reconstituted an implementation of the recursive algorithm, is to make a leap of insight. It's not at all *obvious*; which it would be, one supposes, if the recursive algorithm had *meant* such a process all the time.

I feel that an overemphasis on the 'problem reduction' approach to program design—encouraged in recent years by the structured programming crusade—may have a deleterious effect on our students' abilities to think in process structure terms, rather than in problem reduction terms. The most pernicious effect of such an overemphasis is a belief that an elegant and economical *program* structure is to be equated with an elegant and economical *process* structure. The contrast between MOVESTEEPLE and HANOI would be a good counterexample to such a belief. I am not arguing here against structured programming, but against certain aspects of Structured Programming.

MOVETOWER is doubly recursive and cannot be made iterative by a trivial syntactic transformation. The nonrecursive HANOI is possible because the data structure being manipulated contains enough information to decide what to do at each stage, making the control structure—the stack—superfluous. One might draw a (rather weak) analogy with the use of threaded trees to eliminate recursion in tree traversing programs, where the data structure is modified to include just the information needed by the recursive algorithm. We might say that HANOI is an optimisation of MOVETOWER, in the same sense that an iterative *factorial* program is an optimisation of the recursive algorithm, or the iterative 'backstitching' list reversing algorithm is an optimisation of the obvious recursive list reverser. But, I suggest, it would be a very nontrivial exercise to convert MOVETOWER to HANOI, let alone convert it mechanically. In fact, to close, I hereby offer it as a challenge to optimistic optimisers, and to those who make it their business to prove that equivalent programs are equivalent.

References

DIJKSTRA, (1971). *A short Introduction to the Art of Programming.*

BURSTALL, COLLINS, and POPPLESTONE. (1971). *Programming in POP-2*, Edinburgh.

SUSSMAN and MCDERMOTT. (1974). CONNIVER reference manual. *Memo 259A*, A. I. Laboratory, MIT.

To the Editor
The Computer Journal

Comments on 'Fitting data to nonlinear functions with uncertainties in all measurement variables'

Recently in this *Journal*, W. H. Southwell (1976) presented an expansion of an earlier paper (Southwell, 1969) dealing with the important topic of curve fitting and parameter estimation when the function is nonlinear and when errors are assumed to be present in both the dependent and independent variables. His expansion is, fortunately, largely a clarification of many of the misleading statements and a correction of several errors in his 1969 article (See Powell and Macdonald, 1972). Since much of the 1976 paper appears to be an attack on the least squares method published earlier by the present authors (Powell and Macdonald, 1972), it is desirable that a reply be made so that further confusion may be avoided.

1. We *still* claim that, contrary to Southwell's assertion (p. 69), his 1969 method will not in general converge, for *nonlinear* models, to the least squares solution. We invite details of the 'highly successful' (p. 70) use of the 1969 algorithm for such models. None has yet been given by Southwell. It is important here to appreciate the distinction, glossed over by Southwell, between the method he described in 1969 and that described in his 1976 paper.
2. The proof that the expressions for 'exact' parameter variances given in the 1969 article were themselves approximations for all but linear functions was supplied to Southwell by one of us (Macdonald, private communication, 1973). Southwell's fundamental error in statistical analysis is fortunately cleared up in the recent (1976) paper.
3. Our error of $\sqrt{2}$ in parameter standard deviations was corrected in print in 1974 and in 1975, apparently too late for Southwell to acknowledge in his 1976 paper.
4. Southwell is incorrect (p. 71) in asserting that our method depends on convergence of only the a's. In fact, our method depends on convergence of both the x's and a's, as shown by our Eqs. 2 and 3.
5. Southwell's wording is occasionally misleading. For example, on p. 71, it would have been more appropriate to have said: 'The method here described is, with one difference, equivalent to the earlier Powell and Macdonald method.'.
6. Since Southwell has still not provided any illustrations showing convergence to a least squares solution of his 1969 method for functions non-linear in their parameters, it seems odd for him to imply (pp. 70-71) that our 1972 method, which does yield such convergence, is essentially equivalent to the 1969 method. It is not. Even his present (1976) method is not equivalent to ours since his requires evaluation of third-order mixed partial derivatives while ours required only second-order partials, which we actually evaluate numerically.
7. It is true, as Southwell asserts (p. 71), that we did not use the chain rule when applying our analytical method. We were, in fact, simply indicating that such methods *as described* (O'Neill, Sinclair, and Smith, 1969, and Powell, Macdonald, 1972), would not always converge without implicit mixed partials.
8. We note that the one additional iteration required for our method (Southwell's Table 2, 1976) is a small price to pay for not having to supply analytical derivatives.
9. We did not say (p. 72) that the problem $f(x) = a_1 + a_2x$ with uncertainties in both x and y 'remains linear in the parameters'. We *did* say that the least squares condition $\dfrac{\partial S}{\partial x_i} = 0$ can be solved exactly, thus explicitly eliminating the x's, when one has a model which is linear in the independent variable. Southwell correctly points out that there are other models for which $\dfrac{\partial S}{\partial x_i} = 0$ can be solved analytically.
10. We re-did Southwell's example 3 using the Powell–Macdonald method on an IBM 370 computer using partial double precision. Our results are:

