

АНАЛИЗ ОТВЕТОВ ТИПА ФОРМУЛА В ОБУЧАЮЩЕЙ СИСТЕМЕ ПО МАТЕМАТИКЕ

О.А. Козлов, В.И. Сердюков, Е.В. Садков

Основными данными в обучающей системе по математике являются формулы или математические выражения, так как с их помощью записывается большинство законов и теорем математики, а математические задачи решаются, как правило, с помощью преобразований формул.

Концептуальные знания в математике – это в основном равносильности, формулы и тождества. Например, равносильность $\lg(x) = y \Leftrightarrow x = 10^y$ – знание экспоненцирования, формула $a^2 - b^2 = (a - b)(a + b)$ – знание о преобразовании разности квадратов в произведение, тождество $\arcsin\left(\frac{1}{2}\right) = \frac{\pi}{6}$ – знание табличных тригонометрических углов. Эти знания можно представлять в виде $A \rightarrow B$, где левая часть формулы (А) – предложение-образец, а правая – результирующая формула. На этих знаниях, представленных своего рода продукциями, осуществляется вывод формул, заключающийся в сопоставлении преобразуемой формулы с А и, в случае успешного сопоставления, преобразовании данной формулы в формулу, соответствующую правой (В).

Чаще всего вывод на такой продукционной базе знаний бывает прямой (от данных к поиску цели) или обратный (от цели для ее подтверждения – к данным). Прямой вывод используется при преобразованиях выражений в операционной части системы, а обратный – для доказательства обнаруженной ошибки пользователя в контролирующей части обучающей системы [1].

При представлении формул различают две составляющие: визуальную (достаточную для демонстрационной системы и записи теорем) и логическую, отвечающую за математическое содержание (необходимую для оперирующей системы и преобразований формул).

Рассмотрение математического содержания формулы имеет смысл только для синтаксически правильно записанных арифметических выражений. Синтаксически правильно записанным арифметическим выражением называется выражение, содержащее математические константы, переменные, числа, математические операции, скобки. При этом выражение должно быть правильно структурировано относительно скобок, операция должна иметь нужное число аргументов. Язык правильных арифметических формул описывается грамматикой $\langle E, S, x ::= y \rangle$, где: E – множество терминальных символов алфавита; S – множество нетерминальных символов (метасимволов); $x ::= y$ – множество правил вывода.

В простейшем виде для операций $+, *, -, /$ правила вывода представляются следующим образом [6]:

$$\begin{aligned} \langle \text{формула} \rangle & ::= \langle \text{терм} \rangle \mid \langle \text{терм} \rangle \langle \text{плюс минус} \rangle \langle \text{формула} \rangle \\ \langle \text{терм} \rangle & ::= \langle \text{элемент} \rangle \mid \langle \text{элемент} \rangle \langle \text{умножить разделить} \rangle \langle \text{терм} \rangle \\ \langle \text{плюс минус} \rangle & ::= + \mid - \\ \langle \text{умножить разделить} \rangle & ::= * \mid / \\ \langle \text{элемент} \rangle & ::= (\langle \text{формула} \rangle) \mid \langle \text{число} \rangle \mid \langle \text{переменная} \rangle \end{aligned}$$

Дальнейшие расширения приведенной грамматики заключаются в добавлении знаков операций с учетом их приоритета. В общем случае грамматика формул неоднозначная или недетерминированная.

Для рассматриваемой обучающей системы выбрана правоассоциативная грамматика, включающая тригонометрические функции ($\sin, \cos, \text{tg}, \text{ctg}$ и т.д.), операции над множествами (\cup, \cap и т.д.), статистические функции (мода, медиана), функции из начал анализа (предел, производная) и т.д.

Дерево вывода формульной грамматики определяет дерево формулы или математического (арифметического) выражения. Деревья являются наиболее естественным и часто используемым представлением математических выражений. Каждому листу в таком дереве соответствует операнд, а каждому родительскому узлу – операция.

Узлами дерева являются числа, переменные, операции, характеризуемые тройкой $\langle \text{Sign}, \text{Value}, \text{Arn} \rangle$, где: Sign – тип узла, Value – вид узла, Arn – арность.

Для вычисления числовых формул можно использовать так называемый стековый калькулятор [2] и выполнять операции последовательно соответственно префиксной польской записи, однако вычисление числового значения формулы и соответствующее отношение равенства определены только для числовых формул. Значения формул (или функций), содержащих n -переменных, зависят от значений переменных, поэтому вводится понятие P -равенства формул (по аналогии с P -равенством формул алгебры логики). Если в алгебре логики каждой формуле, содержащей n -переменных, можно сопоставить фигуру в P^n , то вычислительное описание формул, зависящих от n действительных переменных можно представить набором точек в R^{n+1} [4]. Другими словами, P -равенство формул – это равенство множества значений функций, задающих формулы, на заранее определенном или случайно выбранном наборе значений переменных [4; 5].

