Федотов Валерий Павлович

ИНСТРУМЕНТ ДЛЯ ВЫЧИСЛЕНИЙ В ТЕОРЕТИЧЕСКОЙ ИНФОРМАТИКЕ

***Интервальная система счисления** — информационная система счисления, в которой все подмножества числовой оси, определенные несколькими первыми цифрами записи любого числа, являются интервалами.*

В числе атрибутов того или иного этапа развития человеческой цивилизации можно указать на диапазон чисел, востребованных исторической эпохой. Соответственно развиваются языковые средства (прежде всего, количественные числительные) и способы записи чисел – системы счисления. Счет первобытных народов редко выходит за пределы первого десятка. Анализ представления чисел с помощью римских цифр явно указывает на то, что древним римлянам приходилось иметь дело с числами, едва превышавшими 10^3. К концу XX века этот диапазон достиг 10^{80}. Именно таков предел для представления чисел в компьютерах и микрокалькуляторах (если не прибегать к специальным ухищрениям). Примерно такова же оценка для числа элементарных частиц во вселенной. Поэтому естественные науки никогда не потребуют чисел вне этого диапазона, а ни «традиционная» техника, ни экономика к его границе пока даже и не приближаются. Наконец, именно в районе 10^{80} обрывается список имен числительных, которые можно встретить в словарях.

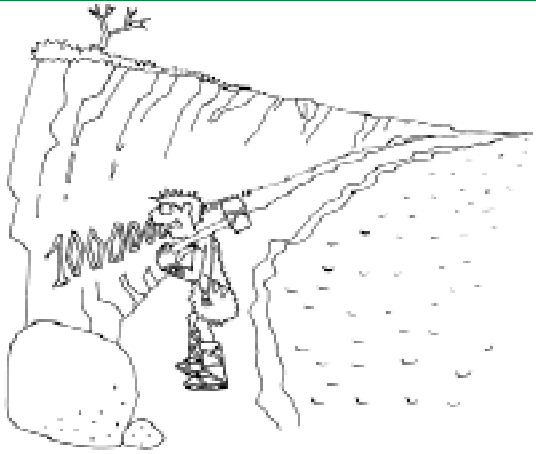
Счет первобытных народов редко выходит за пределы первого десятка.

Но между тем рубеж тысячелетий стал рубежом исторических эпох: произошла информационная революция. Наступившее время характеризуется тотальным внедрением компьютерной техники во все сферы человеческой жизни. Соответственно возникает потребность и в резком расширении диапазона чисел. Ее стимулируют теоретические исследования в области информатики, криптографии и ряда смежных дисциплин, а также борьба с катастрофической потерей точности [1], возникающей в процессе вычислений. Безобидный вопрос: сколькими способами можно заполнить информацией стандартную дискетку, – приводит к числу около 10^{350000}, запись которого в десятичной системе займет более ста толстых томов. Поэтому и встал вопрос о новых системах счисления.

Один из основателей информатики Джон фон Нейман доказал теорему о том, что среди всех основных позиционных систем счисления именно троичная система счисления позволяет наиболее эффективно сворачивать информацию о вещественном числе. Этот факт базируется на том, что именно 3 ближе всех целых чисел к основанию $e \approx 2{,}718$ натуральных логарифмов. Названный эффект можно усилить, если вместо традиционных систем счисления использовать более сложные конструкции башенных [2], итерационных [1] и интервальных [4] систем, построенные автором и моими учениками.

В отличие от традиционных, в основе только что упомянутых систем счисления лежит информационный принцип: каждый очередной бит последовательно уточняет информацию о положении точки на числовой оси [4]. Это становится важным, например, при параллельных вычислениях: первые цифры числа можно передавать очередному этапу алгоритма еще задолго до того, как найдены последующие цифры. При традиционной же записи числа первые цифры вообще не несут никакой информации о величине числа до тех пор, пока неизвестен его порядок или положение разделяющей точки (однако все последующие цифры вполне соответствуют этому требованию).

