

# АЛГОРИТМЫ ПОИСКА ПОЛНОСТЬЮ ПРЕДСТАВЛЕННЫХ ФРАГМЕНТОВ ЗНАНИЙ ДЛЯ МОДЕЛИ УНИФИЦИРОВАННОГО СЕМАНТИЧЕСКОГО ПРЕДСТАВЛЕНИЯ ЗНАНИЙ

*В работе приводится описание системы системы вычисления канонической формы фрагмента sc-текста, основанной на квазиполиномиальном алгоритме решения изоморфизма графов.*

## ВВЕДЕНИЕ

Важной задачей при внедрении эффективных программных систем обработки знаний является поиск в базах знаний. Проще представляется подход к поиску на основе перехода к каноническим формам структур знаний. Алгоритмы сравнения канонических форм проще, чем сравнение структур общих форм. Кроме того, время, необходимое для расчета канонической формы, переносится на фазу, не зависящую от самого поиска, например, на фазу формирования базы знаний. Представляется, что такой подход позволяет сократить время отклика системы на запросы пользователей.

### I. СИСТЕМА ВЫЧИСЛЕНИЯ КАНОНИЧЕСКОЙ ФОРМЫ ФРАГМЕНТА SC-ТЕКСТА.

Долгое время задача изоморфизма графов алгоритма решения, который в худшем случае работал не лучше полного перебора. В 2015 году Ласло Бабаи представил свой алгоритм решающий данную задачу за квазиполиномиальное время [1]. Данный алгоритм положен в основу разработанной системы вычисления канонической формы фрагмента sc-текста. Такие тексты представляются графами в которых ребра могут соединять как вершины, так и другие ребра. Для упрощения алгоритма sc-текст [2] кодируется классическим помеченным ориентированным графом. Для этого используется подход описанный в статье [3]. Основная идея подхода - поставить в соответствие sc-дуге вершину классического графа, для сохранения направления связи данная вершина соединяется двумя встречными

ребрами с вершиной начала, и одним исходящим с вершиной конца связи. Для представленного таким образом графа вычисляется каноническая форма. Она задаёт такую перестановку вершин, что матрицы смежности двух изоморфных графов становятся идентичными. Это позволяет, зная каноническую форму, определять изоморфность двух графов за  $O(N^2)$ , где  $N$  количество вершин. Кроме того, каноническая форма задаёт однозначное представление матрицы смежности одномерной строкой. Что позволяет проводить поиск в библиотеке размеченных графов за логарифмическое время при использовании древовидных структур данных или ещё быстрее при использовании хеш-таблицы.

## ЗАКЛЮЧЕНИЕ

Разработанная система способна позволяет вычислять канонические формы фрагментов sc-текстов. Результат ее работы позволяют решать задачу нахождения дублирования библиотеке таких фрагментов.

## Список литературы

1. Babai L. Graph isomorphism in quasipolynomial time // Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. – 2016. – С. 684-697.
2. Представление и обработка знаний в графодинамических ассоциативных машинах // В. В. Голенков [и др.] – Минск. – 2001. – 413 с.
3. Представление семантических сетей и алгоритмы их организации и семантической обработки на вычислительных системах с массовым параллелизмом. / В. П. Ивашенко [и др.] // Open Semantic Technologies for Intelligent Systems (OSTIS-2015) : материалы V междунар. науч.-техн. конф. / БГУИР. – Минск 2015 – С. 133-140.

*Бруцкий Дмитрий Сергеевич*, магистрант факультета информационных технологий и управления Белорусского государственного университета информатики и радиоэлектроники, dzmitry.brutski@gmail.com

*Научный руководитель: Ивашенко Валерьян Петрович*, доцент кафедры интеллектуальных информационных технологий Белорусского государственного университета информатики и радиоэлектроники, кандидат технических наук, ivashenko@bsuir.by