Формально определение P -равенства записывается следующим образом: Рассмотрим некоторое множество точек в R^n : $P = \{P_1, P_2, \dots, P_m\}$.

Пусть $A(x_1 \dots x_n)$ – формула, представляющая некоторую функцию, зависящую от n переменных, P принадлежит области допустимых значений этой функции. Производя подстановки P_i в A , получим для каждого P_i некоторое значение $A(P_i)$ принадлежащее R . Полученные числа $A(P_i)$ задают множество F из m точек в R . Формулы A и B назовем P -равными, если множество точек F_1 , полученное при подстановках, соответствующих множеству P , равно множеству точек F_2 , полученному при аналогичных подстановках. Отношение P -равенства используется при контроле знаний для сопоставления формул: ответа пользователя и правильного ответа.

Примеры P -равных формул: $a + b$; $b + a$; $2a + b - a$. Из данных примеров видно, что P -равные формулы имеют различное символическое представление, но являются тождественно равными.

Ошибки обучаемого, допущенные при записи формулы, можно разделить на два класса: ошибки в символьном вводе (синтаксические ошибки) и ошибки в математике. К первому классу следует отнести некорректно записанные формулы, не представляющие собой правильные арифметические выражения. Например: $6 + ((x + 9)$, x^2 и т.п. Ко второму классу ошибок относятся формулы, корректно введенные пользователем, но не являющиеся правильными ответами. Например: $3 + x$ при правильном ответе $x - 3$ и т.д. Некоторые математические ошибки такого типа могут быть определены по базе знаний ошибок и зачтены как не совсем верные или почти верные. В приведенном примере $x + 3$ и $x - 3$ ошибка может сопровождаться сообщением «Проверьте знак» и выделением числа 3. Задача сопоставления формул связана также с вопросами определения идентичности выражений (или тождественности 0 их разности), задачей классификации математических выражений.

Для сопоставления формул и определения их идентичности предлагаются следующие основные этапы:

1). Сравнение вычислительных характеристик формул, т.е. проверка Р-равенства формул.

Проверка Р-равенства формул определяет эквивалентность формул. В случае не эквивалентных формул с помощью полученных наборов значений F_1 и F_2 можно найти также коэффициент пропорциональности формул. Данные параметры – эквивалентность и коэффициент пропорциональности учитываются при выдаче сообщения пользователю.

2). Сопоставление структуры деревьев при Р-равенстве формул. Данное сопоставление выявляет наличие в введенных формулах недопустимых операций (log, корень квадратный), не содержащихся в правильном ответе, определяет отсутствие необходимых операций (умножение, степень). Далее проводится проверка на упрощенность. Упрощенная формула (по принятому в системе критерию) представляет собой формулу с минимальным вхождением операндов (листьев). При проверке структур деревьев также можно использовать алгоритм проверки изоморфизма деревьев.

Например, при сопоставлении деревьев, соответствующих выражениям $\sqrt{3}+6$ (правильный ответ) и $\sqrt{5}+6$ (введенный ответ), видно, что оба дерева имеют одинаковую структуру, но различный внешний вид. Так как рассматриваемые объекты – деревья (графы без циклов и кратных ребер), то существует взаимно однозначное отображение вершин правильного ответа на вершины введенного ответа. Этот факт позволяет выделить цветом вершины с расхождением и в диалоге предложить обучаемому проверить выделенные таким образом числа.

3). Сравнение правильного ответа и ответа пользователя по базе знаний ошибок.

Литература

1. Агеев В.Н. Электронные учебники и автоматизированные обучающие системы. М.: 2001. 79 с.

2. Грушевский С.П., Архипова А.И. Проектирование учебно-информационных комплексов. Краснодар, 2000.

3. Карпова И.П. Исследование и разработка подсистемы контроля знаний в распределенных автоматизированных обучающих системах: дис. ... канд. техн. наук. М., 2002. 200 с.

4. Подколзин А.С. Компьютерный решатель математических задач. М.: Изд-во ЦПИ при мех.-мат. ф-те МГУ, 2000. 24 с.

5. Проблемы создания автоматизированных обучающих и тестирующих систем: сборник науч. трудов / редкол. А.И. Иванченко и др. Новочеркасск, 2001. 199 с.

6. Программные средства интеллектуальных систем / А.Е. Городецкий, В.В. Дубаренко, И.Л. Тарасова, А.В. Шерсверов. СПб.: Изд-во СПбГТУ, 2000. 171 с.