

СЕМАНТИЧЕСКАЯ ФОРМУЛИРОВКА ТЕОРЕМЫ ГЕДЕЛЯ О НЕПОЛНОТЕ И ЕЕ ЗНАЧЕНИЕ ДЛЯ ПОСТРОЕНИЯ СИСТЕМ РАСПОЗНАВА- НИЯ

к.т.н. Ю.В. Паржин, к.т.н. В.И. Новиков, К.В. Кувшинов
(представил д.ф.-м.н., проф. С.В. Смеляков)

Рассматриваются причины возникновения сформулированного Геделем понятия неполноты для формализованных языков и естественного языка и обсуждаются некоторые пути решения проблемы выводимости истинных утверждений в выразительных формализмах. Определяется значение данной проблемы для построения систем распознавания образов.

Постановка проблемы. Разнообразие существующих методов и систем распознавания в первую очередь связано с отсутствием единого теоретического подхода к построению подобных систем. Этим обусловлена довольно «жесткая» специализация применяемых систем распознавания, их конкретная направленность на распознавание объектов и процессов определенного вида при заданном множестве ограничений. Рассмотрение теоретических основ построения систем распознавания в большей мере связано с пониманием причин, определяющих невозможность построения в определенной степени «универсальных» систем распознавания на основе использования существующих формализмов. Анализ данных причин, их теоретическое обоснование позволили бы сформулировать подходы к развитию основ единой теории распознавания образов.

Анализ литературы. Проблемы распознавания образов, как и проблемы принятия решений, неразрывно связаны с проблемами представления информации, что в совокупности определяет широкую проблемную область, известную под названием «искусственный интеллект» [1]. Данная проблемная область основывается на глубоких философских законах, нашедших свое отображение в известных математических формализациях. К их числу в полной мере относится теорема Геделя о неполноте. При рассмотрении данного вопроса нас, в первую очередь, будет интересовать семантическая интерпретация этой теоремы в том виде, в каком она была приведена в работе В.А. Успенского [2].

Цель статьи. Исходя из семантической формулировки теоремы Ге-

деля, определить причины возникновения неполноты формализованных языков (ФЯ) и естественного языка (ЕЯ), рассмотреть пути решения данной проблемы, сформулировать выводы относительно подходов к построению систем распознавания, обладающих свойством непротиворечивости и способностью к расширению множества распознаваемых образов.

Результаты исследований. Семантическая формулировка теоремы Геделя о неполноте имеет следующий вид:

при определенных условиях в языке существует недоказуемое истинное утверждение.

Уточним приведенные в данной формулировке понятия.

1. Язык (в том числе ЕЯ) – формализованное отображение реального мира, его объектов, процессов, их качественных и количественных характеристик. Элементы языка, как знаковой системы, являются абстрагированными семантическими эквивалентами (АСЭ) объектов и процессов реального мира, данных нам в ощущениях (за исключением производных языков, о которых будет сказано ниже). Любой язык состоит из алфавита (множества элементарных знаков), слов – множества АСЭ или семантических определителей (СО) и правил их образования и взаимосвязи. Очевидно, что СО, как объекты языка, не могут находиться в причинно-следственной связи с объектами реального мира, так как любые СО не получены в результате какой-либо обработки информации о данных объектах, а просто поставлены им в соответствие. Таким образом, СО не несет полноту информации об объекте, которому он ставится в соответствие, а лишь указывает на возможное наличие этой информации в системе.

2. Исследованию истинности утверждений языка посвящено огромное количество работ, как в области математики, так и в области философии [3].

Для простоты понимания определим, что истинными утверждениями в языке являются правильно построенные выражения языка, состоящие из СО, отражающие факты реального мира.

3. «Недоказуемое» – не являющееся доказуемым. Термин «доказательство» – едва ли не самый главный в математике. В «Началах математики» Н. Бурбаки указывается, что «говорить «математика» – значит говорить «доказательство»». Однако определим, что доказуемое утверждение – утверждение, полученное в результате содержательного рассуждения на основе использования формальных правил, образующих причинно-следственные связи между этапами рассуждения, которые применяются к исходным посылкам (исходным истинным утверждениям –

аксиомам, гипотезам).

