

РАСПРЕДЕЛЕННАЯ ОБРАБОТКА ИНФОРМАЦИИ, КОРПОРАТИВНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ МНОГОМЕРНОГО АНАЛИЗА МИКРОЧИПОВЫХ ДАННЫХ: РЕАЛИЗАЦИЯ ФУНКЦИЙ БАЗЫ ДАННЫХ И ПРЕОБРАЗОВАНИЯ ФОРМАТОВ ДАННЫХ

И. А. Агбаш

Институт цитологии и генетики СО РАН
Новосибирский государственный университет

За счет активного роста технологических достижений в области молекулярной биологии появилась возможность проводить широкомасштабные исследования с помощью микрочипов. Но, несмотря на огромное количество программных средств для анализа микрочиповых данных, оказалось, что на практике ими не очень удобно пользоваться из-за отсутствия технических знаний и трудностей освоения интерфейса. Помимо того, большинство таких программных средств являются коммерческими и не могут быть изменены таким образом, как это необходимо пользователю. Поэтому перед нашей командой разработчиков встала актуальная проблема разработки программного комплекса, достаточного понятного и доступного для биолога.

Все исполняемые модули программы являются независимыми и вызываются при помощи программы диспетчера. Для повышения надежности работы программы предусматривается только пакетная обработка. Обмен информацией – только через текстовые файлы. В основе формата данных лежит представление, что, в отличие от входных языков баз данных, первична не структура, в которую укладываются данные, а, наоборот, некоторые первичные элементы данных в совокупности образуют структуры более высшего порядка, причем одни и те же данные, в зависимости от способа рассмотрения, задаваемого описанием, могут образовывать разные структуры.

Моя изначальная часть работ – это организация функций базы данных внутри программного комплекса, а именно: ввод, вывод таблиц, замена ключей и значений, транспонирование, сортировка, перестановка, выбор строк или столбцов, разделение и слияние таблиц, преобразование векторов в матрицы и обратно, развертка группирующих переменных, а также совместимость с MS Excel (поддержка формата .csv). Все это делается для того, чтобы данный программный комплекс являлся независимым от каких-либо других программных продуктов и пользователь мог задавать всю необходимую обработку, свободно оперируя своими данными и не выходя за пределы программного комплекса.

Научный руководитель – д-р биол. наук, доцент. В. М. Ефимов