Numeric Reasoning with Relative Orders of Magnitude

Philippe Dague

L.I.P.N. Université Paris Nord
Avenue Jean-Baptiste Clément
93430 Villetaneuse - France
e-mail: dague@lipn.univ-paris13.fr

Abstract

In [Dague, 1993], a formal system ROM(K) involving four relations has been defined to reason with relative orders of magnitude. In this paper, problems of introducing quantitative information and of ensuring validity of the results in $\mathbb{R}$ are tackled.
Correspondent overlapping relations are defined in $\mathbb{R}$ and all rules of ROM(K) are transposed to $\mathbb{R}$. Unlike other proposed systems, the obtained system ROM($\mathbb{R}$) ensures a sound calculus in $\mathbb{R}$, while keeping the ability to provide commonsense explanations of the results.
If needed, these results can be refined by using additional and complementary techniques: k-bound-consistency, which generalizes interval propagation; symbolic computation, which considerably improves the results by delaying numeric evaluation; symbolic algebra calculus of the roots of partial derivatives, which allows the exact extrema to be obtained; transformation of rational functions, when possible, so that each variable occurs only once, which allows interval propagation to give the exact results.
ROM($\mathbb{R}$), possibly supplemented by these various techniques, constitutes a rich, powerful and flexible tool for performing mixed qualitative and numeric reasoning, essential for engineering tasks.

Introduction

The first attempt to formalize relative order of magnitude reasoning, with binary relations invariant by homothety, appeared with the formal system FOG [Raiman, 1986] (see also [Raiman, 1991] for a more general set-based framework), based on 3 basic relations and described by 32 inference rules, which has been used successfully in the DEDALE system of analog circuit diagnosis [Dague et al., 1987]. Nevertheless, FOG has several limitations which prevent it from being really used in engineering. A first difficulty arises when wanting to express a gradual change from one order of magnitude to another: only a steep change is possible, due to the non overlapping of the orders of magnitude. This can be solved, as described in [Dague, 1993], by introducing a fourth relation "to be distant from" which allows overlapping relations to be defined and used. This has given a formal system ROM(K) with 15 axioms, consistency of which was proved by finding models in non standard analysis.

But two crucial problems remain: the difficulty to incorporate quantitative information when available (in DEDALE this lack of a numeric-symbolic interface meant writing Ohm's law in an ad hoc form) and the difficulty to control the inference process, in order to obtain valid results in the real world. These problems were pointed out in [Mavrovouniotis and Stephanopoulos, 1987] but the proposed system O(M) does not really solve them. In particular, use of heuristic interpretation semantics just ensures the validity of the inference in $\mathbb{R}$ for one step (application of one rule) but not for several steps (when chaining rules). This paper focuses on solving these two problems by concentrating on how to transpose the formal system ROM(K) to $\mathbb{R}$ with a guarantee of soundness and how to use additional numeric and symbolic algebra techniques to refine the results if needed, in order to build a powerful tool for both qualitative/symbolic and quantitative/numeric calculus for engineering purposes.

The present paper is organized as follows. Section 2 shows through an example how ROM(K), as FOG or O(M), may lead to results that are not valid in $\mathbb{R}$. In section 3 a translation in $\mathbb{R}$ of axioms and properties of ROM(K) is given, which ensures soundness of inference in $\mathbb{R}$. In section 4 the example is revisited with this new formulation; this time correct results are obtained. Nevertheless they may be far from the optimal ones and too inaccurate for certain purposes. In section 5 numeric and symbolic algebra techniques are proposed to refine these results: applications of consistency techniques for numeric CSPs; use of computer algebra to push symbolic computation as far as possible and delay numeric evaluation, which considerably improves the results; symbolic calculus of derivatives and of their roots by using computer algebra alone in order to compute extrema and obtain optimal results; formal transformation of rational functions by changing variables, which allows the exact results to be obtained, in particular cases, by a simple numeric evaluation and opens up future ways of research.

Example: a Heat Exchanger

Let us remember that the formal system ROM(K) (see [Dague, 1993] for a complete description) involves four binary relations $\approx$, $\sim$, $\ll$ and $\not\sim$, intuitive meanings of which are "close to", "comparable to", "negligible w.r.t." and "distant from" respectively. The 15 axioms are as follows:

(A1) $A \approx A$
(A2) $A \approx B \mapsto B \approx A$
(A3) $A \approx B,\ B \approx C \mapsto A \approx C$
(A4) $A \sim B \mapsto B \sim A$
(A5) $A \sim B,\ B \sim C \mapsto A \sim C$
(A6) $A \approx B \mapsto A \sim B$
(A7) $A \approx B \mapsto C.A \approx C.B$
(A8) $A \sim B \mapsto C.A \sim C.B$
(A9) $A \sim 1 \mapsto [A] = +$
(A10) $A \ll B \longleftrightarrow B \approx (B+A)$
(A11) $A \ll B,\ B \sim C \mapsto A \ll C$
(A12) $A \approx B,\ [C] = [A] \mapsto (A+C) \approx (B+C)$
(A13) $A \sim B,\ [C] = [A] \mapsto (A+C) \sim (B+C)$
(A14) $A \sim (A+A)$
(A15) $A \not\asymp B \longleftrightarrow (A-B) \sim A$ or $(B-A) \sim B$

45 properties of ROM(K) have been deduced from these axioms and 7 basic overlapping relations between two positive quantities $A < B$ have been defined:
$A \approx B$, $\neg(A \not\asymp B)$, $\neg(A \approx B) \wedge A \sim B$, $A \not\asymp B \wedge A \sim B$, $A \not\asymp B \wedge \neg(A \ll B)$, $\neg(A \sim B)$, $A \ll B$.
Taking into account signs and identity, these relations give 15 primitive overlapping relations. Adding the 47 compound relations obtained by disjunction of successive primitive relations gives a total of 62 legitimate relations.

