики, и в частности, центрального её понятия — лингвистическая переменная на основе функции принадлежности — используются новые объекты — слова (или термины), которые приобретают новый смысл не только как средство общения, но и как новый предмет для математического исследования и анализа. Слова становятся частью математической модели, а при более целостном рассмотрении — самостоятельной математической моделью.

Данное обстоятельство представляется настолько важным, что из него вытекает принцип организации сложных систем, известный как принцип несовместимости Л. Заде: при росте сложности системы уменьшается возможность её точного описания вплоть до некоторого порога, за которым точность и релевантность (смысловая связанность) информации становятся несовместимыми, взаимно исключаящими характеристиками [2]. Поэтому значение системы семантического анализа терминов для практических приложений трудно переоценить.

### **Решение задачи унификации минимизации неоднозначности толкования терминов методами семантического анализа терминов**

Используем для решения задачи унификации минимизации неоднозначности толкования терминов подход, основанный на семантическом анализе терминов. Конструктивной основой такого подхода является семантика некоторого термина с максимальной истинностью. Уже сейчас можно предвидеть те научно-технические проблемы, которые со всей очевидностью возникнут уже в самом начале разработки – это проблемы организации целенаправленного поиска, агрегации, анализа потоков информации и используемых терминов на самом высоком семантическом уровне.

Для решения всего предполагаемого комплекса задач потребуются специальные технологии, но появятся и такие задачи, облик которых сейчас трудно предвидеть, и их придётся разрабатывать непосредственно в процессе создания базы терминов (БТ).

Из вышесказанного следует, что сложнейшая, по масштабному замыслу БТ, не может не содержать диктуемых внутренней логикой перманентного развития противоборствующих и противоречащих друг другу тенденций, как в алгоритмах своего построения, так и в реализуемых решениях. Поэтому в основу методов функционирования системы должен быть положен как понятный по сути, так и универсальный по применимости такой стержневой критерий, который позволял бы взаимно согласовать и увязывать эти жёсткие, но необходимые требования. Таким понятным и универсальным стержневым критерием построения БТ является критерий на соответствие пропорциям Фибоначчи [2].

Поэтому вполне можно утверждать, что в процессе использования термина необходимо знать динамику его эволюции, чтобы понимать, как его правильно использовать и что от него можно ожидать.

В соответствии с результатами [3] следует, что процесс формирования управленческих решений в условиях воздействия факторов неопределённости рассматривается с формальной точки зрения как эквивалентный ему процесс генерации новых знаний (НЗ).

Процесс генерации НЗ начинается с момента, когда формируется семантическое ядро истинности. Его наличие есть необходимое условие формирования НЗ.

В самом простом варианте семантическое ядро истинности можно наблюдать на примере терминов. Наличие семантического ядра истинности для терминов легко проверяется. Методом простого опроса группы людей был получен ряд распределений частот повторяемости ключевых слов, участвующих в процессе определения заданных терминов. Пример такого распределения для термина «Молодёжь» в масштабе log-log приведён на рис. 1. Из рисунка видно, что для распределения характерно наличие двух зон.

Первая зона образуется первыми двумя-тремя ключевыми словами, здесь наклон аппроксимирующей

прямой - небольшой. Эти большие значения частоты встречаемости ключевых слов можно объяснить наличием семантического ядра чувства истинности понимания термина.

А последующие точки уже образуют зону с большой величиной наклона, которая резко уходит вниз. И в этой зоне ключевые слова встречаются реже и как бы выполняют роль «обертонов», дополняющих основное ядро. Семантическое ядро истинности есть по своей сути результат деятельности правого и левого полушарий мозга человека [3]. Величина энтропии  $H$ , как мера неопределённости семантического ядра истинности, имеет максимальное значение 0,38 [1, 2].

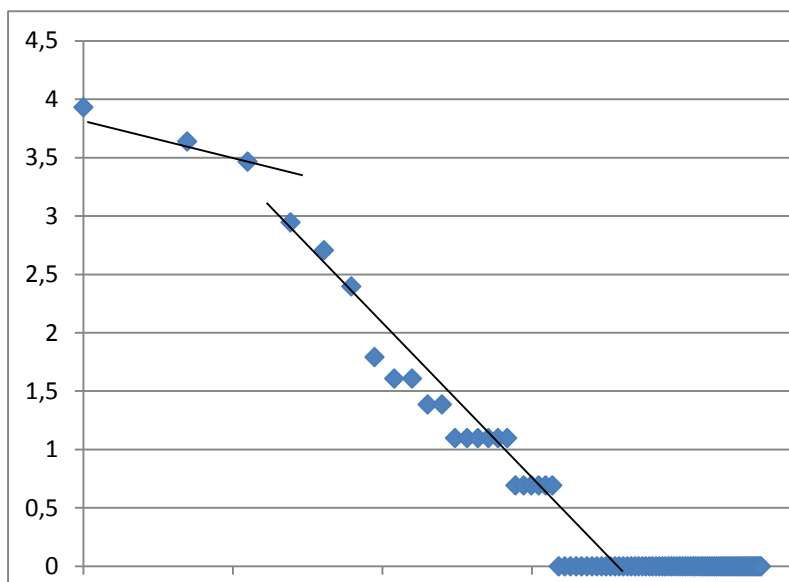


Рис. 1. Иллюстрация ядра истинности термина «Молодёжь», полученная на основе данных опроса группы студентов. По горизонтальной оси отложены ключевые слова, а по вертикальной – частоты их повторения

Всякое увеличение энтропии ядра истинности приведёт к его распаду, т.е. непониманию термина. Значение максимальной энтропии 0,38 может служить точкой отсчёта, уменьшение энтропии относительно которой можно измерять увеличение семантического понимания термина.

