

К.т.н., доц. Бисикало О.В.

Винницкий государственный аграрный университет, Украина

Субъективная единица смысла образных конструкций

Актуальность и значимость тематики предлагаемого исследования можно выразить в виде далеко не риторического вопроса – есть ли смысл рассматривать формальную единицу смысла? Интеллектуальная деятельность людей всегда была неразрывно связана с пониманием смысла, однако для достижения эффективных результатов такой деятельности, как показывает многовековая практика, совершенно не обязательно измерять количество смысла. С другой стороны, существующие единицы измерения исторически определялись исходя из потребностей совершенствующихся в процессе эволюции человеческого общества производственных отношений. Ключевой на сегодня проблемой компьютерной лингвистики и многих других направлений искусственного интеллекта остается семантика и прагматика естественно-языковых конструкций [1]. Смысл речевого высказывания, как и другие феномены человеческого мышления, принадлежит к слабоформализуемым понятиям, но если единица смысла обеспечит прогресс в развитии некоторых приложений искусственного интеллекта, ее целесообразно вводить.

Наиболее явно понятие смысла проявилось в работах Московской семантической школы, стоявшей у истоков модели «Смысл – текст» [2]. Под смыслом здесь понималось последовательное наращивание лингвистических отношений для цепи морфология – синтаксис – семантика в условиях интегрального описания словарной и грамматической компонент языка [3]. Однако формальное определение смысла и по сей день значительно уступает естественному, например, в системах машинного перевода, что ярко иллюстрируется парадоксом понимания смысла речи необразованными людьми, не сведущими в лингвистике.

Исследуемая проблема также тесно связана с теорией информации и близкими к ней теориями вероятностей и нечетких множеств. Ценность

информации по Харкевичу пропорциональна увеличению вероятности достижения цели [4], однако данный показатель отражает прагматическую сторону сообщения в ущерб семантической и не учитывает механизм целеполагания системы. Функция принадлежности в нечеткой логике аккумулирует субъективное мнение эксперта относительно принадлежности элемента к некоему классу элементов [5], но базируется на описательном принципе и не содержит конструктивного алгоритма своего определения.

Формальным языком описания смысла предложения на естественном языке может служить нейронная сеть, при этом нейрон выступает как математическая основа масштабируемой единицы смысла для символа, части слова, слова, словосочетания, предложения, абзаца, всего текста [6]. Данный подход подразумевает под смыслом текста или другого языкового понятия предикат, принимающий значения «истина» или «ложь», однако открытым остается вопрос генерации адекватного отклика системы на внешнее воздействие.

Понятие смысла тесно связано не только с семантическими и прагматическими аспектами передачи информации, но и с когнитивными способностями человека, изучаемыми психологией, физиологией, семиотикой, лингвистикой и другими науками. Научной проблемой можно считать отсутствие обобщающего подхода для уже известных аспектов смысла – вероятностного, нечеткого, прагматического и нейронного. Задача работы состоит в обосновании возможности построения единицы смысла исходя из концепции моделирования образного мышления [7], позволяющей на формальном уровне дать трактовку парадокса необразованных людей.

Рассмотрим инфологическую систему, способную воспринимать n образов из множества $I = \{i_1, i_2, \dots, i_n\}$ в виде дискретного потока внешних событий E и обеспечивать свои внутренние потребности с помощью периодической генерации выходных образных потоков O . Без ограничения общности можно считать, что образы входных и выходных потоков в дискретные моменты времени попарно связаны между собой, т.е. справедливы определения множеств $\{e \in E \mid e \subseteq I^2\}$ и $\{o \in O \mid o \subseteq I^2\}$. При этом все e и o являются

ориентированными деревьями, поскольку в случае языковых образов допускают описание с помощью синтагм – простых повествовательных предложений. В силу своего определения [7] инфологическая система фиксирует количество k_{ij} повторений связи между i -м и j -м образами в результате построения базы знаний $B = \bigcup_{l=1}^L e_l$, где L – общее количество воспринятых системой внешних событий.