Определим, что на множестве всех слов B^∞ в алфавите B , задано подмножество T , называемое множеством «истинных утверждений». Пара $\langle B, T \rangle$ называется фундаментальной парой. Задано множество доказательств D^∞ в алфавите D , в котором выделено подмножество D , элементы которого называются доказательствами. Предполагается наличие алгоритма, позволяющего по произвольному слову в алфавите D узнать, принадлежит оно D или нет. Имеется функция δ (функция выделения доказанного), у которой область определения Δ удовлетворяет соотношению $D \subseteq \Delta \subseteq D^\infty$ и которая принимает свои значения в B^∞ ; предполагается наличие алгоритма, вычисляющего эту функцию; доказательство d из D называется доказательством слова $\delta(d)$. Тройка $\langle D, D, \delta \rangle$ называется дедуктикой над алфавитом B [2].

Уточненная семантическая формулировка теоремы Геделя о неполноте будет иметь вид:

при определенных условиях, налагаемых на функциональную пару $\langle B, T \rangle$, не существует дедуктики над B , полной и непротиворечивой относительно $\langle B, T \rangle$.

При этом дедуктика $\langle D, D, \delta \rangle$ непротиворечива относительно $\langle B, T \rangle$, если $\delta(D) \subseteq T$. Дедуктика полна относительно $\langle B, T \rangle$, если $\delta(D) \supseteq T$.

Доказательство данной теоремы общеизвестно. Уточним только, что необходимым и достаточным условием выполнения теоремы Геделя является неперечислимость множества T .

Интуитивно понятно, что в ЕЯ, как и в любом другом «богатом», «выразительном» языке множества всех истин настолько сложны, что неперечислимы, и потому для этих языков невозможно построить полные непротиворечивые дедуктики.

Это положение нашло свое подтверждение и в интуиционистских исследованиях по данному вопросу. Так, известная формулировка Тарского гласит, что:

истинное высказывание языка L может быть дано в некотором языке ML , причем ML более логически богат, чем L , а именно, $ML \supset L$ в каждой своей части. Необходимо, чтобы в ML имелись выражения более высоких логических типов, чем в языке L .

Данное условие представляется невыполнимым для ЕЯ и ФЯ, о которых идет речь.

Таким образом, можно сделать вывод о том, что ни в одном достаточно богатом формализме невозможно определить вычислимую функцию алгоритма, перечисляющую множество всех истин и, следовательно, невозможно построить разрешающий алгоритм, распознающий принадлежность высказываний языка множеству истинных утверждений. Исключения составляют производные ФЯ, алфавит которых состоит из элементов абстрагированных относительно АСЭ языков более высоких логических типов, например – язык предикатов.

Применительно к системам распознавания образов, из данного заключения следует невозможность построения алгоритма безошибочно распознающего принадлежность любого образа множеству классов распознавания.

Однако если принять во внимание то, что для решения конкретных практических задач нас могут интересовать не все истины языка, а только истины некоторого вида или из некоторого класса, то имеет смысл построение дедуктики, в которой выводятся все истинные утверждения определенного вида. Это позволяет рассмотреть понятия непротиворечивости и полноты в применении к произвольному подмножеству множества B^∞ . Можно доказать следующую теорему:

если V – перечислимое подмножество множества B^∞ и множество истинных утверждений, принадлежащих к V , перечислимо, то существует дедуктика одновременно полная и непротиворечивая применительно к V .

Доказательство данной теоремы очевидно.

Этот идеальный случай редко встречается при решении практических задач распознавания, где зачастую приходится сталкиваться либо с неперечислимостью подмножества V , либо с неперечислимостью множества истинных утверждений, принадлежащих к V .

Еще более сложной представляется задача логико-лингвистического вывода в системах принятия решений, построенных на основе использования ФЯ близких по своим выразительным свойствам к ЕЯ. Для решения данной задачи используют схемы формализованных знаний, применяемые к структурам представления данных (фреймы, семантические и ассоциативные сети и др. ФЯ).

Противоречие между формой представления знаний и данных и их истинным смысловым содержанием выражается в наличии антиномий – проявлении в ходе рассуждения двух противоречащих, но представляющих одинаково обоснованными суждений.

Антиномии подразделяются на логические и семантические, парадоксы. Именно известный парадокс "лжеца" лежит в основе доказательства теоремы Геделя о неполноте формальных аксиоматических теорий. Антиномии возникают в рамках некоторой формализации процесса рассуждения и свидетельствуют об ограниченности этой формализации. Они выдвигают задачу перестройки данной формализации. Решение антиномий означает введение новой более полной формализации, лучше соответствующей отображаемому содержанию.

