Is It Possible to Have a Feasible Enclosure-Computing Method Which Is Independent of the Equivalent Form?*

Marcin Michalak
Institute of Computer Science, Silesian University of Technology, ul. Academicka 16, 44-100 Glivice, Poland
marcin.michalak@polsl.pl

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
vladik@utep.edu

August 22, 2012

Abstract

The problem of computing the range $\boldsymbol{y}$ of a given function $f(x_1, \ldots, x_n)$ over given intervals $\boldsymbol{x}_i$ – often called the main problem of interval computations – is, in general, NP-hard. This means that unless P = NP, it is not possible to have a feasible (= polynomial time) algorithm that always computes the desired range. Instead, interval computations algorithms compute an enclosure $\boldsymbol{Y} \supseteq \boldsymbol{y}$ for the desired range. For all known feasible enclosure-computing methods – starting with straightforward interval computations – there exist two expressions $f(x_1, \ldots, x_n)$ and $g(x_1, \ldots, x_n)$ for computing the same function that lead to different enclosures. We prove that, unless P = NP, this is inevitable: it is not possible to have a feasible enclosure-computing method which is independent of the equivalent form.

Keywords: interval computations, enclosure, equivalent form, NP-hard
AMS subject classifications: 65G20, 65G40, 03D15, 68Q17

1 Formulation of the Problem

One of the main problems of interval computations. One of the main problems of interval computations has the following form:

*Submitted: June 28, 2011; Revised: August 19, 2012; Accepted: ???.

- We are given an algorithm $f(x_1, \ldots, x_n)$ for computing a function of n real variables, and n intervals $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.
- We need to compute the range

$$\boldsymbol{y} = [\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) : x_1 \in \boldsymbol{x}_1, \ldots, x_n \in \boldsymbol{x}_n\}$$

  of the function $f(x_1, \ldots, x_n)$ under given intervals.

This problem is NP-hard. It is known (see, e.g., [3]) that the above problem is, in general, NP-hard.

Comment. It is widely believed that P≠NP. In this case, NP-hardness means that it is not possible to have a feasible (= polynomial time) algorithm that always computes the desired range.

Feasible methods for computing enclosures. Since we cannot always efficiently compute the exact range $\mathbf{y}$, and we want to guarantee that the value $f(x_1, \ldots, x_n)$ is contained in the estimated range, we need to find an *enclosure* $\boldsymbol{Y} \supseteq \boldsymbol{y}$. There exist many feasible techniques for computing enclosures: straightforward interval computations, mean value form, methods combined with bisection, etc.; see, e.g., [4].

Existing feasible methods for computing enclosure are not independent on the equivalent form. In straightforward interval computations, we

- represent the algorithm $f(x_1, \ldots, x_n)$ as a sequence of elementary arithmetic operations such as $+$, $-$, $\cdot$, $/$, min, max, etc. and then
- replace each elementary operation with the corresponding operation of interval arithmetic.

In this method, in general, the resulting enclosure may be different for different algorithms that compute the same function. For example, the algorithms $f(x_1) = x_1 - x_1$ and $g(x_1) = 0$ compute the same function 0 on the interval $[0, 1]$, but:

- for $f(x_1) = x_1 - x_1$ straightforward interval computations leads to an enclosure

$$[0, 1] - [0, 1] = [-1, 1],$$

  while
- for $g(x_1) = 0$, we get the enclosure $\mathbf{Y} = [0, 0] \neq [-1, 1]$.

Similarly, all other feasible methods for computing enclosure – at least those which are known to the authors – are not independent on the equivalent form: for each of these methods, there exist two algorithms that compute the same function but lead to different enclosures.

Natural question. The above fact leads to the following natural question: Is it possible to have a feasible enclosure-computing method which is independent of the equivalent form?