Let us try to apply ROM(K) to a simple example of a counter-current heat exchanger as described in [Mavrovouniotis and Stephanopoulos, 1988]. Let FH and KH be the molar-flowrate and the molar-heat of the hot stream, FC and KC the molar-flowrate and the molar-heat of the cold stream. Four temperature differences are defined: DTH is the temperature drop of the hot stream, DTC is the temperature rise of the cold stream, DT1 is the driving force at the left end of the device, and DT2 is the driving force at the right end of the device. The two following equations hold:
(e1) $DTH - DT1 - DTC + DT2 = 0$,
(e2) $DTH \times KH \times FH = DTC \times KC \times FC$.
The first one is a consequence of the definition of the temperature differences, and the second one is the energy balance of the device. Let us take the following assumptions expressed as order of magnitude relations:
(i) $DT2 \sim DT1$, (ii) $DT1 \ll DTH$, (iii) $KH \approx KC$.

The problem is now to deduce from the 2 equations and these 3 order of magnitude relations the 5 missing order of magnitude relations between quantities having the same dimension (4 for temperature differences and 1 for molar-flowrates). Let us take the axioms (Ai) above and the properties (Pi), viewed as production rules of a symbolic deduction system ROM, as stated in [Dague, 1993].

Consider first the relation between DT2 and DTH. Thanks to (P4) $A \ll B,\ C \sim A \mapsto C \ll B$, ROM infers from (i) and (ii) that (1) $DT2 \ll DTH$.

Consider the relation between DTH and DTC. (P5) $A \ll B \mapsto -A \ll B$ and (P6) $A \ll C,\ B \ll C \mapsto (A+B) \ll C$ applied to (ii) and (1) imply $-DT1 + DT2 \ll DTH$. From this it can be deduced, using (A10), that $DTH \approx DTH - DT1 + DT2$, i.e. using (e1) that (2) $DTH \approx DTC$.

Consider the relation between DT1 and DTC. From (ii) and (2) it can be deduced, using (A6) and (A11), that (3) $DT1 \ll DTC$.

Consider the relation between DT2 and DTC. It results from (i) and (3), by using (P4), that (4) $DT2 \ll DTC$.
Another deduction path can be found to obtain the same result. In fact, from (A10) $A \approx B \mapsto (B-A) \ll A$ and $A \approx C \mapsto (C-A) \ll A$, using (P5) and (P6), (P) $A \approx B,\ A \approx C \mapsto (C-B) \ll A$ can be derived. As, from (3) and (A10), it results that $DTC \approx DTC + DT1$, it can be deduced from this and (2), using (P), that $-DTH + DT1 + DTC \ll DTC$, i.e. using (e1) that (4) $DT2 \ll DTC$.

Consider finally the relation between FH and FC. (A7), applied to (iii) and (2), gives $DTH \times KH \approx DTH \times KC$ and $DTH \times KC \approx DTC \times KC$. Applying (A3) then gives $DTH \times KH \approx DTC \times KC$. Applying (A7) again and using (e2) gives (5) $FH \approx FC$.

The five results (1 to 5) have thus been obtained by ROM (identical to those produced by O(M) because $\not\asymp$ is not used here):
(1) $DT2 \ll DTH$ (2) $DTH \approx DTC$ (3) $DT1 \ll DTC$ (4) $DT2 \ll DTC$ (5) $FH \approx FC$.

We have now to evaluate them in the real world. For this, it is necessary to fix a numeric scale for the order of magnitude relations. Choose for example $\ll$ represented by at most 10%, $\approx$ by at most (for the relative difference) 10 % and $\sim$ by at most (for the relative difference) 80%. Assumptions thus mean that:
(i′) $0.2 \leq DT2/DT1 \leq 5$, (ii′) $DT1/DTH \leq 0.1$, (iii′) $0.9 \leq KH/KC \leq 1.112$.
It is not very difficult in this example to compute the correct results by hand. It is found (see also subsections 5.3 and 5.4) that:
(1′) $DT2/DTH \leq 0.5$ (2′) $0.714 \leq DTH/DTC \leq 1.087$ (3′) $DT1/DTC \leq 0.109$ (4′) $DT2/DTC \leq 0.358$ (5′) $0.828 \leq FH/FC \leq 1.556$.

