ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 681.3.07

В.В. МАРАСАНОВ, А.О. ДЫМОВА, Р.Ю. НЕГРУЦА, В.С.ДЫМОВ
Херсонский национальный технический университет

ОПРЕДЕЛЕНИЕ СМЫСЛА ПРЕДЛОЖЕНИЙ ТЕХНИЧЕСКОГО ТЕКСТА ПО ПРОЕКЦИЯМ ОБЛАСТЕЙ ОПРЕДЕЛЕНИЯ ПРЕДИКАТОВ ПЕРВОГО И ВТОРОГО УРОВНЯ

Решение задачи определения смысла предложений технического текста сведена к представлению предложений в виде логических формул – предикатов всеобщности первого и второго (предикатов от предикатов) уровней и на основе принципа двойственности нахождения для них соответствующих предикатов существования. Теоретико-множественный смысл областей определения предикатов существования для соответствующих предикатов совпадает с их проекциями на одномерные оси для предикатов первого уровня (плоскости – для предикатов от предикатов) и определяет смысл предложений текста.

Ключевые слова: предикаты всеобщности, теоретико-множественный закон, кванторы, принцип двойственности, предикат существования, метод проекций, ортогональные проекции, область определения.

В.В. МАРАСАНОВ, Г.О. ДИМОВА, Р.Ю. НЕГРУЦА, В.С. ДИМОВ
Херсонский национальный технический университет

ВИЗНАЧЕННЯ СЕНСУ РЕЧЕНЬ ТЕХНІЧНОГО ТЕКСТУ ЗА ПРОЕКЦІЯМИ ОБЛАСТЕЙ ВИЗНАЧЕННЯ ПРЕДИКАТІВ ПЕРШОГО І ДРУГОГО РІВНЯ

Рішення задачі визначення сенсу речень технічного тексту зведена до подання речень у вигляді логічних формул – предикатів загальності першого і другого (предикатів від предикатів) рівнів та на основі принципу подвійності знаходження для них відповідних предикатів існування. Теоретико-множинний зміст областей визначення предикатів існування для відповідних предикатів збігається з їх проекціями на одномірні осі для предикатів першого рівня (площини – для предикатів від предикатів) і визначає сенс речень тексту.

Ключові слова: предикати загальності, теоретико-множинний закон, квантори, принцип подвійності, предикат існування, метод проекцій, ортогональні проекції, область визначення.

V.V. MARASANOV, A.O. DYMOVA, R.YU. NEGRUTSA, V.S. DYMOV
Kherson National Technical University

A CERTAIN SENSE, THE PROPOSED TECHNICAL TEXT ON THE PROJECTION OF THE DOMAIN OF DEFINITION OF PREDICATES OF THE FIRST AND SECOND LEVELS

Solution of the problem of determining the meaning of the text is reduced technical proposals to submit proposals in the form of logical formulas – universal predicates first and second (predicates on predicates) levels and on the basis of the duality principle of finding them relevant predicate existence. Set-theoretic sense domains of predicates of existence for the respective predicates coincides with their projections on one-dimensional axis to predicate the first level (the planes – for the predicates of predicates) and determines the meaning of the proposed text.

Keywords: universal predicates, set-theoretic law quantifiers, the principle of duality, the predicate of existence, the method of projection, orthogonal projection domain.

Постановка проблемы

В процессе анализа, распознавания и перевода технического текста при помощи средств вычислительной техники возникает задача определения смысла предложений для повышения качества распознавания или перевода. Основная проблема при этом связана с наличием специфических для каждой предметной технической области терминов и определений, а также с необходимостью построения связных предложений.

Задача сводится к представлению предложений в виде логических формул – предикатов всеобщности первого и второго (предикатов от предикатов) уровней и на основе принципа двойственности нахождения для них соответствующих предикатов существования. Предлагается использовать теоретико-множественный смысл областей определения предикатов существования для соответствующих предикатов для анализа их совпадения с проекциями на одномерные оси для предикатов первого уровня (плоскости – для предикатов от предикатов), что позволит более точно определить смысл предложений текста.