Первый бит информации о числе чаще всего совпадает с его знаком. Он служит ответом на вопрос о сравнении числа с нулем. В качестве второго бита разумно взять знак порядка (то есть знак логарифма; причем совершенно безразлично, идет ли речь о натуральных, десятичных, двоичных логарифмах, либо по другому основанию, большему единицы). Этот бит служит ответом на вопрос о сравнении положительного числа с 1 или отрицательного числа с −1.



Безобидный вопрос: сколькими способами можно заполнить информацией стандартную дискетку, — приводит к числу около 10^{350000}

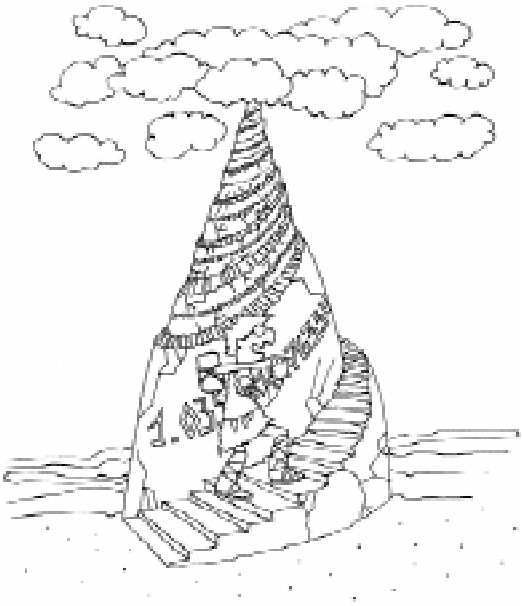
В качестве последующих битов записи числа можно брать ответ на вопрос о сравнении данного числа с числами из некоторой последовательности. Проблема в том, как задать саму последовательность. Алгоритм формирования последовательности нужно выбрать заранее, хотя конкретные ее члены будут строиться уже в зависимости от числа.

Все члены всех таких последовательностей с учетом порядка, в котором они могут появиться, удобно представить в виде дерева базовых точек данной системы счисления.

Чтобы получить какую-либо итерационную систему счисления [2], в качестве чисел для сравнения берут корни последовательных итераций нужной монотонной функции. В частности, если эта функция — логарифм, основание которого не меньше $e^{1/e} \approx 1{,}4447$, то получится башенная система счисления [3]. Именно башенные системы счисления позволяют убить двух зайцев. Во-первых, они лучше всего усиливают эффект названной теоремы фон Неймана. Во-вторых, сравнительно небольшим количеством цифр они позволяют записать числа, абсолютная величина которых не доступна не только словесной формулировке, но даже богатому воображению. Как и в уравновешенной троичной системе *, в ин-

> ***Итерационная система счисления** — интервальная система счисления, в которой в качестве точек разбиения (границ интервалов) выбираются корни последовательных итераций некоторой монотонной функции.*

* Об уравновешенной троичной системе читайте в статье «Троичная система и равновесие» Агафоновой И.В. и Дмитриевой О.М. в этом номере журнала.



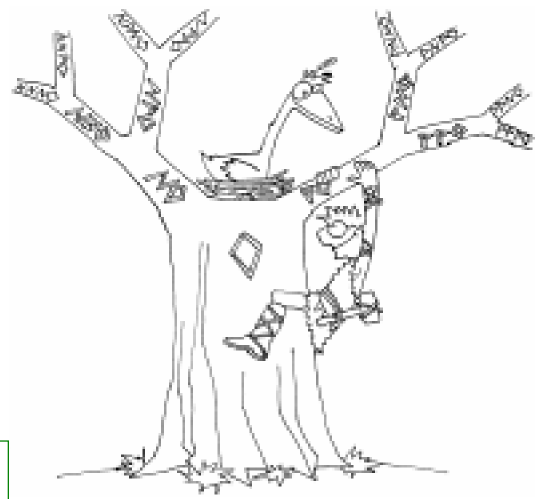
Именно башенные системы счисления позволяют убить двух зайцев.

формационных системах счисления в роли цифр выступают буквы N, O, P со значениями -1, 0 и $+1$ (с учетом модификации).