This shows that only the formal result $DT1 \ll DTC$ is satisfied in practice. For the 4 others, although the inference paths remain short in this example, there is already a non trivial shift, which makes them unacceptable. This is the case for the two $\approx$ relations: DTH may in fact differ from DTC by nearly 30%, and FH may differ from FC by 35%. And the same happens for the two $\ll$ relations: DT2 can reach 35% of DTC and, worse, 50% of DTH. This is not really surprising because we know that there is no model of ROM(K) in $\mathbb{R}$. Here, it is essentially the rule (P4) that causes discrepancy between qualitative and numeric results. Rules such as (P4), or (A11) from which it comes, and also (A3) are obviously being infringed. What this does demonstrate is the insufficiency of ROM for general engineering tasks and the need for a sound relative order of magnitude calculus in $\mathbb{R}$.

Transposing the Formal System to $\mathbb{R}$: ROM($\mathbb{R}$)

In fact, all the theoretical framework developed in [Dague, 1993] is a source of inspiration for this task. Since the rules of ROM capture pertinent qualitative information and may help guide intuition, they will serve as guidelines for inferences in $\mathbb{R}$ ([Dubois and Prade, 1989] addresses the same type of objectives by using fuzzy relations). Let us introduce the natural relations in $\mathbb{R}$, parameterized by a positive real k, "close to the order k":

$A \overset{k}{\sim} B \quad \longleftrightarrow \quad |A-B| \leq k \times \mathrm{Max}(|A|,|B|)$,

i.e. for $k<1$, (I) $1-k \leq A/B \leq 1/1-k$ or $A=B=0$, "distant at the order k":

$A \overset{k}{\not\sim} B \quad \longleftrightarrow \quad |A-B| \geq k \times \mathrm{Max}(|A|,|B|)$,

i.e. for $k<1$, (II) $A/B \leq 1-k$ or $A/B \geq 1/1-k$ or $B=0$, and "negligible at the order k":

$A \overset{k}{\ll} B \quad \longleftrightarrow \quad |A| \leq k \times |B|$,

i.e. (III) $-k \leq A/B \leq k$ or $A=B=0$.

The first one will be used to model both $\approx$ and $\sim$, the second one to model $\not\approx$, and the third one to model $\ll$, by associating a particular order to each relation. When trying to transpose the axioms (Ai) by using these new relations, three cases occur.

Axioms of reflexivity (A1), symmetry (A2,A4), invariance by homothety (A7,A8), and invariance by adding a quantity of the same sign (A12,A13, assuming (A9)) are obviously satisfied by $\overset{k}{\sim}$ for any positive k.

A second group of axioms imposes constraints between the respective orders attached to each of the 4 relations. Coupling of $\sim$ with signs (A9) is true for any order k attached to $\sim$ that verifies $k<1$. The fact that $\sim$ is coarser than $\approx$ (A6) forces the order for $\approx$ to be not greater than the order for $\sim$. Axiom (A14) is true for $\overset{k}{\sim}$ if $k \geq 1/2$. The left to right implication of (A10) has the exact equivalent: $A \overset{k}{\ll} B \mapsto B \overset{k}{\sim} (B+A)$ for $k<1$. We can thus take the same order k_1 for $\ll$ and $\approx$. In the same way, the definition of $\not\approx$ in terms of $\sim$ (A15) has its equivalent: $A \overset{k}{\not\sim} B \longleftrightarrow (A-B) \overset{1-k}{\sim} A$ or $(B-A) \overset{1-k}{\sim} B$ provided $k \leq 1/2$, i.e. $1-k \geq 1/2$. If we call k_2 the order for $\not\approx$, we can thus take $1-k_2$ as the order for $\sim$. All the above thus leads to the following correspondences:

$A \approx B \longleftrightarrow A \overset{k_1}{\sim} B \qquad A \sim B \longleftrightarrow A \overset{1-k_2}{\sim} B$

$A \ll B \longleftrightarrow A \overset{k_1}{\ll} B \qquad A \not\approx B \longleftrightarrow A \overset{k_2}{\not\sim} B$

with $0<k_1 \leq k_2 \leq 1/2 \leq 1-k_2<1$. Note that, as the formal system ROM(K) depends on two relations, its analog in $\mathbb{R}$ has two degrees of freedom represented by the orders k_1 and k_2.

Remaining axioms are those which are not true in $\mathbb{R}$. For these, the loss of precision on the orders is computed exactly in conclusion. The right to left implication of (A10) gives:

$B \overset{k}{\sim} (B+A) \quad \mapsto \quad A \overset{k/1-k}{\ll} B$ ($k/(1-k)<1$ when $k<1/2$).

Transitivity axioms (A3) and (A5) each give:

$A \overset{k}{\sim} B,\ B \overset{k'}{\sim} C \ \mapsto\ A \overset{k+k'-kk'}{\sim} C$ ($k+k'-kk'<1$ when $k<1$ and $k'<1$).

Finally the coupling between $\approx$ (through $\ll$) and $\sim$ (A11) gives:

$A \overset{k}{\ll} B,\ B \overset{1-k'}{\sim} C \ \mapsto\ A \overset{k/k'}{\ll} C$ ($k/k'<1$ when $k<k'$).

Like the axioms (Ai), all properties (Pi) deduced in [Dague, 1993] are demonstrated in the same way by computing the best orders in conclusion, when they are not directly satisfied, and constitute all the inference rules of ROM($\mathbb{R}$). For reasons of lack of space, here are the most significant or useful of the 45 properties for our purpose.

