Grand Challenges and Scientific Standards in Interval Analysis

Arnold Neumaier

Institut für Mathematik, Universität Wien
Strudlhofgasse 4, A-1090 Wien, Austria
email: Arnold.Neumaier@univie.ac.at
WWW: http://www.mat.univie.ac.at/~neum/

April 21, 2002

Abstract. This paper contains a list of 'grand challenge' problems in interval analysis, together with some remarks on improved interaction with mainstream mathematics and on raising scientific standards in interval analysis.

Introduction

Following a question by James Demmel on the Reliable Computing Mailing List, I presented on April 17, 2002 a list of 'grand challenge' problems in interval analysis, together with some remarks on improved interaction with mainstream mathematics. In the weeks before, there were discussions on the Mailing List about scientific standards in interval analysis which complement these remarks. The following is an edited version of my contributions to the Mailing List on these topics, revised in the light of the ensuing discussions. It is a heavily value-driven essay, giving my personal perspective on the issues involved.

I know that a careful posing of the grand challenges would require significantly more detail and references to previous work that would be the basis of corresponding research projects. Only for the first challenge, where an international project is already underway, such a description is available. Unfortunately, due to other commitments, I have no time in the near future to prepare such a detailed version; I give only references immediate at hand, and apologize for work not mentioned. However, upon request by Vladik Kreinovich and Slava Nesterov, I prepared the current paper for publication in the journal Reliable Computing.

Grand Challenges

Significant progress in mathematics has always been application driven, although not necessarily by applications outside mathematics. 'Grand challenges' describe problems that may change the way interval analysis is understood by their advocates and perceived from the outside, and whose purpose is to focus research on a few hot places. 'Grand challenges' are problems where one hopes for a breakthrough in the near future; problems where one can hope to get financial support if one has a good team, good ideas, and a good plan for succeeding. I see several such challenges:

1. Global optimization. Large-scale global optimization and constraint satisfaction. Although this is NP-hard, so is mixed integer programming; nevertheless large mixed integer programs are solved routinely (though not all of them).

Global optimization is in my opinion the only area of wide practical applicability where it is already forseeable that intervals are likely to have a major impact in the near future. Indeed, here interval arithmetic is indispensible to provide good global information about the nonlinearities in the problem formulation; all other ways of getting such information are unreliable or much too crude.

The difficulties to be overcome lie in a successful marriage of analytic (convexity-based) techniques, traditional interval methods, and constraint satisfaction techniques. This is the subject of a current international project called COCONUT [2], funded by the European Union.

2. Uncertainty in finite element calculations. Error bounds for realistic discrete finite element problems with realistic (dependent) uncertainties. Applications abound, and interval methods could become competitive, as replacement for expensive Monte-Carlo studies.

Interval analysis is quantitative sensitivity analysis in the sense of worst case perturbation analysis. (One can express it in terms of intervals or bounds, this is roughly equivalent. The methods and the difficulties are almost the same, though interval formulations usually lead to more accurate results.) In problems where safety is an issue, worst case results are needed, and these are currently obtained by expensive (and therefore usually quite incomplete) Monte Carlo calculations.

Traditional sensitivity analysis is valid only for sufficiently small uncertainties, and it is always unclear when the uncertainties are sufficiently small. For uncertainties above some problem-dependent scale, sensitivity analysis may be severely biased, since it does not account for the nonlinearities in a problem. And in many engineering applications, where the stability of a structure must be guaranteed over all conditions encountered within its life time, uncertainties (due to the environment and to wear of material) are usually substantial.

Problems to be overcome by current interval methods are the size of the systems (thousands of elements in realistic structures) and the nonlinear dependence of the coefficients on uncertain parameters. Some recent progress has been achieved by MUHANNA & MULLEN [16].

3. Sparse linear systems. Solution of large and sparse linear equations with interval coefficients.

Sparse matrices are ubiquitous in applications. Whenever one wants to calculate valid bounds on anything in a high number of variables, one ends up at some stage with a sparse linear system which has at least interval coefficients on the right hand side. (Or, in noninterval terms, one gets expressions involving the absolute value or norm of the inverse.)

There are techniques based on an observation of HAGER [6] (see also NEUMAIER [19, Algorithm 2.5.5]) that frequently produce good approximations or even exact values for the norm of the inverse. But as with sensitivity analysis or Monte Carlo methods, one never knows when this is the case, and occationally Hager-type estimators fail by orders of magnitude without warning. Moreover, the error bounds for solutions computed from the inverse norm bounds are frequently gross overestimates of the actual errors, especially if iterative refinement was used to compute the approximations.

