

**0.1. Смирнова А.Ю. Автоматическое аннотирование с использованием лексических цепочек**

На сегодняшний день в сети Интернет содержатся миллиарды документов, и это число увеличивается с экспоненциальной скоростью. Для того чтобы справиться с этим объемом информации, человеку необходимы инструменты, обеспечивающие своевременный доступ к различным ее источникам и ее систематизацию. Так и родилась идея автоматического аннотирования и реферирования [1]. К сожалению, большинство существующих алгоритмов ориентированы на работу с англоязычными текстами и не применимы к текстам на русском языке.

В ходе данной работы были рассмотрены извлекающие методы аннотирования [2] и реализован метод, использующий лексические цепочки. Лексические цепочки являются попыткой представления темы, которой посвящен текст или его отрывок. Они позволяют отразить семантическую близость между именными группами, с помощью чего можно определить значимость предложений, выбираемых для аннотации [3].

В результате был разработан алгоритм аннотирования, использующий в качестве критерия выбора предложения наличие упоминания элемента частотной лексической цепочки. Также введен ряд критериев для поддержания связности аннотации, основным из которых является упоминание элементов ранее обработанных цепочек наряду с введением в аннотацию новой информации. Анализ результатов работы полученной программной системы показал, что такой подход к аннотированию является эффективным, так как использует знания о мире при обработке текста. Однако узким местом является сильная связь с тезаурусом предметной области и от качества его составления.

*Научный руководитель – к.ф.-м.н. Крючкова Е.Н.*

**Список литературы**

- [1] NENKOVA A., MCKEOWN K. Automatic Summarization // Foundations and Trends® in Information Retrieval. — 2011. — Vol. 5, No 2-3, P. 103–233.
- [2] Осминин П.Г. Современные подходы к автоматическому реферированию и аннотированию // Вестник Южно-Уральского государственного университета. Серия: Лингвистика. — 2012. — №25.
- [3] Лукашевич Н.В., Добров Б.В. Автоматическое аннотирование новостных кластеров на основе тематического представления // Труды Международной конференции «Диалог-2009». — 2009. — С. 299–305.