Анализ последних исследований и публикаций

Исходя из структуры предложений, рассмотрим множества, соответствующие их логическим выражениям. Пусть W – некоторое множество (область), на котором определены предикаты. Каждому предикату одной переменной $F(x)$ можно поставить в соответствие множество тех элементов a из области W, для которых $F(a)$ истинно. Обозначим это множество E_F. Обратно, каждому множеству $E\subset W$ можно поставить в соответствие предикат $P(x)$, представляющий собой высказывание, истинное, если $x\in E$. Предикат $P(x)$ принимает значение И (истина) на E и Л (ложь) на CE (дополнение множества E). Следовательно E есть E_P. [1, 2].

Пусть:

$$P(x) \equiv P_1(x) \vee P_2(x).$$

Тогда:

$$E_P = E_{P_1} \cup E_{P_2}$$

и если $x\in E_P$, то $P(x)$ истинно.

Значит $P_1(x)$ или $P_2(x)$ истинны, следовательно:

$$x \in E_{P_1} \cup E_{P_2} .$$

Аналогичным образом можно показать, что если:

$$P(x) \equiv P_1(x) \wedge P_2(x),$$

то:

$$E_P = E_{P_1} \cap E_{P_2} .$$

Множество, отвечающее предикату $\overline{P}(x)$, является дополнением к множеству, отвечающему предикату $P(x)$, т.е. $E_{\overline{P}} = CE_P$ и $W = E_P \cup CE_P$.

В силу соответствия между логическими функциями и множествами законам логики предикатов соответствуют известные законы для теоретико-множественных операций (изоморфизм логики Буля и логики Кантора).[1-3, 5, 6]. Например, первому закону дистрибутивности:

$$F(x)[G(x) \vee H(x)] \equiv F(x)G(x) \vee F(x)H(x)$$

соответствует теоретико-множественный закон:

$$P \cap (Q \cup S) = P \cap Q \cup P \cap S ,$$

где P, Q, S – произвольные множества.

Пусть теперь W^2 – множество всех пар (x, y) множества W. При этом пары различаются не только составом элементов, но и порядком.

Предикату $P(x, y)$ поставим в соответствие множество тех пар $(x, y)\in W^2$, для которых $P(x, y)$ истинно. Обозначим это множество по-прежнему E_P.

Для выявления смысла введем кванторы и определим их теоретико-множественный смысл.

Формулировка цели исследования

Обосновывается возможность определения смысла предложений технических текстов на основе ортогональных проекций областей определения предикатов первого и второго (предикатов от предикатов) уровня.

Изложение основного материала исследований

Пусть:

$$F(x) = \exists y\, P(x, y),$$

где $\exists y$ - квантор существования.

Множество E_P, соответствующее предикату F, состоит только из тех элементов области W, для которых истинно $F(x)$, т.е. $\exists y\, P(x, y)$, но $F(x) = \exists y\, P(x, y)$ истинно для конкретного x_0, если существует такое y, что $P(x_0, y)$ истинно. [1, 2].

Функции $P(x, y)$ отвечает часть E_P множества W^2.

Итак, E_P состоит из всех тех элементов x области W, для каждой из которых найдется пары (x, y), принадлежащая E_P. Назовем x_0 проекцией любой пары (x_0, y), а проекцией множества – совокупность проекций принадлежащих ему пар. Отсюда следует, что E_F есть проекция E_P. [2]

Дадим геометрическое определение смысла наших рассуждений (интерпретацию).

Пусть W – множество действительных чисел. Будем рассматривать W^2 как плоскость, точки которой имеют координаты x, y, а W – как ось OX этой плоскости. В этом смысле точка x является проекцией точки (x, y) в правильном геометрическом смысле. Так мы подошли к смыслу $\exists y\, P(x, y)$, согласно которого множество E_F, отвечающее области определения логической функции $F(x) = \exists y\, P(x, y)$, совпадает с обычной ортогональной проекцией множества E_P на ось OX. [1]

Обозначим проекцию произвольного множества H на W в виде $пр_x\, H$.

Тогда $E_F = пр_x\, E_P$.

Для окончательного определения теоретико-множественного смысла проектирования в соответствии с [1, 2] определим смысл квантора всеобщности. Для этого применим закон действия отрицания на квантор в соответствии с принципом двойственности.

