

ОНЛАЙН РАСПОЗНАВАНИЕ РУКОПИСНЫХ МАТЕМАТИЧЕСКИХ ФОРМУЛ

Надеран Эдрис

Аннотация: рассматривается разработанный алгоритм структурного анализа для онлайн распознавания рукописных математических формул. Вводятся следующие понятия: допустимое и обязательное расположение символа, алгоритм размещения, алгоритм чередования, алгоритм группировки.

Ключевые слова: структурный анализ, математические формулы, распознавание рукописных символов, онлайн распознавание, сегментация.

ACM Classification Keywords: I.5.2 Computing Methodologies - PATTERN RECOGNITION - Design Methodology - Pattern analysis. G.4 Mathematics of Computing – MATHEMATICAL SOFTWARE - Algorithm design and analysis. I.5.1 Computing Methodologies - PATTERN RECOGNITION - Models - Structural.

Вступление

Задачу распознавания математических формул можно разделить на два основных этапа: распознавание символов и структурный анализ. Такое разделение на подзадачи имеет ряд преимуществ, в частности подзадачи могут быть решены независимо друг от друга, что позволяет проанализировать результаты и сделать улучшения на каждом из этапов, при этом сохранится целостность всей системы. Целью данной статьи является подробно рассмотреть разработанный метод структурного анализа, который позволяет определить пространственные отношения между символами и полностью отобразить структуру, написанной пользователем формулы. Предлагаемый алгоритм является гибким к внесению изменений, поскольку пересматривает формулу после каждого внесенного изменения, что обеспечивает дополнительный комфорт для пользователя при введении формулы онлайн. Работоспособность алгоритма была протестирована с помощью разработанного редактора форм. Полученные результаты экспериментов приводятся в третьем разделе статьи.

1 Краткий обзор процесса распознавания

Дадим определение термина математическая формула - это символическая запись законченного логического суждения, которая может состоять из цифр, букв и различного вида символов, при этом все цифры, буквы и символы формируют внутреннюю иерархическую структуру. Следуя из этого определения можно говорить о том, что для того, чтобы правильно распознать формулу необходимо точно знать значение каждого из символов, которые вошли в формулу и определить расположение этих символов относительно друг друга.

В данной статье будет подробно рассмотрен алгоритм структурного анализа, позволяющий решить вторую из поставленных задач, а именно определить расположение символов относительно друг друга. Но сначала кратко рассмотрим процесс сегментации и непосредственно распознавания каждого отдельно сегментированного символа, и после перейдем к описанию самого метода структурного анализа. Такой подход позволит создать полную картину процесса онлайн распознавания рукописных формул. Необходимо заметить, что все три этапа: сегментация, распознавание и структурный анализ происходят друг за другом для каждого отдельно сегментированного образца.

В предлагаемой системе, используется набор из 77 символов, который позволяет писать тригонометрические, логарифмические функции, интегралы, греческие и латинские буквы. В работе

рассматривается система онлайн распознавания, как наиболее перспективная область развития в ближайшие несколько лет, по причине увеличения использования планшетных ПК.

0	7	e	l	s	z	{	.	Ω	$\sqrt{\quad}$	Δ
1	8	f	m	t	+	}	,	π	!	∇
2	9	g	n	u	-	[Σ	α	\approx	λ
3	a	h	o	v	/]	∞	β	γ	σ
4	b	i	p	w	*	<	\geq	μ	θ	\int
5	c	j	q	x	(>	\leq	φ	ε	∂
6	d	k	r	y)	%	=	δ	\in	\exists

Рис. 1 Рассматриваемые символы в системе онлайн распознавания формул

На первом этапе происходит сегментация символов. В рукописном математическом выражении символы почти всегда печатные, не слитные. То есть каждый символ написан отдельно от всех остальных. На этом предположении основывается алгоритм разделения. Необходимо отметить, что при добавлении новых штрихов или удалении уже существующих, может значительно измениться значение символа, поэтому происходит динамическая интерпретация символа после каждого внесенного пользователем изменения.

