Table I lists the results of solving problems Z1 and Z2 corresponding to integer linear programs with 292 and 532 variables, respectively. We see from the table that the solution time is relatively short even for fairly large problems. It thus seems that the proposed algorithms can be used to solve these problems in computerized transportation scheduling systems.

LITERATURE CITED


PROBABILISTIC ANALYSIS OF FLOATING-POINT ADDITION

A. I. Mikov

The addition of long sequences of numbers is one of the widespread macro-operations for different calculations. Finding a definite integral, estimation of statistical means, calculation of the scalar product of vectors, and other problems include the calculation of sums. Errors associated with equalization of the orders of components are inevitably accumulated in floating point summation, where the most rapid growth takes place in the case of components of one sign.

Within the framework of the probabilistic approach, the characteristics of cumulative error during floating point addition of positive random numbers \( \xi \) that are mutually independent and not certainly identically distributed are studied in this paper. Different addition methods are examined: in arbitrary, decreasing, and increasing order. Since exact compact expressions for the cumulative error characteristics that are suitable for calculation practice are not accessible in the majority of cases, the fundamental results are represented in the form of inequalities or in the form of limit relations.

1. REPRESENTATION ERROR DISTRIBUTION

Numbers are represented in a machine with floating point in the form \( \mu \beta^p \), where \( \mu < 1 \) is the \( m \)-digit mantissa, \( p \) is the order, and \( \beta \) is the base of the counting system. For insertion and round-off the number \( \xi \) is identified with one of the two adjacent normalized machine numbers of order \( p = \lfloor \log \xi \rfloor + 1 \), between which the spacing is \( \beta^{p-m} \). Here \( \lfloor x \rfloor \) is the integer part of the number \( x \).

Besides the insertion error in addition, an additional round-off error occurs because of the necessity to reduce the components to one order. It appears because the number is now represented by one of the two adjacent unnormalized machine numbers of order \( k \geq p \), between which the spacing is \( \beta^{k-m} \) (\( k \) is the order to which \( \xi \) is reduced). We shall examine the round-off process from precisely this viewpoint, i.e., when \( \xi \) is reduced to an arbitrary order \( k \). It is here necessary to take into account that only numbers \( \xi < \beta^k \) can be reduced to the order \( k \); their conditional distribution is \( P \{ \xi \leq x \mid 0 \leq \xi \leq \beta^p \} = F(x) / F(\beta^p) \), \( F(x) = P \{ \xi \leq x \} \) for \( x \leq \beta^k \).

It follows from the above that the error of a number \( \xi \) prepared for addition is distributed in an interval of length \( \tau = \beta^{k-m} \) whose location on the real axis depends on the kind of roundoff.

Three kinds of round-off of positive numbers are of interest: round-off downward when \( D_\xi \leq \xi \) (\( D_\xi \) is the rounded off number), upward when \( U_\xi > \xi \), and round-off to the nearest machine number (n.m.n.). We will find the error distribution in each case.

Let us consider the more general rounding off. Let

$$\delta = \begin{cases} \lfloor \frac{x}{\tau} \rfloor, & \text{if } \{\frac{x}{\tau}\} \leq a, \\ \lceil \frac{x}{\tau} \rceil + 1, & \text{if } \{\frac{x}{\tau}\} > a. \end{cases}$$

Here \(\{x\}\) is the fractional part of the number \(x\). Round-off to the n.m.n. is a particular case \((a = 0.5)\), for round-off upward \((a = 0)\), and downward \((a = 1)\). The error distribution function has the form

$$1 + \sum_{n=0}^{\infty} \frac{|F(n\tau + x) - F(n\tau + ax)|}{F(b^\tau)}, \quad 0 \leq x < a\tau,$$

$$\sum_{n=0}^{\infty} \frac{|F(n\tau + \tau + x) - F(n\tau + ax)|}{F(b^\tau)}, \quad (a-1)\tau < x < 0.$$