Пусть:

$$F(x) = \forall y\, P(x, y),$$

тогда по принципу двойственности:

$$\forall y\, P(x, y) \equiv \overline{\exists y\, \overline{P}(x, y)}.$$

Операция отрицания соответствует теоретико-множественной операции дополнения:

$$E_F = C\, пр_x\, CE_P,$$

т.е. множество, отвечающее функции $\forall y\, P(x, y)$, есть дополнение к проекции на W дополнения к E_P.

При этом справедливо и обратное: всякое множество R, являющееся проекцией на W множества D, содержащегося в W^2:

$$R = пр_x\, D$$

и может быть представлено как E_F, где $F(x)$ есть $\exists y\, P(x, y)$, причем $D = E_P$. Действительно, множеству R отвечает предикат $F(x)$, определенный на W, а множеству D – предикат $P(x, y)$, определенный на W^2, и, очевидно $F(x) = \exists y\, P(x, y)$. [1, 2]

Если же R' – дополнение к проекции D, то R' соответствует предикату $\forall y\, \overline{P}(x, y)$.

В самом деле, предикату $\forall y\, \overline{P}(x, y)$ соответствует множество $C\, пр_x\, CE_{\overline{P}}$, но $CE_{\overline{P}} = E_P = D$ и следовательно $C\, пр_x\, CE_{\overline{P}} = C\, пр_x\, D$.

Итак, кванторы, связаны с геометрической операцией проектирования и, обратно, проектирование имеет указанный логический смысл. [1]

При лингвистическом анализе предложений технического текста математического аппарата исчисления предикатов первого уровня для определения смысла сложноподчиненного предложения оказывается недостаточно. Уточним, что будем понимать под формулой исчисления предикатов, а именно под формулой мы понимаем выражение, построенное осмысленным образом из знаков для переменных с помощью знаков $\wedge$, $\vee$, $\overline{}$, $\rightarrow$, $\sim$, (с учетом старшинства), связывающих высказывания, и знаков общности и существования (кванторов). Но соблюдение аксиоматической точки зрения, при которой доказательства приводятся по чисто формальным правилам, не прибегая к истолкованию логических знаков, делает необходимым характеризовать выражения называемыми формулами, только путем описания их формального построения и избегать через неуточненные понятия вроде "осмысленный".

С учетом этого можно считать, что для нахождения смысла предложения технического текста исчисление предикатов первого уровня явно недостаточно и требует соответствующего расширения понятия формулы. Для содержательной установки основы логического содержания исчисления предикатов было существенным и строго отделялись высказывания и предикаты от предметов, рассматриваемых как значения аргументов для предикатов. Однако для расширения исчисления ничто не препятствует рассматривать сами предикаты и высказывания как предметы, которые служат аргументами предикатов [2]. Например, логическое выражение $(x)(A \rightarrow F(x))$ можно понимать как предикат второго уровня $P(A, F)$, где первое пустое место занято высказыванием A, а второе пустое место – одномерным предикатом F.

Ложное высказывание A находится к каждому F в отношении $P(A, F)$; истинное высказывание только к таким F, для которых $(x)F(x)$ имеет место. Другими примерами служат свойства рефлексивности, симметричности и транзитивности двуместных предикатов. [1, 2]. Этим свойствам соответствуют три предиката: $Ref(R)$, $Sym(R)$ и $Tr(R)$, аргументом R которых является предикат с двумя пустыми местами. Символически эти три свойства выражаются следующим образом:

$$Ref(R): \quad (x)R(x,x),$$

$$Sym(R): \quad (x)(y)(R(x,y) \rightarrow R(y,x)),$$

$$Tr(R): \quad (x)(y)(z)(R(x,y) \wedge R(y,z) \rightarrow R(x,z)).$$

Предикат $\equiv(x, y)$ (x тождественно с y) обладает всеми ранее приведенными свойствами; предикат же $<(x, y)$ только свойством транзитивности. Таким образом формулы $Ref(\equiv)$, $Sym(\equiv)$, $Tr(\equiv)$, $Tr(<)$ представляют истинные высказывания, а формулы $Ref(<)$ и $Sym(<)$ – ложные высказывания.