Для решения задачи распознавания изолированных символов был выбран нечеткий классификатор NEFClass. Система NEFClass прекрасно справляется с необходимой задачей классификации символов, поскольку сочетает в себе нейронные сети и нечеткие системы логического вывода, при этом обладает высокой скоростью работы алгоритмов обучения ([1], [2]). Для задачи распознавания рукописных символов был построен нечеткий классификатор с 3 входами и 77 выходами. Входам системы соответствуют признаки, которые были подобраны специально для математических символов, цифр, латинских и греческих букв, а выходам соответствует набор из 77 символов, которые пользователь может использовать при написании формул. В качестве алгоритма обучения был выбран алгоритм сопряженных градиентов, который обеспечивает сходимость к оптимальному решению за конечное число шагов ([1], [2]). В результате этапа распознавания, на выходе нейронной сети будет получен один из символов.

2 Структурный анализ

Основной задачей этапа структурного анализа является определить расположение символов в формуле относительно друг друга.

В разработанном методе структурного анализа было выделено несколько основных понятий:

- Допустимое и обязательное расположение
- Алгоритм размещения
- Алгоритм чередования
- Алгоритм группировки

Рассмотрим предлагаемый метод более подробно.

Все символы были разделены на группы, в зависимости от допустимых позиций размещения других символов относительно них. На Рисунке 2 схематически изображены все восемь возможных позиций.



Рис. 2 Схематическое изображение допустимых позиций расположения символов относительно текущего рассматриваемого символа

Для разработанного метода были выделены следующие группы символов:

1. Цифры. Принимают значения от 0 до 9, π , ∞ . Символы могут быть размещены относительно данной группы в следующих позициях – 1, 2.

Математические знаки. К этой группе были отнесены такие символы: *, +, -, /, =, <, ≤, >, ≥, %, ≈, €, ∃, Δ,

∇. Допустимо размещение символов во 2 позиции.

2. Скобки открывающие. В эту группу вошли следующие символы: (, {, [. Относительно этой группы символы могут располагаться во 2 позиции.
3. Скобки закрывающие. К этой группе относятся такие символы:), },]. Допустимое размещение символов относительно рассматриваемого символа: 1, 2 позиции.
4. Знаки препинания. К знакам препинания были отнесены: «.», «,», «!». Символы могут располагаться в 1, 2 позиции.

Интеграл \int . Допустимое размещение символов во 2, 4, 8 позиции.

5. Извлечение корня $\sqrt{\quad}$. Допустимое размещение: 1, 2, 6 позиции.
6. Сумма \sum . Допустимое расположение символов: 2, 4, 8 позиции.
7. Греческие и латинские буквы, рассмотренные на Рис.1. Допустимое размещение символов: 1, 2, 3 позиции.
8. Тригонометрические функции. Рассматриваются следующие тригонометрические функции \sin , \cos , \tan/tg , $\cotg/ctn/cot$. Допустимое размещение символов: 1, 2 позиции.
9. Логарифмические функции \ln , \lg , \log . Допустимое размещение символов: 1, 2, 3 позиции.

Под допустимой позицией расположения символов понимается возможная, но не обязательная к заполнению символом позиция. Под термином «обязательная позиция» имеется в виду такая позиция, которая должна быть обязательно заполнена одним из следующих символов, такая позиция получит приоритет при добавлении нового символа. Например, знак \int или $\sqrt{\quad}$ предполагает обязательное наличие зависимого символа в позиции 2.

Алгоритм размещения нового символа относительно других символов работает по следующему принципу.

1. Формула перезаписывается слева направо каждый раз при добавлении нового символа.
2. Выбирается позиция для нового символа А по следующему правилу:
 - Проверяется наличие пустой обязательной позиции, в случае ее наличия данная позиция получает приоритет и в нее помещается символ А.
 - Если пустая обязательная позиция отсутствует, то выбирается ближайшая допустимая позиция, в которую помещается новый символ А.
3. В случае, если позиция, в которую был помещен новый символ А, была занята символом Б, то выбирается новая позиция для символа Б по правилам описанным в пункте 2.

Стоит отдельно рассмотреть знаки «.» и «-». Так точка в выражении может быть десятичной запятой или оператором умножения в зависимости от положения точки и ее соседних символов. Горизонтальная линия может быть дробной чертой или знаком минус в зависимости от длины линии и наличия символов выше и ниже линии. В предлагаемом методе для решения для решения такого рода неопределенности были использованы следующие подходы:

- Знак «-» будет рассматриваться как дробная черта, в том случае, если имеет символы в 4 и/или 8 позициях.
- Знак «.» рассматривается как оператор умножения, если центр масс его сегмента находится ближе к 2 позиции соседнего слева символа.