Apart from a paper by RUMP [21], nothing has been done on the interval side.

4. Error bounds for large systems of ODEs.

The existing interval codes (LOHNER [12, 13], NEDIALKOV et al. [17], HOEFKENS et al. [11]) only handle very small systems; already modeling the sun and planets (without asteroids) over 10 years is probably too hard for present codes. But a long term goal in the applications should be to produce verified ephemerides (which are tables of positions of all celestial objects in the solar system throughout past and future centuries).

5. Stiff systems. Error bounds for stiff systems and singularly perturbed systems of ODEs.

Many dynamical systems in the applications are stiff, and current interval methods do not cope well with stiffness. But the books by HAIRER, NORSETT & WANNER [7, 8] contain enough analytical results that could be used to get started, by creating variants of the results with constructively verifiable assumptions.

6. Elliptic PDEs. Error bounds for elliptic partial differential equations on fine, irregular geometries. In contrast to Problem 2, which is concerned with discrete systems, here the discretization error is to be accounted for; but the problem is assumed to involve no uncertainties.

Most applications have their mesh adapted to the problem; current interval work is mostly on toy problems with very simple domains and meshes. In its general form, this might need a solution of Problem 3; but some applications are in the inverse monotone regime, where differential inequality techniques should be sufficient.

Tractability and success criteria. I have good reasons to think that each of these problems is of high significance, nontrivial, and tractable if it is made the focus of a group of good and dedicated mathematicians. Apart perhaps from Problem 3, it requires the dedicated collaboration of several people to get the work done at a sufficiently high quality level.

All problems require a solid understanding of the corresponding techniques from the analytical side, since they are not simply solvable by applying interval recipes to a problem, but only by translating the theorems and techniques of the application into sharper versions that allow one to add quantitative aspects to previously qualitative or asymptotic reasoning, and analytic rigor to previously approximate computations.

In each case, success would be measured by the availablility of an easy to use software system, perhaps based on INTLAB [22], that can be given to people working in applications to check it out. (As in the past, without that, interval techniques are doomed to a fringe existence.)

To qualify as a solution of the mathematical challenges, the software must provide realistic bounds for realistic-sized problems; it need not compete in speed with nonrigorous software.

In the past, potential users from real applications frequently tried methods suggested by interval researchers to their problems and found them wanting – not because they were too slow, but because they were limited to toy sizes, because they broke down at uncertainties well below realistic levels, or because they produced intervals that were unrealistically wide. *We* know that intervals must be handled with care; users often don't and go away disappointed, unless we give them tools that they can use with *their* experience, and which have *our* experience and expertise built in.

The breakthrough is when a method is shown to work well (even if slow) on significant problems of significant size; this attracts potential users to put in efforts to make it actually fully fitting their needs, and this includes speeding the methods up.

Tuning a method that satisfies the mathematical demands to achieve maximal speed is more computer science than mathematics. But to reach the level of usage in production codes which users in applications will be prepared to use routinely, the overhead of the reliable computation should probably be within a factor of $3-5$ of competing nonrigorous methods.

Interaction with mainstream mathematics

To work successfully on these grand challenges and to attract the resources needed for their successful solution is difficult for the present interval community because of a lack of deep interaction with mainstream mathematics, which must provide the missing tools.

There is a still prevailing sectarian isolation under the unspoken motto *'we have the solutions but no one is recognizing it'.* People or groups with such a motto usually underestimate the gulf between an idea and a solution, and are not prepared to take the trouble and invest the energy needed to turn good ideas into working solutions. In our case, exceptions to this attitude usually came from the outside, not from the interval camp.

Interval analysis (I do not like the presumptuous term 'interval mathematics' and never use it) is nothing else than mathematical analysis and linear algebra on the computer. To think of it as a separate discipline is a mistake, and progress in interval analysis is always based on better understanding of the relevant mathematical analysis in the various applications. What is special about interval analysis is only the perspective.

The foundations on intervals are already in a very good mathematical shape. The book by ALEFELD & HERZBERGER [1] and my book (NEUMAIER [18]) constitute surveys of interval work full of nontrivial mathematics; and much more has been done since these were published – there is no doubt that interval analysis can generate good mathematics.

But it can do so only if it maintains close contact to the rest of mathematics, from which it must derive its depth. To overcome the current isolation,

- *we* must become better mathematicians (i.e., better grounded in noninterval techniques of challenging mainstream mathematics);

- *we* have to do the innovative work ourselves and not leave it to others;
- *we* have to cross the communication barrier and learn to listen and speak in the language of those we dream of benefitting; and
- *we* have to understand the mathematics in the applications better than those working in the applications. This is not impossible, since the latter are frequently under time pressure and are therefore limited to the techniques they learnt during their College days. Moreover, the goal of verification gives us a different and better perspective for the tools in the field.