(P4), as (A11), gives $A \overset{k}{\ll} B,\ A \overset{1-k'}{\sim} C \ \mapsto\ C \overset{k/k'}{\ll} B$. (P6) gives $A \overset{k}{\ll} C,\ B \overset{k'}{\ll} C \ \mapsto\ (A+B) \overset{k+k'}{\ll} C$ (which can be improved if $[A]=-[B]$ by taking $\max(k,k')$ as order in conclusion) and (P) gives $A \overset{k}{\sim} B,\ A \overset{k'}{\sim} C \ \mapsto\ C-B \overset{k''}{\ll} A$ with $k''=(k+k'-kk')/(1-\max(k,k'))$ (which can be improved if $[C-B]=[A]$ by taking $k''=(k+k'-kk')/(1-k')$).

Transitivity of $\ll$ (P11) obviously improves the degree $A \overset{k}{\ll} B,\ B \overset{k'}{\ll} C \ \mapsto\ A \overset{k \times k'}{\ll} C$.

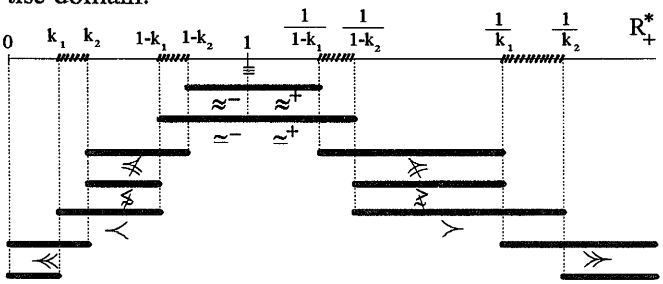
The incompatibility of $\ll$ and $\sim$ (P14) $A \overset{1-k_2}{\sim} B,\ A \overset{k_1}{\ll} B \ \mapsto\ A=B=0$ and of $\approx$ and $\not\approx$ (P34) $A \overset{k_1}{\sim} B,\ A \overset{k_2}{\not\sim} B \ \mapsto\ A=B=0$ are ensured provided $k_1 < k_2$. These two properties will be used to check the consistency of the set of relations describing, for example, an actual behavior of a physical system, or on the contrary to detect inconsistencies coming, from example, from discrepancies between modeled and actual behaviors of a system for tasks such as diagnosis.

Relations between $\approx$ and $\not\approx$ (P37, P38) give $A \overset{k_2}{\not\sim} B,\ C \overset{k_1}{\sim} A \ \mapsto\ C \overset{(k_2-k_1)/(1-k_1)}{\not\sim} B$ and $A \overset{k_1}{\sim} B,\ C \overset{k_2}{\not\sim} A \ \mapsto\ (C-A) \overset{k_1/k_2}{\sim} (C-B)$.

Finally, the completeness of the description is obtained: $A \overset{1-k_2}{\sim} B$ or $A \overset{k_2}{\not\sim} B$ provided that $k_2 \leq 1/2$.

Moreover, it can be proved that adding the assumption $A \overset{1-k_2}{\sim} B$ or $B \overset{1-k_2}{\sim} A$ or $A \overset{k_1}{\ll} B$ or $B \overset{k_1}{\ll} A$ would be equivalent to adding $A \overset{k_1}{\sim} B$ or $A \overset{k_2}{\not\sim} B$ and also equivalent to $k_2 \leq k_1$. This would imply $k_1=k_2=e$ as in FOG or in the strict interpretation of O(M) [Mavrovouniotis and Stephanopoulos, 1987], i.e. only one degree of freedom instead or two. In the same way that formal models of ROM(K) could not be reduced to FOG or O(M), the same is obtained for ROM($\mathbb{R}$) by choosing $k_1<k_2$, i.e. $0<k_1<k_2 \leq 1/2$. Note that, in relation to non standard models of ROM(K), one degree of freedom corresponds to what is chosen as the analog in $\mathbb{R}$ of the infinitesimals, and the other to the choice of the analog of the parameter ε [Dague, 1993] of the model, where ε corresponds to k_1/k_2.

The above gives the exact counterpart in $\mathbb{R}_+$ of the 15 primitive relations of [Dague, 1993] (in fact inference rules analog to the previous ones can be defined for these relations, often with some better orders in conclusion by taking into account the signs of the quantities, e.g. (P6) and (P) above), for describing the order of magnitude of A/B w.r.t. 1. These relations correspond to real intervals which overlap (in contrast with the strict interpretation of O(M)) and are built from the landmarks k_1, k_2, $1-k_2$, $1-k_1$, 1, $1/(1-k_1)$, $1/(1-k_2)$, $1/k_2$, $1/k_1$ of $\mathbb{R}_+$ (see Fig. 1). Note that, as in the formal model, intervals (k_1, k_2), $(1-k_2, 1-k_1)$ and their inverse are not acceptable relations. The heuristic interpretation of O(M) [Mavrovouniotis and Stephanopoulos, 1987], which only allows the correct inference to be made for one rule, corresponds to the particular case where $k_1 = k_2/(1+k_2)$. But here k_1 and k_2 are chosen independently, according to the expertise domain.