Comment. For straightforward interval computations $S(f, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, we have an even stronger result about algorithms that compute the same function but lead to different enclosures (see, e.g., [1, 2]). This result is as follows: for each algorithm $f(x_1, \ldots, x_n)$, for each set of intervals $\boldsymbol{x}_i$, and for each interval $\boldsymbol{Y} \supseteq \boldsymbol{y}$ containing the actual range $\boldsymbol{y}$, there exists an algorithm $g(x_1, \ldots, x_n)$ with the following two properties:

- for values $x_i \in \boldsymbol{x}_i$, the algorithm $g(x_1, \ldots, x_n)$ computes the same function as $f(x_1, \ldots, x_n)$, i.e.,

$$\forall x_1 \in \boldsymbol{x}_1 \ \ldots \ \forall x_n \in \boldsymbol{x}_n \, (f(x_1, \ldots, x_n) = g(x_1, \ldots, x_n));$$

- straightforward interval computations, when applied to the algorithm $g(x_1, \ldots, x_n)$, return the interval $\boldsymbol{Y}$:

$$S(g, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \boldsymbol{Y}.$$

2 Main Result

Definition. *By a* feasible enclosure-computing method*, we mean a feasible algorithm A that, given an algorithm $f(x_1, \ldots, x_n)$ and n intervals $\boldsymbol{x}_1$, ..., $\boldsymbol{x}_n$, computes an interval $A(f, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ that contains the range*

$$\boldsymbol{y} = \{f(x_1, \ldots, x_n) : x_1 \in \boldsymbol{x}_1, \ldots, x_n \in \boldsymbol{x}_n\}.$$

Proposition. *If $P \neq NP$, then for every feasible enclosure-computing method A, there exist two algorithms $g(x_1, \ldots, x_n)$ and $h(x_1, \ldots, x_n)$ and n intervals $\boldsymbol{x}_1$, ..., $\boldsymbol{x}_n$ for which*

- *$g(x_1, \ldots, x_n) = h(x_1, \ldots, x_n)$ for all $x_1 \in \boldsymbol{x}_1$, ..., $x_n \in \boldsymbol{x}_n$, but*
- *$A(g, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \neq A(h, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$.*

Comment. Thus, it is not possible to have a feasible enclosure-computing method which is independent of the equivalent form.

Proof.

1°. In our proof, we will use the following result from [3]: that even if we fix the intervals $\boldsymbol{x}_1 = \ldots = \boldsymbol{x}_n = [0, 1]$, then, for a certain class of polynomial functions $f(x_1, \ldots, x_n)$, we would always have $\overline{y} \geq 1$ or $\overline{y} \leq 0$, and the problem of checking whether $\overline{y} \geq 1$ or $\overline{y} \leq 0$ is NP-hard. (This result is proved at the end of the proof of Theorem 3.1 from [3].) In our proof, we will consider these same intervals $\boldsymbol{x}_1 = \ldots = \boldsymbol{x}_n = [0, 1]$.

2°. We will prove this result by contradiction. Let us assume that there exists a feasible enclosure-computing method A for which, if two algorithms $g(x_1, \ldots, x_n)$ and $h(x_1, \ldots, x_n)$ compute the same function on a box $[0, 1] \times \ldots \times [0, 1]$, then $A(g, [0, 1], \ldots, [0, 1]) = A(h, [0, 1], \ldots, [0, 1])$.

Let us show that in this case, we will be able to feasibly check, given an algorithm $f(x_1, \ldots, x_n)$ and n intervals $\boldsymbol{x}_1 = \ldots = \boldsymbol{x}_n = [0,1]$, whether the upper endpoint $\overline{y}$ of the range $[\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) : x_1 \in \boldsymbol{x}_1, \ldots, x_n \in \boldsymbol{x}_n\}$ is ≥ 1 or is ≤ 0.

3°. Let us first apply the method A to the algorithm $f_0(x_1, \ldots, x_n)$ that always returns 0. Let us denote the result $A(f_0, [0,1], \ldots, [0,1])$ of applying A to this algorithm f_0 by $[z_-, z_+]$. Since the method A is enclosure-computing, the range $[z_-, z_+]$ is an enclosure for the actual range $[0,0]$ of the function f_0. Thus, $z_- \leq 0 \leq z_+$.

4°. Our main idea is that, based on the algorithm $f(x_1, \ldots, x_n)$, we can compute a new algorithm $g(x_1, \ldots, x_n) = \max((z_+ + 1) \cdot f(x_1, \ldots, x_n), 0)$.

Let us consider two possible cases: $\overline{y} \geq 1$ and $\overline{y} \leq 0$.