Таким образом, источником антиномии является отсутствие либо неполнота представления информации о проблемной области. Законно сделать вывод о том, что эта же причина является источником существования недоказуемых истинных утверждений в ЕЯ и ФЯ, о которых идет речь. Ввиду того, что данные формализмы содержат непополнимые дедуктики, т.е. невозможно построить их расширения в рамках самого формализма (это вытекает из теоремы Геделя и приведенной формулировки Тарского), то проблема кажется неразрешимой.

Однако, как указывалось выше, при рассмотрении СО не учитываются их денотативная (описательная) информационная составляющая, формируемая в процессе распознавания образов [4]. Денотативные элементы S_i , полученные в результате детектирования структурных элементов распознаваемого образа, в процессе последовательных этапов абстрагирования образуют многоуровневую иерархическую систему A обобщенных элементов распознавания (абстрактов) a_i . Вершине данной системы в процессе обучения ставится в соответствие СО. Совокупность пересекающихся систем A_i обеспечивает транзитивные взаимосвязи элементов множества СО на уровне сформированных абстрактов. Данные взаимосвязи определяют родовидовые, причинно-следственные, пространственно-временные отношения, в которых находятся распознаваемые образы и, следовательно, поставленные им в соответствие СО.

Таким образом, представляется возможным создание языка абстрактов AL. Алфавитом данного языка является множество детекторов распознавания структурных элементов образов, а выражения данного языка представляют собой множество абстрактов высших уровней системы A , образующих концептуальную структуру распознаваемых образов. Выражениями языка AL являются также абстракты высшего уровня системы A , которым ставятся в соответствие СО, и подмножества взаимосвязанных абстрактов этого уровня. Данный язык обладает признаками формализма теории типов:

1) наличием отношений типа $y(x_1, \dots, x_n)$, где y выполняется на

x_1, \dots, x_n ;

- 2) расчленением предметной области на слои или типы, образующие иерархию типов;
- 3) наличием теоретико-типовых аксиом свертывания (или их эквивалентов).

Можно предположить, что $AL \supset NL$ в каждой своей части, где NL – язык, словами которого является множество CO , а выражениями – множество определенным образом взаимосвязанных CO . Тогда, при определении перечислимого множества детекторов и перечислимого множества CO , возможно построить полную и непротиворечивую дедуктику для языка AL , причем данная дедуктика будет пополнимой.

Важной особенностью данного подхода является то, что множество распознаваемых образов, как и множество выражений, состоящих из взаимосвязанных CO , могут быть неперечислимы.

Выводы. Таким образом, на основе приведенных рассуждений можно сделать вывод о том, что причиной существования недоказуемых истинных утверждений в выразительных формализмах является неполнота представления информации о проблемной области. Путем решения данной проблемы является создание языка абстрактов AL с определенными свойствами. Данный язык может служить основой построения систем распознавания образов.

ЛИТЕРАТУРА

1. Хант Э. *Искусственный интеллект*. – М.: Мир, 1978. – 558 с.
2. Успенский В.А. *Теорема Геделя о неполноте*. – М.: Наука, 1982. – 111 с.
3. *Философия, логика, язык: Пер. с англ. и нем. / Сост. и предисл. В.В. Петрова; Общ. ред. Д.П. Горелова и В.В. Петрова*. – М.: Прогресс, 1987. – 336 с.
4. Паржин Ю.В., Кайбышев М.С., Гайдаров С.Ю. *Построение элементов концептуальной структуры распознаваемых образов и принятие решения о распознавании // Информационные системы*. – Х.: НАНУ, ПАНИ, ХВУ. – 1999. – С. 106 – 108.

Поступила 23.05.2003

ПАРЖИН Юрий Владимирович, кандидат технических наук, старший научный сотрудник, начальник НИЛ ХВУ. Окончил в 1982 году ХВВКИУ. Область научных интересов – распознавание образов.

НОВИКОВ Владимир Иванович, кандидат технических наук, заместитель начальника кафедры ХИ ВПС. Окончил в 1983 году ХВВКИУ, в 1994 году – ВА им. Дзержинского. Область научных интересов – системы управления.

КУВШИНОВ Константин Владимирович, ст. научный сотрудник НИЛ ХВУ. Окончил в 1973 году Куйбышевский авиационный институт. Область научных интересов –

распознавание образов.