

**РАЗРАБОТКА И РЕАЛИЗАЦИЯ АЛГОРИТМОВ
КОМПЬЮТЕРНОГО АНАЛИЗА МЕТРОРИТМИЧЕСКИХ
ХАРАКТЕРИСТИК РУССКИХ ПОЭТИЧЕСКИХ
ТЕКСТОВ**

Кузнецова Ирина Владимировна

Аспирант

Факультет ФИТ НГУ, Новосибирск, Россия

E-mail: Kirina92@rambler

Научный руководитель —

В литературоведении существует необходимость автоматизации анализа различных уровней структуры стиха, а также автоматизированного составления на основе такого анализа метрических справочников к корпусам стихов, словарей рифм и конкордансов.

Целью настоящей работы является разработка и практическая реализация подходов и технологий для автоматизации комплексного анализа русского поэтического текста. Результаты такого анализа позволят существенно расширить возможности филологов, в том числе уменьшить зависимость качества сравнительного анализа от личной эрудиции исследователя, а также применять различные методы интеллектуального анализа данных. В работе представлена реализация программного модуля, отвечающего за анализ структурного уровня – выделение метrorитмики стихотворения (метр, стопность и рифма) в русских поэтических текстах. Алгоритм определения метра и стопности учитывает проблему неоднозначности расстановки ударений в омографах, возможные пропуски схемных ударений (пиррихии), наложение сверхсхемных ударений (спондеи), которые решаются методом “по аналогии”. Описан алгоритм выявления случаев переноса ударения с одной части речи на другую (проклитик). Представлен алгоритм поиска рифм, результатом работы которого является установление строфики стихотворения. Простейший алгоритм определения размера был предложен в работе [1]. Именно этот алгоритм и используется в разрабатываемой в ИВТ СО РАН системе автоматизированного анализа поэтических текстов [3]. Однако в работе [2] был предложен намного более подробный алгоритм, учитывающий некоторые возможные нарушения стандартного размера и перечисленные выше проблемы. Реализация этого алгоритма является основной целью данной работы.

Для реализации алгоритма [1] была произведена:

1. акцентуация слов в тексте согласно словарю А.А.Зализняка,
2. разбиение слов в тексте на слоги,
3. получение метрико-ритмической схемы стиха,
4. сравнение полученной метрико-ритмической схемы с набором метрико-ритмических шаблонов.;

Другой задачей настоящей работы являлось тестирование реализуемого алгоритма и алгоритма системы автоматического анализа поэтических текстов [3], созданной в ИВТ СО РАН, а также сравнение результатов их работы. В качестве тестовой выборки был использован корпус текстов лирических произведений А. С. Пушкина, предварительно размеченный по метру, стопности, рифме.

Проект реализуется на языке программирования Python 3.4.

Литература

1. Бойков В. Н., Каряева М. С., Соколов В. А., Пильщиков А. И. Об автоматической спецификации стиха в информационно-аналитической системе // CEUR Workshop Proceedings. 2015. V. 1536. P. 144-151.
2. Козьмин А.В. Автоматический анализ стиха в с системе Starling / Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции "Диалог 2006"(Бекасово, 31 мая – 4 июня 2006 г.). М.: Издательский центр РГГУ, 2006. С. 265-268.
3. URL: <http://poem.ict.nsc.ru>