This is hard work, and it may deter many of us – it is easier to work on well-defined pet problems that change little over the years, have little competition and little relevance. But hard work is the hallmark of success, and if we are not willing to do it, we do not deserve success.

And als long as we are not willing to do it, the mathematical successes of interval methods (for a review see FROMMER [5]) will be – like in the past – achieved by people outside the interval camp (ECKMANN et al. [3], FEFFERMANN & SECO [4], HALES [9], HASS et al. [10], MEHLHORN & NÄHER [14], MISCHAIKOW & MROZEK [15], RAGE et al. [20]) who understand their field and learn just enough basic interval technique to reach their goal, and then go on to other interesting avenues in their fields and leave us alone, while taking the credit for their achievement with them. We may continue to claim their achievement as ours and bolster our self-esteem that intervals are after all important in other fields, but we deepen our isolation with the resulting self-pity of not being recognized.

Raising scientific standards

A long time ago, Jesus said to his would-be-great disciples, "I tell you that unless your righteousness surpasses that of the Pharisees and the teachers of the law, you will certainly not enter the kingdom of heaven." (Matth. 5:20)

Similarly, unless our mathematical ability and horizon surpasses that of the speakers of an International Congress of Mathematicians, we will certainly not enter the list of invited speakers. In the old days, the disciples of Jesus learnt their lesson; it is up to us to do the same.

To be recognized by the worldwide community of mathematicians as a vital part of mathematics, our standards must meet (and if we want our results to be honored surpass) the standards of publications set by high quality journals such as those published by SIAM. I want to strongly encourage good scientific practices for publishing research in interval analysis.

The purpose of a scientific publication is not to gather items for one's publication list or to show off with one's work, but to enable others to learn from one's insight and experience, and to allow them (in principle) to duplicate the work and improve upon it. This is possible only with good documentation of the relevant details. Moreover, scientific progress depends on acknowledging the work of others and building upon it, instead of brushing it aside which implicitly communicates the message that previous work by others was irrelevant.

- Never make exaggerated general claims based on few examples and no theory to support the statements made.

- Give complete documentation of the details that are critical for a reliable performance, so that it is possible to check corresponding claims, and avoid suspicions about the quality of the work.

- Do not use vague or ambiguous terminology that increases the likelihood of misunderstandings; check the best of the recent literature about the precise usage of concepts.

- Compare new techniques with the state of the art and not with the most naive methods. Compare them both on simple (often traditional) examples that give insight into the differences, and on difficult examples that show how far the new techniques may improve upon the state of the art.

- Get fully acquainted with related past work by others, and acknowledge the merits of other people and methods by giving quotations of all directly related work and ideas.

- The greater the claims and the better the work, the more important it is to ensure that the highest standards are respected.

On rigor in interval analysis

Publications on problems involving intervals cannot always be fully rigorous; for example, for problems that currently defy rigorous methods, it may be interesting to propose approximation algorithms. Indeed, less rigorous methods may contain the germs for more rigorous ones, and problems thast can be solved now by less rigorous methods pose challenges for more rigorous attacks.

In publications related to interval techniques, the degree of rigor used should be clearly visible (and mentioned in papers at prominent places – abstract, introduction, conclusion), and there should be sufficient documentation or references to allow readers to check that the claimed degree of rigor is actually achieved.

There are several increasingly demanding stages of rigor:

1. With uncontrolled approximations. For problems involving functions, this is the realm of traditional numerical analysis, where asymptotic estimates are all that is available to justify a method.

2. Without uncontrolled approximations but with unverified assumptions. This is the case, e.g., for methods that assume that the intervals are sufficiently narrow, without giving precise criteria for how narrow is enough.

3. Without uncontrolled approximations or unverified assumptions, but without error control. This would be rigorous in exact arithmetic, but is not in floating point arithmetic. An example is Gaussian elimination for exact data. For bounding interval linear systems, Rohn's sign accord algorithms belongs to this category, since so far no one has proposed a safe strategy for making the algorithm behave correctly in floating points.

4. Without uncontrolled approximations or unverified assumptions, and with rounding error control. These are the only algorithms that deserve the designation 'rigorous' or 'validated'.

5. Exact solutions. This usually requires symbolic computation and multi-precision rational arithmetic, and is outside the scope of interval analysis, though the use of interval analysis within exact algorithms is often useful (e.g., to verify signs of expressions).