Here \(\tau = \frac{a^n}{\tau} = \frac{b^m}{\tau}\). This expression can be simplified if the z-transform \(F(z) = \sum_{n=0}^{\infty} z^{-n} \times F[n(n+1)]\) of the function \(F(x)\) is known:

$$1 + \sum_{n=0}^{\infty} \Phi_e(z, x/\tau) z^{-1}, \quad 0 \leq x < a\tau,$$

$$\sum_{n=0}^{\infty} \Phi_e(z, x/\tau + 1) z^{-1}, \quad (a-1)\tau < x < 0,$$

where \(\Phi_e(z, x/\tau) = [F(z, x/\tau) - F(z, a)]/[F(z) - F(b^\tau)]\).

2. MARKOV CHAIN

Let us turn to a calculation of the cumulative error. Round-off errors occur for each
addition. However, this process is inhomogeneous. The next addition to the cumulative error
is statistically nonconstant and depends on the order of the sum being formed. If the pro-
cess of sum formation \(S_0, S_1, S_2, \ldots, S_n, \ldots\) \(\{S_n = \delta(S_{n-1} + \xi_n)\}\) is considered as a pro-
cess with discrete time, then the times \(\tau_0, \tau_1, \tau_2, \ldots, \tau_k, \ldots\) can be extracted so that \(\eta(S_{k+1}) \geq b^k\) under the condition \(P(S_{k+1}) < b^k\) (here \(P(x)\) is the order of the number \(x\) when it
is represented in floating point form), i.e., \(\tau_k\) is the time of sum intersection with the
level \(b^k\).

Let us consider the time segments between two adjacent intersection times. The equality
\(S_n = \delta(S_{n-1} + \xi_n) = \delta S_{n-1} + \xi_n\) holds in this intersection. In other words, if the order does
not change upon the addition of a new component to a partial sum, then the addition and round-
off operations can interchange places. Hence, there results directly that the error \(\xi_n = \delta(S_{n-1} + \xi_n) - \delta(S_{n-1} + \xi_n) = \xi_n - \xi_n\) originating in addition is independent of the magnitude of
the sum and is governed only by the quantity \(\xi_n\) and the level at which the process is. There-
fore, successive round-off errors, occurring between adjacent intersection times, are inde-
pendent, and in the case \(P(\xi_n \leq x) = F(x)\) are also identically distributed. Hence, the for-
mation of a cumulative error within each level is a (stationary) process with independent
increments. This fact simplifies the analysis substantially.

To evaluate the total cumulative error to the time of intersection of the level \(b^k\), the
residence time (the number of additions \(N_k = \tau_{k+1} - \tau_k\) at each level must be known.

Let the process start with a certain initial value \(w\) with order \(p(w) = v + 1\). Then the
total error at a time directly preceding the intersection of the level \(b^k\) equals

$$E_k = \sum_{i=0}^{N_k} \xi_i^{(i)} + \sum_{i=0}^{N_{k+1}} \xi_i^{(i+1)} + \ldots + \sum_{i=0}^{N_{k-1}} \xi_i^{(k-1)},$$

where \(\xi_i^{(j)}\) is the error of one round-off at the level \(b^j\); \(\xi_0^{(j)} = 0\); \(N_j\) is the time passed by
the process \(\{S_n\}\) at the level \(b^j\) (i.e., from the intersection of \(b^j\) to the intersection of
\(b^{j+1}\)). If the generating functions of the probabilities for the integer random variables \(N_j\):

$$P_j(z) = \sum_{n=0}^{\infty} z^n P(N_j = n)$$

and the Laplace–Stieltjes transforms of the distribution functions

$$\psi_{E_k}(s) = \int_{-\infty}^{+\infty} e^{-sx} dP(E_k \leq x), \quad \psi_j(s) = \int_{-\infty}^{+\infty} e^{-sx} dP(\xi_j \leq x),$$

402