$$\begin{aligned} a_1 &= 5{\cdot}914859 \\ a_2 &= -0{\cdot}6035114 \\ a_3 &= -0{\cdot}07996518 \\ a_4 &= 0{\cdot}02619385 \\ a_5 &= -8{\cdot}086482 \times 10^{-4} \\ a_6 &= -1{\cdot}685054 \times 10^{-4} \end{aligned}$$

with $2S = 0{\cdot}4503256$. We used the increment $\Delta = 10^{-4}$ to calculate our numerical derivatives, and convergence was achieved in seven iterations. It is not necessary, as Southwell implies (p. 72), to employ full extended precision on a CDC 6600-type machine to use our estimated partial derivatives program. The solution set obtained from our program is indeed somewhat different from Southwell's results, probably owing to the relatively large parameter standard deviations found in the quintic model. We believe that ours is as good a least squares solution as Southwell's.

Finally, we would like to point out that a technique's value depends

in large measure upon its 'track record', its availability, and its ease of use. The Powell–Macdonald method in FORTRAN code has been (and is) available to anyone who wishes to use it since 1972. In some respects, there are advantages to the Britt–Luecke method (1973), but we have been pleased with the wide variety of successful employment of our algorithm by a large number of users. Despite Southwell's claims of superiority in method to both Powell–Macdonald and Britt–Luecke, a potential user would be well advised to compare all available methods carefully before deciding which would be 'best' (in some sense) for his own problems.

Yours faithfully,

D. R. POWELL* and J. ROSS MACDONALD‡

*Texas Instruments
PO Box 6015
Dallas
Texas 75222
USA

‡Department of Physics and Astronomy
University of North Carolina
Chapel Hill
North Carolina
USA

19 October 1976

References

BRITT, H. I., and LUECKE, R. H. (1973). The estimation of parameters in nonlinear, implicit models, *Technometrics*, Vol. 15, pp. 233-247.

O'NEILL, M., SINCLAIR, I. G., and SMITH, F. J. (1969). Polynomial curve fitting when abscissas and ordinates are both subject to error, *The Computer Journal*, Vol. 12, pp. 52-56.

POWELL, D. R., and MACDONALD, J. R. (1972). A rapidly convergent iterative method for the solution of the generalized nonlinear least squares problem, *The Computer Journal*, Vol. 15, pp. 148-155; (1973), Vol. 16, p. 51.

SOUTHWELL, W. H. (1969). Fitting experimental data, *J. Comp. Phys.*, Vol. 4, pp. 465-474.

SOUTHWELL, W. H. (1976). Fitting data to nonlinear functions with uncertainties in all measurement variables, *The Computer Journal*, Vol. 19, pp. 69-73.

Dr. Southwell replies:

No attack was intended; however, I felt, for example, the number of iterations comparison (Table 4, Powell and Macdonald, 1972), which lists, 'Southwell . . . did not converge,' was rather harsh in view of the fact (Southwell, 1969) that the method converged as fast or faster than any other method listed.

My reply will be to each of the numbered comments above.

