Математические структуры и моделирование
1999. Вып. 3, с.141-142.

УДК 530.51

О ЧИСЛОВОМ АНАЛИЗЕ СТРОЕНИЯ ЗНАКОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ И МАССИВОВ УПОРЯДОЧЕННЫХ ДАННЫХ

А.С. Гуменюк

Number analysis of structure of symbol sequences and massives of ordered datas is discussed

Любая знаковая последовательность представляет особое (по)строение элементов по линейно расположенным координатам-позициям. Строение конкретной последовательности задано мощностью алфавита, распределением чисел вхождения его элементов, удалением позиций одинаковых и разных знаков друг от друга и от финишной позиции наблюдателя. Таким образом, строение последовательности определяется составом и композицией элементов.

Известные математические средства применяются для анализа локальной структуры и не приспособлены для исследования полного строения знаковых последовательностей, других массивов данных. Предлагается в качестве первичной информационной компоненты последовательности использовать, кроме собственно знака, интервал между соседними «одинаковыми» знаками или «словами». Перемножение всех интервалов такой однородной последовательности определяет объем «символа» или «понятия»; относительный объем символа определяется от позиций другого знака; перемножение всех объемов символов определяет объем целого «произведения».

Логарифмирование представленных величин дает набор удобных для практики аддитивных информационных характеристик (мер). При этом интервал заменяется «удаленностью», объемы – суммарными удаленностями, соответственно названными: глубиной «содержания» символа или понятия и глубиной содержания произведения. Полезны оценки в виде средних арифметических глубин содержаний (отображающих средние геометрические интервалы) соответственно – «содержательности» символа, понятия и целого произведения. Отношение содержательности символа и произведения дает характеристику, которая дополняет частоту вхождения и названа относительной содержательностью символа. При этом для любой полной знаковой последовательности (без

© 1999 А.С. Гуменюк
Омский государственный технический университет

пустых позиций) сумма произведений частот вхожения элементов на их относительные содержательности равна единице. Содержательность целого произведения (средняя удаленность знаков) является функцией, которая «перекрывает» функцию энтропии, так как чувствительна к расположению знаков в последовательностях с одинаковым их составом.

Учет «весов» символов позволяет получить еще более чувствительные к расположению знаков информационные характеристики последовательности.

При определении предлагаемых характеристик для произвольного массива данных требуется его предварительное преобразование в соответствующую конечную знаковую последовательность.