Несколько букв написанных вместе могут формировать единое значение. Это могут быть тригонометрические функции \sin , \cos , \tan/tg , $ctg/ctn/cot$ или логарифмические функции \ln , \lg , \log . Рассматривая группу букв написанных подряд алгоритм проверяет, не образуют ли они название функции. В этом случае происходит чередование нескольких букв на одну из известных функций. Кроме того, алгоритм делает проверку различных вариантов, так, если на этапе распознавания одна из букв была распознана неправильно, на данном этапе ошибка будет исправлена. На Рис. 3 приведен пример работы описанного алгоритма чередования:

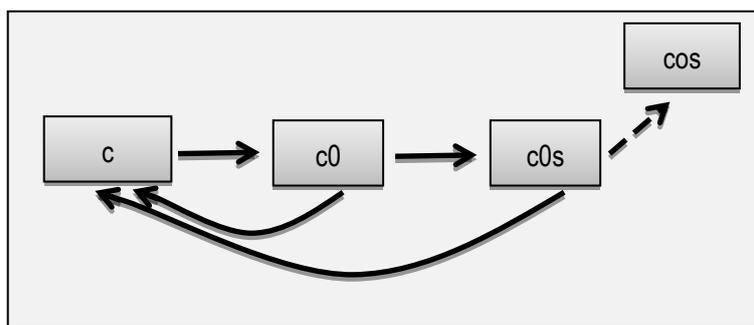


Рис. 3 Пример работы алгоритма чередования

На Рис.3 видно, что каждый раз при добавлении нового символа формула начинает просматриваться сначала, в данном случае буква «o» на этапе распознавания была ошибочно распознана как цифра «0». После добавления следующего символа алгоритм просматривает все возможные комбинации, имеющиеся в базе, и, найдя наиболее подходящую, исправит ошибку.

Следующее понятие, которое было введено в разработанном методе – это понятие группы. Специальные символы, к ним были отнесены различного вида скобки, дробная черта и знаки препинания: точка, запятая, позволяют сгруппировать несколько отдельных значений в одну группу. Алгоритм группировки позволяет корректировать длину скобок и дробной черты, выделить дробное число, как результат такой группировки соответствующая оригиналу формула будет отображена наиболее верно.

Использование описанного метода позволяет определить пространственные отношения между символами и полностью отобразить структуру, написанной пользователем формулы. Кроме, того полный анализ структуры формулы после добавления каждого нового символа делает систему гибкой к внесению изменений.

3 Экспериментальные исследования

Для проверки эффективности работы рассматриваемого метода структурного анализа были проведены тесты в разработанном редакторе формул. Для тестирования были отобраны 50 различных формул, включающих в себя тригонометрические, логарифмические функции, операции возведения в степень, извлечения корня, дробного деления, индексы. На Рис. 4 видно, что редактор хорошо структурирует достаточно сложную формулу с использованием тригонометрических функций.

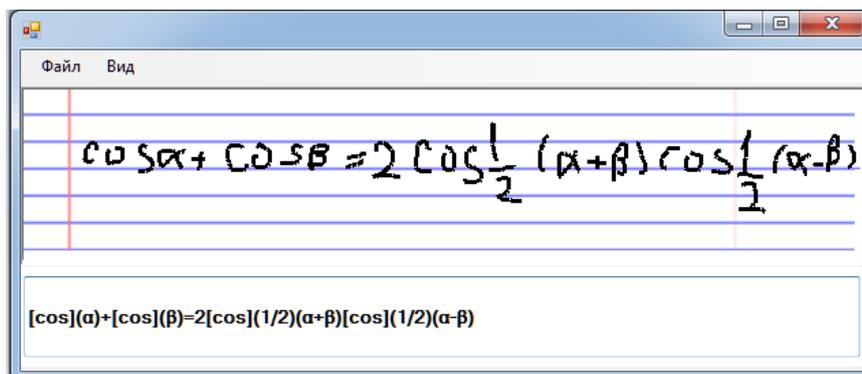


Рис. 4 Пример работы редактора формул

В таблице 1 приводятся результаты работы алгоритма по количеству неправильно структурированных символов. Можно сказать, что такие символы как буквы и знаки операций имеют меньший процент ошибки, в отличие от логарифмических функций и знака суммы.