Семантическое ядро истинности можно представить, как ЖС, в виде двух составляющих в соответствии с формулой (4):

$$\text{Нсемант. ядро} = \text{Нсем} - \text{Ноберт},$$

где:

Нсемант. ядро – энтропия семантического ядра истинности термина, или, другими словами, энтропия семантического ресурса, Нсем – максимальное значение энтропии истинности термина, Н оберт – энтропия «обертонов».

Механизм понимания семантики терминов эквивалентен процессу понимания семантики НЗ, поскольку описание НЗ всегда включает в своём составе термины, требующие аргументации. Это означает, что нам всё равно, идёт ли речь о генерации НЗ или понимания термина, в основе данного механизма всегда лежит семантическое ядро чувства истинности.

А если эти рассуждения пролонгировать на управленческие решения, то можно распространить понятие семантического ядра истинности и на управленческие решения. Действительно, любое сгенерированное управленческое решение требует однозначного понимания, а без семантического ядра истинности этого понимания мозг человека данное решение не сможет воспринять.

В целом можно заключить, что всякое умственное действие, которое сопровождает термины, НЗ, управленческие решения, имеет в своей основе семантическое ядро чувства истинности, как живую систему, жизненный цикл которой описан в [4].

Вот поэтому в силу своего универсализма рассмотрим семантическое ядро истинности в качестве функции принадлежности при решении задач когнитивного моделирования экономических процессов.

Как показано в [4], функция принадлежности, заданная на семантическом ядре истинности, всегда имеет вид гауссового распределения.

Поэтому для данной функции принадлежности справедливы теоремы Гаусса-Маркова и, как следствие, в анализе и моделировании экономических систем применимы все эконометрические конструкции метода наименьших квадратов.

Такой подход соответствует результатам, которые получены рядом авторов в части моделирования чувственной стороны процессов познания.

Основная идея состоит в том, что знания – это есть не только информация сущность, но и неотъемлемое наличие чисто человеческого чувства осознания истинности этой информации.

Данный подход может оказаться весьма полезным при формировании алгоритмов одного из самых главных блоков современных интеллектуальных роботов – классифицирующего автомата. И здесь функция распределения чувства истинности на множестве аргументов, подтверждающих истинность, может служить функцией принадлежности.

Таким образом, ФП, заданная как распределение истинности на множестве аргументов, представляет собой отдельную задачу классификации знаний по признаку истинности.

## Выводы

1. Основная идея предлагаемого подхода состоит в доказательстве того, что для унификации терминов и определений необходимо использовать интеллектуальную систему семантического анализа энтропии распределения величины разброса параметров истинности. Эта система должна имитировать когнитивные процессы мозговой деятельности человека.

2. Если величина разброса параметров истинности мала, то это говорит о стабильности семантического ядра термина и здесь есть возможность минимизации этого разброса.

3. Если же величина разброса велика, то это может служить подсказкой в наличии процесса зарождения нового термина или нового направления в толковании термина. Устранение неоднозначности понимания термина в этом случае невозможно, но возможна минимизация этой неоднозначности. При этом можно искусственно ускорить этот процесс и таким образом сформировать новое толкование термина или даже на этой основе – сформировать предполагаемую область возможного существования нового знания.

4. Данный подход позволяет также ранжировать существующие инновационные направления с точки зрения их перспективности или неперспективности.

## Список литературы

---

1. Заличев Н.Н. Энтропия информации и сущность жизни. М.: Радиоэлектроника, 1995. – 192 с.
2. Иванус А.И. О ключевой роли истинности знаний в задаче принятия управленческих решений в экономике. Хроноэкономика, № 2 (4), 2017. Стр. 13-24.
3. Иванус А.И. Гармоничное управление инновационной экономикой в условиях неопределённости. -М., Книжный дом «Либроком», 2012. –208 с.
4. Иванус А.И. Об использовании принципов функционирования живых систем в моделировании экономических систем. М.: Хроноэкономика, № 5 (7), 2017. Стр. 14-22.
5. Трусов Н.А., Башелханов И.В., Иванус А.И., Заличев Н.Н., Костиков П.А., Каратеев А.Д., Мартынов С.В. О необходимости создания технологии семантического анализа терминов в интересах задач управления. Хроноэкономика, № 3 (5), 2017. Стр. 15-22.

## References

---

1. Salichev N. N. The entropy of information and the essence of life. M.: Radio Electronics, 1995. – 192 p.
2. Ivanus A. I. On the centrality of truth of knowledge in the problem management decisions in the economy. Chronosequence, No. 2 (4), 2017. p. 13-24.
3. Ivanus A. I. Harmonious management of the innovation economy in conditions of uncertainty. Moscow, Knizhnyy Dom "Librokom", 2012. -208 p.
4. Ivanus A. I. On the use of the principles of functioning of living systems in modeling economic systems. M.: Chronomania, No. 5 (7), 2017. p. 14-22.
5. Panties N.A. Bushlanov I. V., A. I. Ivanus, Salychev N. N., Kostikov, A. P., Karataev A. D., Martynov S. V. About necessity of creation of technology of semantic analysis of terms in the interests of the management tasks. Chronosequence, No. 3 (5), 2017. p. 15-22.