

Построение алгоритмов подсчета числа единиц в двоичном слове

Зинченко И.С., студент; Кулик И.А., доцент
Сумский государственный университет, г. Сумы

Подсчет числа двоичных единиц является одной из самых распространенных операций над двоичными последовательностями. Данная операция широко используется в различных кодовых приложениях, например при формировании равновесных и квазиравновесных кодов, кодов Бергера и Левенштейна, АВ-сжатии, а также при решении ряда программных задач.

Тривиальный алгоритм подсчета – последовательный просмотр каждого разряда двоичного слова, и если он единичный, переменной единиц задавать соответствующее приращение. Недостатком такого алгоритма является значительное время (равное n , где n – число разрядов слова a_i) подсчета единиц. Данное время можно уменьшить, если ввести еще одну переменную – переменную нулей и одновременно с подсчетом единиц проводить подсчет нулей.

Более быстрым способом вычисления числа k единиц является способ, основанный на арифметических и логических операциях вида

$$s_k \leftarrow s_k \wedge (s_k - 1), s_k \leftarrow a_i \text{ и } s_m \leftarrow s_m \wedge (s_m - 1), s_m \leftarrow \bar{a}_i. \quad (1)$$

Алгоритм, использующий операции (1), эффективен для "разреженных" слов, т.е. двоичных слов с малым числом единиц (или нулей), и время его работы будет $\min(k, n - k)$. Недостатком алгоритма является зависимость времени его работы от значения k .

Перспективным способом подсчета числа k единиц является способ, основанный на том, что несколько операций сложения разрядов в коротких частях производятся путем одной операции в длинных частях двоичного слова. При этом последовательно разбивают исходное слово на равные части, представляя их как двоичные записи чисел единиц, далее, дополняют полученных части нулями до длины исходного слова и складывают. В этом случае время подсчета единиц соответствующим алгоритмом составляет постоянную величину $\lceil \log_2 n \rceil$.