4.1°. If $\underline{y} \leq 0$, this means that $f(x_1, \ldots, x_n) \leq 0$ for all $x_i \in [0,1]$. Thus, for all values $(x_1, \ldots, x_n)$ from the corresponding box $[0,1] \times \ldots \times [0,1]$, we have $g(x_1, \ldots, x_n) = \max((z_+ + 1) \cdot f(x_1, \ldots, x_n), 0) = 0$. Hence, in this case, $g(x_1, \ldots, x_n)$ computes the same function as $f_0(x_1, \ldots, x_n)$ for all the values from the box. So, due to our assumption, $A(g, [0,1], \ldots, [0,1]) = A(f_0, [0,1], \ldots, [0,1]) = [z_-, z_+]$, and thus,

$$\overline{A}(g, [0,1], \ldots, [0,1]) = z_+.$$

4.2°. On the other hand, if $\overline{y} \geq 1$, this means that there exists values $x_i \in [0,1]$ for which $f(x_1, \ldots, x_n) \geq 1$. For these values x_i, we have

$$g(x_1, \ldots, x_n) = \max((z_+ + 1) \cdot f(x_1, \ldots, x_n), 0) \geq z_+ + 1.$$

Thus, the upper endpoint of the range of the function g on the box $[0,1] \times \ldots \times [0,1]$ is also $\geq z_+ + 1$. Since the interval $A(g, [0,1], \ldots, [0,1])$ is an enclosure for this range, its upper endpoint greater than or equal to $z_+ + 1$:

$$\overline{A}(g, [0,1], \ldots, [0,1]) \geq z_+ + 1.$$

4.3°. Summarizing:

- if $\overline{y} \leq 0$, then $\overline{A}(g, [0,1], \ldots, [0,1]) = z_+$;
- if $\overline{y} \geq 1$, then $\overline{A}(g, [0,1], \ldots, [0,1]) \geq z_+ + 1$.

5°. According to Part 4 of this proof, by applying the method A to the function $g(x_1, \ldots, x_n)$ and the intervals $\boldsymbol{x}_1 = \ldots = \boldsymbol{x}_n = [0,1]$, and by checking whether $\overline{A}(g, [0,1], \ldots, [0,1]) = z_+$ or $\overline{A}(g, [0,1], \ldots, [0,1]) \geq z_+ + 1$, we will be able to detect, in feasible time, whether $\overline{y} \leq 0$ or $\overline{y} \geq 1$.

However, we know that the problem of detecting whether $\overline{y} \geq 0$ or $\overline{y} \geq 1$ is NP-hard. Thus, the fact that we can solve this problem in polynomial time means that P = NP – and we assumed that P≠NP.

This contradiction shows that our assumption is wrong, i.e., that for every feasible enclosure-computing method A, there exist two algorithms $g(x_1, \ldots, x_n)$ and $h(x_1, \ldots, x_n)$ that compute the same function for given n intervals $\boldsymbol{x}_1 = \ldots = \boldsymbol{x}_n = [0,1]$ but for which $A(g, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \neq A(h, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$.

The proposition is proven.

Acknowledgments. The first author was supported by the European Union from the European Social Fund (grant agreement number: UDA-POKL.04.01.01-106/09). This work was also partly supported by the US National Science Foundation grants HRD-0734825 and DUE-0926721, and by Grant 1 T36 GM078000-01 from the US National Institutes of Health.

The authors are thankful to all the participants of the 13th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing RSFDGrC'2011 (Moscow, Russia, June 25–30, 2011) and to the anonymous referees for valuable discussions.

References

[1] Koshelev, M.: *Every Superinterval of the Function Range Can Be An Interval-Computations Enclosure*, The Chinese University of Hong Kong, Department of Mechanical & Automation Engineering, Technical Report CUHK-MAE-99-003, January 1999.

[2] Koshelev, M.: *Every superinterval of the function range can be an interval-computations enclosure*, Reliable Computing 6, pp. 219–223, 2000.

[3] Kreinovich, V., Lakeyev, A., Rohn, J., Kahl, P.: *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1997.

[4] Moore, R. E., Kearfott, R. B., Cloud, M. J.: *Introduction to Interval Analysis*, SIAM Press, Philadelphia, Pennsylvania, 2009.