

ББК Ч 480.28

Л. В. Макарова, С. А. Маркелов

К ВОПРОСУ ОЦЕНКИ СЕМАНТИЧЕСКОЙ ИНФОРМАЦИИ В
ОБРАЗОВАТЕЛЬНЫХ СИСТЕМАХ

(Самарский государственный аэрокосмический университет)

На современном этапе развития образовательных систем резко возросло количество данных, фактов, правил и других лексических единиц, которыми перегружены тексты лекций, учебных программ и т. п. Такая перегрузка ведет к ухудшению качества образования, которое часто пытаются "улучшить", вводя дополнительные данные, факты, правила, якобы оптимизирующие качество образования. Выходом из такого порочного круга могут стать образовательные технологии, использующие для оптимизации образовательных процессов процедуры определения вклада в образовательный процесс лексических единиц, объективно необходимых в контексте данной предметной области, то есть процедуры измерения семантической информации.

В процессе общения переданная информация может быть оценена двояким образом - по количеству переданных терминов (так называемая "лексическая информация") и по его смысловой значимости. Однако если количество переданных терминов абсолютно, смысловая значимость текста зависит от контекстуальных особенностей передаваемого сообщения, то есть наличия структурированных знаковых последовательностей, сопровождающих текст и определяющих смысловую нагруженность терминов. Если подходить к сообщению, как к последовательности знаков (подход Шеннона), то количество лексической информации, связанное с сообщением, можно определить по формуле

$$H = N \star N_0,$$

где N_0 - мера неопределенности, связанная с одним знаком в последовательности и определяемая для алфавита в M знаков соотношением

$$N_0 = \text{LOG}(M);$$

N - длина последовательности (в знаках - включая пробелы).

С другой стороны, для определения объема семантической информации в сообщении оказывается важным фиксировать контекст сообщения и установить зависимость смысла от этого контекста.

Авторами разработан элемент представления ПО, называемый

далее элементарной ячейкой логико-категориального тезауруса (ЛКТ), и пригодный для решения указанной задачи.

Авторская концепция представления знаний на основе ЛКТ, позволяет: представлять знания и умения на высших уровнях - системном и концептуальном, допускающих стратегическое планирование образовательной деятельности, а также формирование новых понятий на основе старых; оперативно осуществлять переход от декларативной компоненты знаний к операциональной компоненте; определять шаги по оперативной коррекции содержания обучения при изменении условий обучения (времени обучения, стоимости обучения, интеллектуального потенциала обучаемого и т. п.); измерять объем семантической информации и осуществлять расчет количества часов, отводимых на изучение дисциплины; осуществлять квалиметрию динамических образовательных систем; рассчитывать системные параметры образовательных программ.

В выбранном способе представления знаний фиксация контекста возлагается на так называемый коммутационный блок элементарной ячейки ЛКТ, обеспечивающий навигацию в базе данных, представляющей систему ограничений, условий и постусловий реализации отношений между терминами предметной области.

Семантическое (смысловое) содержание сообщения может быть представлено совокупностью возможных переходов от начального состояния уровня знаний (в рамках контекста) о тематике сообщения к его конечному состоянию. Если N - число всех возможных путей переходов (образующих семантическую сеть), N_0 - число априорно значимых путей (например, предустановленных), то объем семантической информации может быть найден в рамках концепции кросс-энтропии из соотношения Кульбака-Лейблера:

$$H = \text{LOG} (N/N_0) .$$

Наличие декларативной семантики однозначно определяет структуру ЛКТ (через внутритезауральные ссылки, фиксирующие совокупность заранее сформулированных правил вывода и отобранных фактов). Это, в свою очередь, обеспечивает однозначность и осуществимость определения числа N_0 и, следовательно, однозначность расчета семантической информации в рамках контекстов, задаваемых ЛКТ.