$\ll$ Fig. 1 $\gg$

The orders appearing in each rule of ROM($\mathbb{R}$) have to be considered as variables, with order in conclusion symbolically expressed in terms of orders in premises. Inference by such a rule is made simply by matching relations patterns in premises with existing relations, i.e. by instantiating the orders to real numbers, and computing the orders in conclusion to deduce a new relation. A sound calculus in $\mathbb{R}$ is thus ensured whatever the path of rules used, i.e. any conclusion that can be deduced by application of rules of ROM($\mathbb{R}$) from given correct premises is correct when interpreted in $\mathbb{R}$. This may be entirely hidden from the user, with both data and results being translated via k_1 and k_2 in symbolic order of magnitude relations as in FOG, but this time correctly. But the user may also incorporate numeric information as required by using k orders directly or even introducing numeric values for quantities. In this last case, binary relations between two given numeric quantities are automatically inferred from the equivalent formulation of these relations in terms of intervals (I,II,III).

Control of the inference in order to avoid combinatory explosion remains a problem for large applications. A pragmatic solution is to call on the expert who will possibly prevent a rule from being applied if he considers in the light of its underlying qualitative meaning that the conclusions are not relevant.

The Example Revisited

Let us take the example of the heat exchanger again and follow the same reasoning paths, but this time using the inference rules of ROM($\mathbb{R}$). Take, as in the numeric application, $k_1 = 0.1$ and $k_2 = 0.2$. As (1) is inferred by using (P4) we obtain:

(1a) DT2 $\overset{k3}{\ll}$ DTH with $k3 = k_1/k_2 = 0.5$.

DT2 and $-$DT1 having opposite signs, application of (P6) gives $-$DT1 + DT2 $\overset{k4}{\ll}$ DTH

with $k4 = \max(k_1, k3) = 0.5$. Using left to right implication of (A10) and (e1), we get:

(2a) DTH $\overset{k4}{\approx}$ DTC with $k4 = \max(k_1, k3) = 0.5$.

From (ii) and (2a), it can be deduced, using (A11), that:

(3a) DT1 $\overset{k5}{\ll}$ DTC with $k5 = k_1/(1-k4) = 0.2$.

The first deduction path leading to (4) has as an equivalent, using (P4), DT2 $\overset{k6}{\ll}$ DTC with $k6 = k5/k_2 = 1$, which cannot be used because the order of $\ll$ is assumed to be < 1. We may, however, use the other deduction path. As, from (3a) and (A10), we have DTC $\overset{k5}{\approx}$ DTC + DT1, it can be deduced from this and (2a), using (P) and (e1), that:

(4a) DT2 $\overset{k7}{\ll}$ DTC

with $k7 = (k4 + k5 - k4k5)/(1-k5) = 0.75$.

Note that two different paths leading to the same formal result in ROM may lead here to different results in ROM($\mathbb{R}$). This poses the as yet unsolved problem of defining heuristics in order to obtain the best result, by pruning the paths that lead to less precise ones (the length of the deduction path not necessarily being a good criterion).

From DTH$\times$KH $\overset{k1}{\approx}$ DTH$\times$KC and DTH$\times$KC $\overset{k4}{\approx}$ DTC$\times$KC, (A3) gives DTH$\times$KH $\overset{k8}{\approx}$ DTC$\times$KC, from which, using (A7) and (e2), we obtain:

(5a) FH $\overset{k8}{\approx}$ FC with $k8 = k_1 + k4 - k_1k4 = 0.55$.

Results (1a to 5a) have thus been obtained by applying rules of ROM($\mathbb{R}$). In terms of intervals, these relations are expressed as:

$DT2/DTH \leq 0.5$ $\quad 0.5 \leq DTH/DTC \leq 2$
$DT1/DTC \leq 0.2$ $\quad DT2/DTC \leq 0.75$
$0.45 \leq FH/FC \leq 2.223$.

They can automatically be translated in terms of formal order of magnitude compound relations w.r.t. scales k_1 and k_2 (with obviously a loss of precision in general), giving:

DT2 $\ll .. \prec$ DTH $\quad$ DTH $\not\preceq .. \not\succeq$ DTC $\quad$ DT1 $\lll$ DTC $\quad$ DT2 $\ll .. \prec$ DTC $\quad$ FH $\not\preceq .. \not\succeq$ FC.

In contrast with results of FOG or O(M) (1 to 5), this time the results are correct. We thus have a sound calculus in $\mathbb{R}$, with all the advantages of the qualitative meaning transmitted by rules of ROM. In particular, each of the above results has its explanation in commonsense reasoning terms given by the rules applied to obtain it. Using symbolic computation, the formal k orders of the results can even be expressed in terms of k_1 and k_2. This can be used for other purposes such as design, so as to tune the values of k_1 and k_2 and thus make sure desired orders of magnitude relations are satisfied.

Nevertheless it can be noticed, by comparing them with exact results (1′to 5′), that these relations are not in general optimal. In fact, in qualitative terms, the exact results would give:

DT2 $\ll .. \prec$ DTH $\quad$ DTH $\not\preceq .. \approx^+$ DTC $\quad$ DT1 $\ll$ DTC $\quad$ DT2 $\ll .. \prec$ DTC $\quad$ FH $\simeq^- .. \not\succeq$ FC,

improving 3 of the 5 above results. The improvement is even more obvious when comparing ranges given by numeric orders k3 to k8 with exact results. Only DT2/DTH is correctly estimated (i.e. (1a) and (1′) are the same). In fact, this is not surprising. Although each rule of ROM($\mathbb{R}$) has been computed with the best estimate for order in conclusion, so that each rule taken separately cannot be improved, this does not guarantee optimality through an inference path using several rules that share common variables. In some way, what we have is local optimality, not global optimality. If we estimate that the obtained results, although sound, are not accurate enough for our purpose, we have to supplement ROM($\mathbb{R}$) with other techniques.