Двуместным предикатом предикатов является эквивалентность $Aeg\ (F, G)$, которая определяется выражением $(x)(F(x) \sim G(x))$; она состоит в том, что предикаты F и G истинны (соответственно ложны) для одних и тех же значений аргументов. Аналогично двуместными предикатами являются: несовместимость $Unv\ (F, G)$ и импликация $Imp\ (F, G)$, которые символически определяются через:

$$(x)(F(x) \vee \overline{G}(x)),$$

$$(x)(F(x) \rightarrow G(x)),$$

При этом пока еще не получается расширения символики, так как приведенные предикаты предикатов могут выражены средствами исчисления предикатов первого уровня (узкого исчисления предикатов) и такие формулы как $Ref(R)$, $Sym(R)$ и т.д. надо понимать как сокращения [2, 7]. Расширение наступает лишь тогда, когда вводятся переменные для предикатов от предикатов, частными значениями которых являются приведенные индивидуальные предикаты. Первое их применение получено при определении смысла количества. Согласно этому числа выступают в качестве свойств предикатов. Пример этого является определение булеана β (множества всех подмножеств некоторого исходного множества). Например, $\beta(1, 2, 3) = \{(\varnothing), (1), (2), (3), (1,2), (1,3), (2,3), (1,2,3)\}$. Согласно этому, числа выступают в качестве свойств предикатов и для исчисления предикатов от предикатов; определенное число представляет собой индивидуальный предикат от предикатов. Этот способ выражения чисел основывается на том, что предикат от предикатов, которые образуют числа, могут быть полностью выражены при помощи логической символики. Для чисел 0, 1, 2, … и т.д., т.е. для предикатов от предикатов $0(F)$, $1(F)$, $2(F)$ их выражения имеют вид:

$$0(F): (\overline{E\,x})\ F(x),$$

что означает "не существует x, для которого бы выполнялось $F(x)$",

$$1(F): (Ex)\,[F(x) \wedge (y)(F(y) \rightarrow\ \sim(x,y)],$$

означающее – “существует x, для которого выполняется $F(x)$ и каждое y, которое удовлетворяет $F(y)$, тождественно с этим x”.

Сложение чисел можно свести к дизъюнкции предикатов.

Если F и G несовместные предикаты и предикату F принадлежит число m, а предикату G – число n, то предикату $F \vee G$ соответствует число $m + n$.

Например, равенство 1 + 1 = 2 выражается чисто логической формулой:

$$(F)(G)([Unv\,(F,G) \wedge 1(F) \wedge 1(G)] \rightarrow 2(F \vee G)) \ .$$

Что очевидно, если вместо предиката от предикатов Unv и вместо предикатов от предикатов 1, 2 подставить выражающие их соотношения.

Декартовые произведения множеств и их подмножества, выражающие упорядочные отношения в грамматических предложениях естественного языка позволят выразить смысл, содержащийся в техническом тексте [3–7].

Предикаты формальной арифметики при использовании аксиомы бесконечности теоретико-числовые аксиомы превращаются в логические доказуемые предложения и с использованием грамматических правил языка PROLOG позволяют производить анализ смысла технических текстов. [1-3, 7]

На основе принципа двойственности для перечисленных двуместных предикатов всеобщности можно перейти к предикатам существования $\exists$ и, используя метод проекций, определить их смысл.

Выводы

Задача нахождения смысла предложений технического текста формулируется в виде предикатов всеобщности и на основе принципа двойственности сводится к нахождению соответствующих предикатов существования $\exists$ и нахождения смысла в виде ортогональных проекций их областей определения.

Список использованной литературы

1. Новиков П.С. Элементы математической логики / П.С.Новиков – М.: Наука, 1973.
2. Гильберт Д. Основы теоретической логики / Д. Гильберт, В. Аккерман – М.: И.Л., 1947.
3. Лорьер Ж.-Л. Системы искусственного интеллекта / Ж.-Л. Лорьер – М.: Мир, 1991.
4. Люгер Д.Ф. Искусственный интеллект / Д.Ф. Люгер – М.: Вильямс, 2003.
5. Осуга С. Обработка знаний / С. Осуга – М.: Мир, 1989.
6. Уэно Х. Представление и использование знаний. / Под ред. Х. Уэно, М. Исидзука. – М.: Мир, 1989.
7. Братко И. Алгоритмы искусственного интеллекта на языке PROLOG / И. Братко – М.: Вильямс, 2004.