

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ТЕХНОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ» (МИРЭА)
Проспект Вернадского, д. 78, Москва, Россия, 119454

«УТВЕРЖДАЮ»

Первый проректор
 Соколов В.В.
«29» февраля 2016г.


Отзыв
ведущей организации на диссертацию Петюшко Александра Александровича
«Биграммные языки»
на соискание ученой степени кандидата физико-математических наук
по специальности 01.01.09—
дискретная математика и математическая кибернетика.

Диссертационная работа Петюшко А.А. посвящена изучению и классификации формальных языков, заданных матрицами кратностей биграмм, т.е. пар соседних букв слова изучаемого языка. Еще в начале 20 века выдающимся русским ученым А.А.Марковым (старшим) был создан аппарат цепей, которые позднее получили название цепей Маркова. Последовательность данных, или событий, аппроксимируется с помощью вероятностей переходов в цепи Маркова. Если же построить аналогию между дискретной и вероятностно-непрерывной точкой зрения на цепи Маркова, то аналогом вероятности переходов являются частоты (или кратности) появления пар соседних элементов изучаемой последовательности. Поэтому задача описать формальные языки, имеющие одну и ту же матрицу кратностей биграмм, является в некотором роде обратной к традиционной – нахождения по заданной последовательности вероятностей переходов. При этом классификация изучаемых языков согласно иерархии Н.Хомского формальных языков, или грамматик, является классической и общепринятой практикой в теории изучения формальных языков.

Изучение биграммных языков было начато в семидесятых годах 20 века, когда были опубликованы работы по подсчёту точного числа ДНК-последовательностей, заданных наборами кратностей биграмм и одиночных символов (Дж.Хатчинсон и Х.Уилф, 1975), а также была получена верхняя асимптотическая оценка числа ДНК-последовательностей (К.Ким и Ф.Рауш, 1979). Указанные работы изучают спектральные характеристики исключительно конечных биграммных языков. Все вышеперечисленное позволяет заключить, что изучение бесконечных биграммных языков, в чем и состоит

основная часть диссертационной работы, относится к актуальной теме исследования.

Диссертация состоит из введения, трех глав, заключения и списка цитируемой литературы и публикаций автора. Общий объем составляет 121 страницу.

Во введении дается обзор полученных ранее результатов, относящихся к исследуемой теме, обосновывается актуальность работы, приводятся формулировки основных понятий и полученных в диссертации результатов, а также излагается краткое содержание диссертации по главам.

В первой главе введены основные понятия, касающиеся матрицы кратностей биграмм, конечных и бесконечных биграммных языков. Показаны основные свойства матрицы кратностей биграмм. Получены условия непустоты, конечности и счетности биграммных языков. Приведены критерии регулярности, контекстно-свободности и контекстно-зависимости биграммных языков в терминах эйлеровых циклов графа матрицы кратностей биграмм.

Во второй главе изучаются спектральные характеристики биграммных языков. Получена точная формула мощности множества слов конечного биграммного языка только на основе знания матрицы кратностей биграмм (в работе Дж.Хатчинсона и Х.Уилфа дополнительно предполагалось знание кратностей также и отдельных букв), а также доказана асимптотически точная оценка мощности (в работе К.Кима и Ф.Рауша дается верхняя асимптотическая оценка). В этой главе рассмотрен вопрос о доле матриц, которые задают пустые, конечные или счётные биграммные языки; также рассмотрен вопрос о доле матриц, задающих регулярные, контекстно-свободные или контекстно-зависимые биграммные языки.

В третьей главе изучены так называемые "биграммные языки с закольцовыванием", в которых к набору биграмм, задающему биграммный язык, добавляется биграмма, состоящая из последней и первой буквы рассматриваемого слова. В работе изучены свойства биграммных языков с закольцовыванием, а также построена их классификация. Выясняется, что, несмотря на общую идентичность понятий, есть и различия: так, не бывает конечных биграммных языков с закольцовыванием. В конце третьей главы показано, что разобранные выше задачи о мощности и классификации в случае n -граммных языков (когда задана не двухмерная матрица биграмм, а n -мерная матрица n -грамм) при $n > 2$ сводятся к соответствующим задачам для биграммных языков.

В диссертации получены следующие основные результаты:

1. Введено понятие биграммного языка. Получены аналитические формулы, а также точные асимптотические оценки мощности для простейших биграммных языков, а также получены критерии непустоты, конечности и счётности этих языков.
2. Получены простые графовые критерии для выделения подклассов бесконечных биграммных языков согласно иерархии Н.Хомского: регулярные, контекстно-свободные и контекстно-зависимые. Доказано, что других подклассов нет.
3. Введено понятие биграммных языков с закольцовыванием и установлено взаимно-однозначное соответствие с биграммными языками при одинаковом эйлеровом графе.
4. Получены простые графовые критерии для выделения подклассов конечных биграммных языков с закольцовыванием согласно иерархии Н.Хомского: регулярные, контекстно-свободные и контекстно-зависимые. Доказано, что других подклассов нет.
5. Сведены как задачи о мощности, так и задачи о выделении среди бесконечных языков подклассов согласно иерархии Н.Хомского для языков, заданных n -граммами при $n > 2$, к решённым задачам в случае биграммных языков.

Полученные результаты являются интересными, новыми, снабжены полными математическими доказательствами. В диссертации используются методы классической дискретной математики, теории автоматов, комбинаторики и теории графов.

Основные результаты диссертации опубликованы в 11 работах, все они написаны автором самостоятельно. 7 статей опубликованы в журналах, входящих в перечень ВАК. В диссертации имеются ссылки на все эти работы. Содержание диссертации соответствует содержанию автореферата и опубликованных работ.

Полученные результаты имеют теоретический характер и могут быть использованы в дальнейших исследованиях по изучению формальных языков, заданных матрицей кратностей биграмм. Также результаты диссертации могут быть использованы в прикладных задачах поиска похожих фрагментов данных в системах хранения.

В качестве замечаний к диссертационной работе следует отметить слишком краткое рассмотрение сведения n -граммных задач к биграммным. Данные замечания не снижают общей положительной оценки работы.

Диссертация Петюшко А.А. «Биграммные языки» представляет собой законченное научное исследование на актуальную тему, полностью отвечает требованиям п.п. 9, 10, 11, 13, 14 «Положения о порядке присуждения ученых степеней ВАК Министерства образования и науки РФ» от 24 сентября 2013 г. № 842, предъявляемым к кандидатским диссертациям. Диссертация соответствует специальности 01.01.09 – «Дискретная математика и математическая кибернетика» и удовлетворяет всем требованиям, предъявляемым к кандидатским диссертациям, а ее автор Петюшко Александр Александрович заслуживает присуждения ему ученой степени кандидата физико-математических наук.

Отзыв утвержден на заседании кафедры ОП-5 «Информатика» 16.02.2016 г., протокол № 4.

Голосовали: За - 9, против - 0, воздержался - 0.

И.о. заведующего кафедрой
ОП-5 «Информатика»



Карташов С.И.