Using Numeric or Symbolic Algebra Techniques

Once sound results of ROM($\mathbb{R}$), with the obvious qualitative meaning of the inference paths, have been obtained, several supplementary techniques can be used in order to refine them if needed. These techniques come from two different approaches: numeric ones which transpose well-known consistency techniques for CSPs to numeric CSPs, and symbolic ones which use computer algebra. These approaches are not exclusive and can be usefully combined.

Applying Consistency Techniques for Numeric CSPs

A first way of refining the results is to start from definitions (I,II,III) of the fundamental relations of ROM($\mathbb{R}$) in terms of intervals, a technique that can easily be extended to all 15 primitive relations. Numeric values are also naturally represented by intervals to take into account precision of observation. Interval computation thus offers itself. Moreover, we are not limited to intervals representing the 15 primitive relations or the compound ones, i.e. to the scale of ROM($\mathbb{R}$); we can in fact express any order of magnitude binary relation between two quantities by an interval encompassing the quotient of the quantities. In particular, intervals do not need the specific symmetry properties of those of ROM($\mathbb{R}$) such as in (I,II,III). Since using intervals is thus more accurate when expressing data, it should also be so for the results. But, unfortunately, interval propagation is rarely powerful enough: in the heat exchanger example nothing is obtained by this method.

The idea is to generalize interval propagation in the same way that, in CSPs, k-consistency with $k > 2$ extends arc consistency. This has been done in [Lhomme, 1993], who shows that the consistency techniques that have been developed for CSPs can be adapted to numeric CSPs involving, in particular, continuous domains. The way is to handle domains only by their bounds and to define an analog of k-consistency restricted to the bounds of the domains, called k-B-consistency. In particular 2-B-consistency, or arc B-consistency, which formalizes interval propagation, is extended by the notion of k-B-consistency. The related algorithms with their complexity are given for $k=2$ and 3. They have been implemented in Interlog [Dassault Electronique, 1991], above Prolog language. In this section, these techniques are evaluated w.r.t. the heat exchanger example.

Starting from equations (e1) and (e2) and assumptions (i′,ii′,iii′), bounds for the 5 remaining quotients are looked for. In this case, as already seen, arc B-consistency gives no result. But 3-B-consistency gives the following results for the first 4 quotients (nothing is obtained for FH/FC) with parameters characterizing the authorized relative imprecision at the bounds $w1=0.02$ and $w2=0.0001$ (in about 75s on an IBM 3090):

(1″) $DT2/DTH \leq 0.508$
(2″) $0.665 \leq DTH/DTC \leq 1.120$
(3″) $DT1/DTC \leq 0.112$
(4″) $DT2/DTC \leq 0.559$.

It can be noticed that estimates (2″,3″,4″) are better than corresponding results of ROM($\mathbb{R}$) (2a,3a,4a) and, for the first two, not far from optimal ones (2′,3′). 4-B-consistency has also been tried, although execution time increases considerably. For example, $0.710 \leq DTH/DTC \leq 1.090$ and $DT2/DTC \leq 0.362$, which well approximate (2′) and (4′), are obtained in a few minutes with $w1=0.01$ and $w2=0.05$.

Although interval propagation alone is in general insufficient, k-B-consistency techniques with $k \geq 3$ may thus provide very good results, but some difficulties remain (here, nothing can be done with equation (e2), unless considering at least 5-consistency with efficiency problems).

Using Symbolic Algebra first

The above results reach the limits of purely numeric approaches. If we want to progress towards optimal results, we have to use computer algebra in order to push symbolic computation as far as possible and delay numeric evaluation In a great number of real examples, the total number of equations expressing the behavior of the system and of order of magnitude assumptions equals the number of order of magnitude relations asked for, and the desired dimensionless quotients can be solved in terms of the known quotients, using these equations. These solutions are very often expressed as rational functions and this symbolic computation can be achieved by computer algebra.

For example, from equations (e1) and (e2), known relations

$DT2=Q1\times DT1$, $DT1=Q2\times DTH$, $KH=Q3\times KC$,

and searched relations

$DT2=X\times DTH$, $DTH=Y\times DTC$, $DT1=W\times DTC$, $DT2=Z\times DTC$, $FH=U\times FC$,

MAPLE V [Char, 1988] immediately deduces the formulas (F):

$X = Q1\times Q2$, $Y = 1/(1-Q2+Q1\times Q2)$,
$W = Q2/(1-Q2+Q1\times Q2)$,
$Z = Q1\times Q2/(1-Q2+Q1\times Q2)$,
$U = (1-Q2+Q1\times Q2)/Q3$,

with $1/5 \leq Q1 \leq 5$, $0 \leq Q2 \leq 1/10$, $9/10 \leq Q3 \leq 10/9$.