1. The basic methods of the 1969 paper and of the more recent paper (Southwell, 1976) are the same. The fundamental approach is to minimize the sum-of-the-squares function with respect to the **a** parameters, with x_i being regarded as a function $x_i(\mathbf{a})$ through the relationship $\partial S/\partial x_i = 0$. The 1969 paper suggests a Newton–Raphson approach (a well-known and 'highly successful' technique when starting points are near the solution) for the multi-parameter vector **a**. For any fixed **a**, a subproblem is solved for **x**, which involves finding a solution to a single variable transcendental equation. The 1969 paper suggests solving for the x_i using the iterative scheme $x_i^{new} = F(x_i)$ (Eq. 3, 1969). This is also a well-known technique and is easy to implement, since it requires no additional derivatives. The problem is that it doesn't always converge, depending on the nature of the function $F(x_i)$. Thus, in solving the transcendental equation for x_i, the 1976 paper states that:
   (*a*) Newton's method (also well-known) is preferred.
   (*b*) Convergence should be achieved rather than relying on a single step adjustment.

   The 1976 paper also points out that the basic method may be implemented using other nonlinear least squares techniques instead of the Newton–Raphson.
2. The error in the 'exact' parameter variance derivation in the 1969 paper was first revealed to the author by J. R. Macdonald, however, it was in an unsigned referee's report.
3. The 1976 paper was submitted in March 1974.
4. only a single step adjustment for the x_i. Also, no convergence tests are given for the x_i.
5. In view of the fact that there are several ways to implement the method, what I said (p. 71) was, 'It should be pointed out that with one difference Powell and Macdonald's method is equivalent to the Newton–Raphson, numerical derivative application of this method.' I don't think that is misleading. The point raised seems to be, which was the earlier method?
6. The 1976 paper gives examples of convergence of nonlinear models using the improved implementation (see item 1 in this reply) of the 1969 method. The belabored point appears to be whether or not the modifications (*a*) and (*b*) above constitute a change in method. Perhaps the answer depends on the reader. One looking for a package with which to fit his data might say that it is a change. A numerical analyst, however, would certainly not solve a transcendental equation using $x_i^{new} = F(x_i)$ without checking its convergence. If it failed, he would use Newton's or another method and still be able to achieve convergence using the 1969 paper.

   The numerical derivative approach (described on p. 70 in the 1976 paper) does not require evaluation of third order mixed partial derivatives.
7. Table 4 (p. 152, Powell and Macdonald, 1972) states that convergence was poor using exact analytical derivatives. Apparently, they only used partial derivatives as exact derivatives would have given different results.
8. The analytical derivatives on that example are trivial. But, generally, the choice depends on the application. For a dedicated program to fit a specific function, the extra effort may be well worth it in terms of accuracy and computer running time. For an easy-to-use general purpose program, the numerical derivatives would probably be preferred. Fortunately, the theory is well understood so that either approach may be used.
9. They indeed did say that (p. 148 and p. 149). This would be minor, except that the model $f = a_1 + a_2x$ is not a trivial case and has a long history in the literature. Because of this, it was used to illustrate the basic approach in the 1969 paper.
10. I took the above IBM 370 double precision results and used them as input parameters in CURFIT on a Hewlett-Packard 9830A calculator. (CURFIT is a commercially available (Cybernetic Service Co., 3508 Fifth Avenue, Pittsburgh, Pennsylvania 15213, USA) interactive curvefitting program that also handles data with uncertainties in both x and y.) The $2S$ value on input was 0·4503256739. After one iteration, the results were:

$$a_1 = 5{\cdot}91482$$
$$a_2 = -0{\cdot}603146$$
$$a_3 = -0{\cdot}0803391$$
$$a_4 = 0{\cdot}0263287$$
$$a_5 = -8{\cdot}28747 \times 10^{-4}$$
$$a_6 = -1{\cdot}67448 \times 10^{-4}$$

with $2S = 0{\cdot}4503256674$. These results are in close agreement with my example 3 (1976) and indicate a better least squares solution.

Finally, with regard to comparisons and superiority of methods. I now feel that for a general program, the Newton–Raphson optimisation is not the best approach. I am satisfied that the approach used by CURFIT is better. (CURFIT is basically a highly efficient least squares optimisation algorithm that requires only first derivatives, which are computed numerically. For problems with x weights, CURFIT uses the same least squares algorithm for the **a** parameters, but eliminates the x_i using the approach described in my work.) Thus, a potential user is well advised to consider all available methods. Or, if innovative, he may wish to consider other possible implementations (i.e., damped least squares or variable metric) of the approach outlined in my papers.

To the Editor
The Computer Journal

Sir

I should like to raise three points concerning the paper by Macnaghten and Hoare entitled 'Fast Fourier transform free from tears' on page 78 of your February 1977 edition.

1. The paper discusses many interesting...