Предположим, что относительный вес v_{ij} ассоциативной по своей сущности связи (i, j) в базе знаний B характеризует элементарный смысл использовать для реакции системы именно эту упорядоченную пару образов. Тогда можно потребовать, чтобы функция смысла $Se_{ij} = v_{ij} \in [0,1]$ аналогично функциям вероятности и принадлежности. При этом $v_{ij} = 0$, если связь (i, j) ни разу зафиксирована не была и $v_{ij} = 1$ в случае $v_{ij} \gg \lambda$, где λ – среднее количество повторений одной связи в базе знаний. Определив количество n' ненулевых связей из матрицы B и общее количество повторений всех связей $m = \sum_{i=1}^n \sum_{j=1}^n k_{ij}$, нетрудно найти значение $\lambda = m/n'$.

Потребуем, чтобы функция смысла Se учитывала кроме физиологической статистики в виде k_{ij} еще два фиксируемых системой параметра:

- a) Степень успеха/неуспеха при использовании данной связи для реакции на аналогичные воздействия в предыдущем опыте.
- b) Уровень эмоционального состояния позитив/нейтральное/негатив как показатель близости достижения актуальной потребности системы.

Для определения функции смысла ассоциативной связи (i, j) возьмем за основу сигмоидальную функцию, которую запишем в виде:

$$Se_{ij} = 1 / (1 + e^{-\frac{k_{ij} - \lambda_{ij}}{\mu}}),$$

где μ – показатель эмоционального состояния (b). При этом параметр a) можно регулировать с помощью алгоритма отклонения λ_{ij} от среднего значения λ .

Принимая во внимание природу рассмотренного формального определения функции Se_{ij} назовем единицу измеряемого количества смысла как *Saw* (*Синтагматической ассоциации вес*) – данная аббревиатура имеет схожее звучание на украинском (*Синтагматичної асоціації вага*) и английском (*Syntagmatic Association Weight*) языках. Кроме того, английское слово *Saw* (вторая форма неправильного глагола *to See*) переводится как «посмотренное», «увиденное» и, тем самым, указывает на образный путь появления смысла в инфологической системе.

Предложенный подход к определению единицы смысла *Saw* позволил формально связать вероятностную природу восприятия образов с физиологией условного рефлекса и психологическими механизмами потребностей–эмоций. Перспективными задачами можно считать введение метрики пространства ассоциативных пар на основе функции смысла *Se* и построение алгоритма функционирования инфологической системы.

Литература:

1. Попов, Э.В. Общение с ЭВМ на естественном языке [Текст] / Э.В. Попов. – М.: Наука, 1982. – 360 с.
2. Мельчук И.А. Русский язык в модели "Смысл - Текст" [Текст] / И.А. Мельчук. – М.: Языки русской культуры, 1995. – 682 с.
3. Апресян, Ю.Д. Лексическая семантика (синонимические средства языка) [Текст] / Ю.Д. Апресян. – М.: Наука, 1974. – 324 с.
4. Харкевич, А.А. О ценности информации [Текст] / А.А. Харкевич // Проблемы кибернетики. – М.: Физматгиз, 1960. – Вып. 4. – С.53–57.
5. Заде, Л.А. Размытые множества и их применение в распознавании образов и кластер-анализе [Текст] / Л.А. Заде // В кн.: Вопросы анализа и процедуры принятия решений. – М.: Мир, 1976. – С. 172–215.
6. Дударь, З.В. Семантическая нейронная сеть, как формальный язык описания и обработки смысла текстов на естественном языке [Текст] / З.В. Дударь, Д.Е. Шуклин // Радиоэлектроника и информатика. Х.: Изд-во ХТУРЭ, 2000. – № 3. – С. 72–76.
7. Бісікало, О.В. Концептуальні основи моделювання образного мислення людини [Текст] : моногр. / О.В. Бісікало. – Вінниця: ПП Балюк І.Б., ВДАУ, 2009. – 163 с.