Numeric CSP techniques can now be applied directly to these symbolic equations. This time,

results are obtained just with arc B-consistency, even for U:
(1s) $X \leq 0.5$ (2s) $0.666 \leq Y \leq 1.112$ (3s) $W \leq 0.112$ (4s) $Z \leq 0.556$ (5s) $0.810 \leq U \leq 1.667$.

It can thus be seen that, when starting from solved symbolic expressions, the most simple numeric technique, i.e. analog to interval propagation, gives results which are close to the exact ones (1′to 5′) and, in all cases, much better than those given by ROM(ℝ) (1a to 5a). Obviously, using 3-B-consistency improves the results still further, in particular for Z, as follows (with $w1 = 0.001$ in 10s):
(1s′) $X \leq 0.5$ (2s′) $0.713 \leq Y \leq 1.088$ (3s′) $W \leq 0.110$ (4s′) $Z \leq 0.358$ (5s′) $0.827 \leq U \leq 1.556$,
which are practically optimal.

Using Symbolic Algebra Alone for Computing Optimal Results

Symbolically expressing searched quotients in terms of known ones (Qi) leads to expressions which are continuously differentiable in Qi and most often algebraic (rational functions such as in (F)). The problem to be solved can thus generally be expressed as that of finding the absolute extrema of these expressions on n-dimensional closed convex parallelepipeds defined by the ranges of the known intervals $mi \leq Qi \leq Mi$ for $1 \leq i \leq n$. It is well-known that these extrema occur at points where partial derivatives are null. Thus this is a way to compute them exactly from roots of derivatives by using computer algebra.

More precisely, a necessary (not sufficient because it can correspond in particular to a local extremum) condition for an absolute extremum in a neighborhood is the nullity of all the partial derivatives at the given point. A difficulty arises because extrema may be obtained on a face of dimension $< n$ rather than in the interior of the parallelepiped. Thus derivatives on all faces have to be considered. But, thanks to computer algebra, it is sufficient to symbolically compute partial derivatives once and for all and then, in order to obtain derivatives on any face, to fix the Qi, which determine the face, to their numeric values. Roots of all derivatives (in our case roots of a system of polynomials) are computed, firstly in the interior and then on the different faces in decreasing order of dimension, and the corresponding numeric values of expressions at these points are evaluated up to the vertices. These values are finally compared and only the highest and lowest are kept, which correspond to the absolute extrema.

Let us now apply this method, implemented in MAPLE V, to the heat exchanger example. Expressions X, Y, W and Z depend on the 2 variables Q1 and Q2 and are thus considered w.r.t. the rectangle $1/5 \leq Q1 \leq 5$, $0 \leq Q2 \leq 1/10$; U, which depends on the 3 variables Q1, Q2 and Q3 is considered w.r.t. the parallelepiped based on the previous rectangle with $9/10 \leq Q3 \leq 10/9$. Results are computed immediately and summarized below.

For X, Y, W and Z, it is found that only their derivatives w.r.t. Q1 are null on the edge $Q2=0$. Corresponding constant values $X=0$, $Y=1$, $W=0$ and $Z=0$ are shown, after inspection of vertices, to be the minima for X, W and Z, but not an absolute extremum for Y. Looking now at the vertices, it is found that the maximum of X is obtained at the vertex $Q1=5$, $Q2=1/10$ and is equal to $1/2$; the minimum of Y is reached at $Q1=5$, $Q2=1/10$ and is equal to $5/7$, and its maximum is reached at $Q1=1/5$, $Q2=1/10$ and is equal to $25/23$; the maximum of W is obtained at $Q1=1/5$, $Q2=1/10$ and is equal to $5/46$ and the maximum of Z is obtained at $Q1=5$, $Q2=1/10$ and is equal to $5/14$.

The derivative of U w.r.t. Q1 is null both on the edge $Q2=0$, $Q3=9/10$ corresponding to the constant value $U=10/9$ and on the edge $Q2=0$, $Q3=10/9$ corresponding to the constant value $U=9/10$. But it is finally found that the minimum occurs at the vertex $Q1=1/5$, $Q2=1/10$, $Q3=10/9$ and is equal to $207/250$, and that the maximum occurs at the vertex $Q1=5$, $Q2=1/10$, $Q3=9/10$ and is equal to $14/9$.

Finally computer algebra, which works with rational numbers, gives the exact solutions (S) to our problem:
$0 \leq X \leq 1/2$, $5/7 \leq Y \leq 25/23$, $0 \leq W \leq 5/46$, $0 \leq Z \leq 5/14$, $207/250 \leq U \leq 14/9$.
Floating point approximation with 3 significant digits gives (1′ to 5′).

The method of roots of derivatives, processed by computer algebra, is thus a very powerful technique to automatically obtain the exact ranges. But, in addition to the complete loss of the qualitative aspect of the inference and the necessity, as in the above subsection, for the system of equations to be algebraically solvable, there are two other drawbacks to this approach. The first one is that roots of a polynomial system cannot in general be obtained exactly. This is solved in practice in a large number of cases by using the most recent modules of computer algebra which are able to deal with algebraic numbers (represented as a couple of a floating point interval and a polynomial, coefficients of which are algebraic numbers, such that the considered number is the only root of the polynomial belonging to the interval). The second one is the exponential complexity of the method: in an n-dimensional space we have 3^n systems of polynomials to look for, from the interior to the vertices. The method becomes intractable very rapidly unless the number of variables (assumed order of magnitude relations) remains very small.