Таблица 1. Результаты работы алгоритма структурного анализа

	Кол-во символов, содержащихся в протестированных формулах	Кол-во неправильно структурированных символов	% неправильно структурированных символов
Цифры	371	14	6,46
Знаки операций	194	10	5,15
Скобки	92	8	8,69
Знаки препинания	17	3	17,6
Интеграл	25	3	12
Извлечение корня	23	2	8,69
Сумма	15	2	13,3
Греческие и/или латинские буквы	160	7	4,37
Тригонометрические функции	58	4	6,89
Логарифмические функции	15	2	13,3

На Рис. 5 приводится пример неправильно определенной структуры формулы. Здесь алгоритм структурного анализа неправильно определил пространственные отношения между символами. Это ошибка могла произойти по причине слишком близкого расположения значения «2» к полю обязательному полю тригонометрической функции sin.

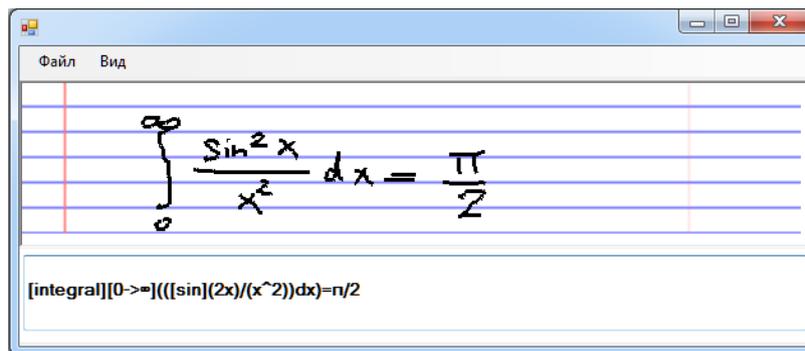


Рис. 5 Пример работы редактора формул

В результате тестирования было получено общее значение ошибочно классифицированных символов 6,7% с проверочного набора данных. Данное значение свидетельствует о достаточной эффективности использования предложенного в работе метода структурного анализа к онлайн распознаванию рукописных формул.

Выводы и перспективы дальнейших исследований

В результате проведенного исследования был разработан алгоритм структурного анализа, который позволяет достаточно точно определить пространственные отношения между символами и как результат полностью отобразить структуру, написанной пользователем формулы. Общее значение ошибочно классифицированных символов равно 6,7% свидетельствует об эффективности его применения к поставленной задаче. Как показали исследования такие символы как логарифмические функции, знак суммы, знаки препинания имеют самый большой процент неправильно структурированных символов. Это можно объяснить наличием нескольких обязательных позиций у данных символов. Необходимо провести дальнейшие исследования, которые позволят улучшить алгоритм. Улучшения будут достигнуты путем доработки алгоритма размещения. На данном этапе модель алгоритма структурного анализа разработана с учетом правила, что пользователь может ввести только одну формулу в редакторе. Для того, чтобы пользователь мог ввести нескольких формул одновременно алгоритм должен быть усовершенствован, введением дополнительных символом, которые будут разделять формулы. Кроме того необходимо разработать систему подсказок, которая позволит пользователю выбирать одну из предлагаемых в списке структур. Предлагаемые улучшения позволят еще больше снизить процент ошибочно классифицированных символов.

Литература

1. Зайченко Ю.П. Нечёткие модели и методы в интеллектуальных системах. Учебное пособие для студентов высших учебных заведений. – К.: «Издательский Дом «Слово»», 2008. – С. 344
2. Зайченко Ю.П. Основы проектирования интеллектуальных систем. Учебное пособие. – К.: «Издательский Дом «Слово»», 2004. – С. 352

Информация про автора



Надеран Эдрис – аспирант Национального технического университета Украины «КПИ», адрес электронной почты: e.naderan@gmail.com

Основные сферы научных исследований автора: применение нечеткого классификатора к онлайн распознаванию рукописных формул.