Базовой (узловой) точкой системы счисления называется число, которое имеет в этой системе счисления конечную (по количеству цифр) запись. Минимально возможное число цифр такой записи (без учета крайних нулей и вспомогательных знаков) называется уровнем базовой точки. Например, в десятичной системе счисления дробь $1/2 = 0{,}5$ является базовой точкой первого уровня, дробь $1/4 = 0{,}25$ является базовой точкой второго уровня, а дробь $1/3 = 0{,}333333...$ вообще не является базовой точкой.

Если же речь идет об интервальных системах счисления [4], то их базовые точки естественным образом (с учетом как неравенств между ними, так и вхождений записи одного числа в качестве подстроки в запись другого числа) выстраиваются в двоичное дерево. Вершиной дерева (и единственной на нулевом уровне базовой точкой) служит число, записываемое одной лишь цифрой O (в любой башенной системе счисления это 0). На первом уровне лежат два числа – NO и PO (в любой башенной системе счисления это -1 и 1). На втором уровне лежат четыре числа – NNO, NPO, PNO, PPO (они равны $-d$, $-1/d$, $1/d$ и d, где через d обозначено основание башенной системы счисления). На третьем уровне лежат восемь чисел – $NNNO$, $NNPO$, $NPNO$, $NPPO$, $PNNO$, $PNPO$, $PPNO$ и $PPPO$ (равные соответственно $-d^d$, $-d^{1/d}$, $-1/d^{1/d}$, $-1/d^d$, $1/d^d$, $1/d^{1/d}$, $d^{1/d}$ и d^d). Аналогично на каждом последующем уровне число базовых точек удваивается (на k-ом уровне имеем 2^k точек). Запись каждого из этих чисел содержит $k+1$ цифру. Последней цифрой записи является O, а k предыдущих цифр могут быть любой последовательностью цифр N и P.

В случае башенных систем счисления [5] множество базовых точек имеет довольно интересную геометрию. Оно обладает обоими атрибутами фракталов: самоподобием и дробной размерностью. Причем оба свойства относятся не только к множеству базовых точек, но и продолжаются на свя-



Если же речь идет об интервальных системах счисления, то их базовые точки естественным образом ... выстраиваются в двоичное дерево.

***Башенная система счисления** – итерационная система счисления, в которой каждый очередной бит в записи числа имеет смысл знака логарифма от абсолютной величины мантиссы, полученной на предыдущем шаге.*

зи между ними (иерархия дерева и неравенства).

В последние годы все основные параметры компьютерной техники ежегодно вырастали в среднем в 4 раза (в тысячу раз за 5 лет). Ясно, что их экспоненциальный рост не может продолжаться бесконечно (в частности, этому препятствует и атомное строение вещества). Как только он прекратится, производители начнут искать другие пути усовершенствования компьютерной техники. И тогда они вынуждены будут отказаться от двоичной системы счисления в пользу башенных систем.

Еще одним направлением применения новых систем счисления может стать криптография. Файл любого назначения представляет собой последовательность байтов. Поэтому его можно прочесть как число в системе счисления с основанием 256. Перевод его в любую из информационных систем счисления затрудняет чтение файла. А если еще прочитать затем ту же последовательность цифр из другой системы счисления, то такой способ шифрования создает абсолютную защиту информации.

На диске к журналу публикуется словарь терминов, относящихся к системам счисления, подготовленный Федотовой М.В.

Литература

1. *Федотов В.П.* Башенные системы счисления // Информационные технологии в образовании. СПб.: РГПУ, 1998.

2. *Федотов В.П.* Новые системы счисления как альтернатива интервальным вычислениям // Межд. конф. по мягким вычислениям и измерениям SCM–2003. СПб, 2003.

3. *Баранова Н.В., Федотов В.П.* Итерационные системы счисления // Актуальные проблемы современной науки. Ч. 1. Самара, 2001, с. 21.

4. *Федотова М.В., Федотов В.П.* Интервальные системы счисления // Актуальные проблемы современной науки. Ч. 1. Самара, 2001, с. 55.

5. *Федотова М.В.* Информационные системы счисления // Межд. конф. «Интел-Юниор, 2001». М.: МИФИ, 2001.

Федотов Валерий Павлович,
чл.-корр. МАИ, Санкт-Петербург.