Syntactically Transforming Rational Functions: a Line of Research

There are cases where, after having judiciously syntactically transformed rational functions which are solutions of the set of equations, the simple interval propagation technique gives the exact optima, as illustrated in the example.

Let us consider symbolic formulas (F). The exact result (1s) can be obtained simply by interval propagation for X because variables Q1 and Q2 have only one occurrence in X. It is not the case for the other 4 formulas, which is why, in this case, interval propagation does not give exact results (2s to 5s). However, a simple trick may be found by hand to satisfy this condition. In fact the expression $1-Q2+Q1\times Q2$ in Y, W and U may be rewritten as $1+Q2\times(Q1-1)$, which boils down to changing a variable: $Q1-1$ instead of Q1. A simple interval propagation gives $23/25 \leq 1+Q2\times(Q1-1) \leq 7/5$, from which exact solutions (S) for Y, W and U are immediately obtained. It is not the case for Z because Q1 appears also in the numerator. But Z can be rewritten as $Z=1/(1+(1/Q1)(1/(Q2-1)))$ where each new variable $1/Q1$ and $1/(Q2-1)$ appears only once. The exact result (S) $0 \leq Z \leq 5/14$ follows immediately. This interval propagation may be achieved exactly by manipulating rational numbers, or with a given approximation by manipulating floating point numbers, as is done by Interlog with 10 exact significant digits.

It can be concluded that, when expressions can be rewritten by changing variables, such that each new variable occurs only once, simple interval propagation gives exact solutions. This transformation is obviously not always possible. A line of research would be to characterize the cases where such a transformation of rational functions (or at least a partial one which minimizes the number of occurences of each variable) is possible and to find algorithms to do this.

Conclusion

It has been shown in this paper that the formal system ROM(K) [Dague, 1993] can be transposed in $\mathbb{R}$ in order to incorporate quantitative information easily, and to ensure validity of inferences in $\mathbb{R}$. Rules of ROM($\mathbb{R}$) thus guarantee a sound calculus in $\mathbb{R}$ (which was not the case with FOG, O(M) or ROM(K)), while keeping their qualitative meaning, thus guiding research and providing commonsense explanations for results.

If the loss of precision through inference paths is such that some of these results are judged to be too imprecise for a specific purpose, several complementary techniques can be used to refine them. k-consistency algorithms for numeric CSPs, which generalize for $k>2$ interval propagation, generally improve the results but may require a large k, in which case they are very time consuming. A better approach is first to use computer algebra to express dimensionless quotients for which approximation is searched in terms of quotients for which given bounds are assumed, and then to apply k-consistency techniques to the symbolic expressions obtained. It has also been shown that computer algebra alone may be used to obtain exact results, by computing roots of partial derivatives in order to obtain the extrema of the expressions on n-dimensional parallelepipeds although this method, which is exponential in n, is tractable only for a small number of variables (i.e. known quotients). Finally, future work would consist in formally modifying rational functions in order to have a minimal number of occurrences of each variable, thus making interval computation more precise; in particular, when it is possible to have only one occurrence for each variable, simple interval computation gives the exact results.

All this assortment of tools, with ROM($\mathbb{R}$) as the basis, is now available to perform powerful and flexible qualitative and numeric reasoning for engineering tasks, and will be tested soon on real applications in chemical processes.

Acknowledgments

This research was done partly at the Paris IBM Scientific Center, to whom I extend my thanks, and partly at the University Paris 6, where I had very fruitful discussions with Daniel Lazard on computer algebra. I would also like to thank Olivier Lhomme for experiments with Interlog and Emmanuel Goldberger for applying this work to chemical processes. Finally, my thanks go to Rosalind Greenstein for the English version of this paper and to Gilles Dague for the figure.

References

B.W. Char, *MAPLE Reference Manual*, Watcom, Waterloo, Ontario, 1988.

P. Dague, "Symbolic reasoning with relative orders of magnitude," *Proceedings of the Thirteenth IJCAI*, Chambéry, August 1993.

P. Dague, P. Devès, and O. Raiman, "Troubleshooting: when Modeling is the Trouble," *Proceedings of AAAI Conference*, Seattle, July 1987.

Dassault Electronique, INTERLOG 1.0, User's Guide, 1991.

D. Dubois and H. Prade, "Order of magnitude reasoning with fuzzy relations," *Proceedings of the IFAC/IMACS/IFORS International Symposium on Advanced Information Processing in Automatic Control*, Nancy, 1989.

O. Lhomme, "Consistency techniques for numeric CSPs," *Proceedings of the Thirteenth IJCAI*, Chambéry, August 1993.

M.L. Mavrovouniotis and G. Stephanopoulos, "Reasoning with orders of magnitude and approximate relations," *Proceedings of AAAI Conference*, Seattle, July 1987.

M.L. Mavrovouniotis and G. Stephanopoulos, "Formal order of magnitude reasoning in process engineering," *Comput. chem. Engng. 12*, 1988.

O. Raiman, "Order of magnitude reasoning," *Proceedings of AAAI Conference*, Philadelphia, August 1986.

O. Raiman, "Order of magnitude reasoning," *Artificial Intelligence 51*, 1991.