Progress consists in bringing new problems into stage 1, or bringing older or larger problems than before one stage higher.

References

[1] G. Alefeld and J. Herzberger, Introduction to Interval Computations, Academic Press, New York 1983.

[2] C. Bliek, P. Spellucci, L.N. Vicente, A. Neumaier, L. Granvilliers, E. Monfroy, F. Benhamou, E. Huens, P. Van Hentenryck, D. Sam-Haroud and B. Faltings, Algorithms for Solving Nonlinear Constrained and Optimization Problems: The State of the Art. A progress report of the COCONUT project, 2001.
`http://www.mat.univie.ac.at/~neum/glopt/coconut/StArt.html`

[3] J.-P. Eckmann, H. Koch, and P. Wittwer, A computer-assisted proof of universality for area-preserving maps, Amer. Math. Soc. Memoir 289, AMS, Providence 1984.

[4] C.L.Fefferman and L.A.Seco, Interval Arithmetic in Quantum Mechanics, pp. 145–167 in: Applications of Interval Computations (R.B. Kearfott and V. Kreinovich, eds.), Kluwer, Dordrecht 1996.
`http://www-risklab.erin.utoronto.ca/seco/publ.htm`

[5] A. Frommer, Proving conjectures by use of interval arithmetic, pp. 1–13 in: Kulisch, Ulrich; Lohner, Rudolf; Facius Axel: Perspective on Enclosure Methods (U. Kulisch et al., eds.), Springer, Wien 2001.

[6] W.W. Hager, Condition estimates, SIAM J. Sci. Statist. Comput. 5 (1984), 311–316.

[7] E. Hairer, S.P. Norsett, and G. Wanner, Solving Ordinary Differential Equations, Vol. 1, Springer, Berlin 1987.

[8] E. Hairer and G. Wanner, Solving Ordinary Differential Equations, Vol. 2, Springer, Berlin 1991.

[9] T.C. Hales, The Kepler conjecture, Manuscript (1998). math.MG/9811071
http://www.math.lsa.umich.edu/~hales/countdown/

[10] J. Hass, M. Hutchings, and R. Schlafli, Double bubbles minimize, Ann. Math. (2) 515 (2000), 459–515.
http://math.ucdavis.edu/~hass/bubbles.html

[11] J. Hoefkens, M. Berz and K. Makino, Efficient High-Order Methods for ODEs and DAEs, pp. 341–351 in: Automatic Differentiation: From Simulation to Optimization (G. Corliss et al., eds.), Springer, New York 2001.

[12] R.J. Lohner, Einschliessung der Lösung gewöhnlicher Anfangs- und Randwertaufgaben und Anwendungen, Dissertation, Fakultät für Mathematik, Universität Karlsruhe, 1988.

[13] R.J. Lohner, AWA – software for the computation of guaranteed bounds for solutions of ordinary initial value problems.
ftp://ftp.iam.uni-karlsruhe.de/pub/awa/

[14] K. Mehlhorn and S. Näher, The LEDA Platform of Combinatorial and Geometric Computing, Cambridge University Press, 1999.

[15] K. Mischaikow and M. Mrozek, Chaos in the Lorenz equations: A computer assisted proof. Part II: Details, Math. Comput. 67 (1998), 1023–1046.

[16] R.L. Muhanna and R.L. Mullen, Uncertainty in mechanics problems – interval-based approach, J. Engin. Mech. 127 (2001), 557–566.

[17] N.S. Nedialkov, K.R. Jackson and G. Corliss, Validated Solutions of Initial Value Problems for Ordinarly Differential Equations, Appl. Math. Comput. 105 (1999), 21–68.

[18] A. Neumaier, Interval Methods for Systems of Equations, Cambridge Univ. Press, Cambridge 1990.

[19] A. Neumaier, Introduction to Numerical Analysis Cambridge Univ. Press, Cambridge 2001.

[20] T. Rage, A. Neumaier and C. Schlier, Rigorous verification of chaos in a molecular model, Phys. Rev. E. 50 (1994), 2682-2688.

[21] S.M. Rump, Verification methods for dense and sparse systems of equations, pp. 63-136 in: J. Herzberger (ed.), Topics in Validated Computations - Studies in Computational Mathematics, Elsevier, Amsterdam 1994.

[22] S.M. Rump, INTLAB – INTerval LABoratory, pp. 77–104 in: Developments in reliable computing (T. Csendes, ed.), Kluwer, Dordrecht 1999. `http://www.ti3.tu-harburg.de/rump/intlab/index.html`