

Вероятностная модель языка

По трудам В.В. Налимова

Проблемы семантики

- Как знак связан с тем, что он обозначает?
- Как знаковые системы используются в интеллектуальной деятельности?
- Чем различаются в семантическом отношении знаковые системы, созданные в разные культурные эпохи?
- Чем они отличаются от единственной не социальной созданной знаковой системы биологического кода?

Характеристики знаковых языковых систем

Функциональные:

- передача информации;
- свертка, хранение и воспроизведение.

Структурные:

- алфавит,
- грамматика.

Знак и значение

«Знак и значение не покрывают друг друга полностью. Один и тот же знак имеет несколько функций, одно и то же значение может быть выражено несколькими знаками. Всякий знак является потенциально омонимом и синонимом одновременно, т.е. он образован скрещением этих двух рядов мыслительных явлений.»

«Если бы знаки были неподвижный и каждый из них выполнял только одну функцию, язык стал бы простым собранием этикеток. Но также невозможно представить себе язык, знаки которого были бы подвижны до такой степени, что они ничего бы не значили за пределами конкретных ситуаций.»

С. О. Карцевский

Размерность языка

Размерность языка – количество знаков, участвующих в формировании значения одного знака.

Размерность языка может не совпадать с размерностью мира, который он описывает.

Проблема перевода одного языка в другой усложняется если языки имеют различные размерности.

Обыденный язык семантически и семиотически нелинеен.

Иерархия языков и знаковых систем

- Один язык может быть представлен в различных знаковых системах, образующих иерархию.
- Язык может быть объектом другого, иерархически выше стоящего языка – метаязыка.

Вероятностная теория языка

Задача: построить модель, отражающую как логическую структуру языка, так и существование неоднозначной связи между знаком и значением.

Проблема (вывод из теоремы Гёделя): Если теория первого порядка непротивоерчива – то она неполна.

Байесовская модель языка

Теорема Байеса:

$$P(a | b) = k P(b | a) P(a)$$

$P(a)$ – априорная функция распределения
смыслового значения (знание о слове),

$P(b | a)$ – функция изменения значения процессом
чтения (употребление слова),

$P(a | b)$ – апостериорная функция распределения
смыслового значения,

k – коэффициент нормирования.

Байесовская модель языка

$$P(a|b) = k P(b|a) P(a)$$

Модель задает обучающий алгоритм.

В модели заложены понятия:

- точность термина,
- синонимия,
- ассоциация,
- абстракция.

Язык абстрактной живописи

Первичный алфавит (с примерами операций):

- линия (прерывание, изгиб, излом, пересечение),
- полоса (прерывание, изгиб, излом, пересечение),
- фигура (дробление, деформация, наложение),
- фон (дробление, плавное изменение),
- цвет (чистый цвет, дробление, перетекание),
- диффузные облака (изменение тона, наложение).

Алфавит и грамматика в семантически пусты.

Язык биологического кода

Алфавит (РНК): 4 рибонуклеиновых кислоты

Грамматика: триплеты РНК фиксированным образом кодируют 20 аминокислот.

Язык биологического кода избыточен, есть синонимы, есть бессмысленные коды.

В. В. Налимов

Вероятностная структура языка

О соотношении естественных и
искусственных языков

Изд. Наука, 1979г.