A CONTRIBUTION TO THE INVESTIGATION OF ALGEBRAIC MODEL STRUCTURES IN QUALITATIVE SPACE

Franjo Jović [(1)], Alan Jović [(2)], Velimir Rajković [(3)]
[(1)]University of Osijek, Faculty of Electrical Engineering
Kneza Trpimira 2b, 31000 Osijek; Tel: 091/ 224 60 49; fjovic@inet.hr
[(2)]University of Zagreb, Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb; Tel: 091/ 767 02 03; alan.jovic@gmail.com
[(3)] Zagreb Polytechnic for Technical Sciences, Electrotechnical Department
Konavoska 2, 10000 Zagreb; Tel,Fax: 01/3665-011; vrajkov@vtszg.hr

Abstract - Quantitative and qualitative models. Basic forms of qualitative space. Code and numerical patterns in qualitative space. Qualitative process models. Latent and contradicting models. Completeness and acceptability measure of qualitative models. Examples of qualitative models of production parameters of the oil well log, seismic signals from the oil deposits and laser welding process. Quacol algebra and possibilities of its extensions. Criteria for usage of specific algebraic operations. Comparison of model structures from the standpoint of accuracy, simplicity and solution stability. Conversion of qualitative models to its quantitative counterparts and determination of mean value and amplification factor of the quantitative model. Comparison of the interim step of the Medusa-T and Medusa-H algorithms. Numerical limits of circular qualitative models.

"Reality resists imitation through a model."
Erwin Schrödinger in "The Present Situation in Quantum Mechanics", part of Quantum Theory and Measurement, Wheeler & Zurek, *Princeton University Press*, New Jersey, 1983.

I Introduction

Quality as a term seems to have more marketing then mathematical conotation. Strangely enough precision – the most seeken mathematical feature – is by far on the side of quality. The theoretical reason for this lays in the fact that a limit procedure in evaluation of differentiating certain process values depends on the quanta in mathematics and on shapes difference in the qualitative engineering – quanta being defined by the process noise and shapes by virtue of mental process like a pattern. Principal power of analytical difference between noise and pattern needs no further explanation.

While not possessing the notion of zero and negative number the qualitative engineering approach circumvents the seventeen meaning of zero and each specific meaning of negative – specific to each process value. Finally the quantitative engineering manipulation of algebra in the named negative region is not correct with process values where for instance (-a)*(- b) = -ab that is a correct form for processes and a complete disaster in mathematics. For two simple sets of process variable data a = {5, 2} and b = {4, 1} their multiple pairs are a*b = {20, 2}. Now if both variables are decreased for the same constant 3 the corresponding multiple pairs are {(5-3)* (4-3) = 2, (2-3)*(1-3) = 2 } and the rank order has been distroyed by the multiplication operation. Applying the multiplication rule (-) * (-) = (-) one will get {2, -2} with the preservation of rank order during multiplication. This feature is enormously important in solving inequalities [1]. Qualitative modeling has been quite popular in mid-ninetees of the 20th century. Due to the inappropriate treatment of the algebra development, like SR1 or QUASIM, it gradually faded away. Still its results obtained in modeling were sound and very much appropriate. The reason for its relative decay was dropping out the basic circular feature of the information arising in information tools. Following Železnikar's work on information nature and by practicaly postulating the circularity feature in 1997, the QUACOL algebra is an usable tool for qualitative process modeling and thus enabled reinvention of qualitative modeling [2],[3].

II Qualitative model – a similarity logic function

What is a qualitative model? A simple differential equation given with general parameters represents a qualitative model of a process. By exchanging general parameters with specific values i.e. at the very moment of calculation qualitative analytical model becomes a quantitative analytical model. By this very act it obtains the analytical expressiveness but loses the abstraction.

Still simpler qualitative model is a behavioral similarity of process variables. If for instance by seeking a model for the process variable A one finds another process variable B that behaves in a similar manner then B becomes a candidate for the model of A. Thus

$$B = Model(A) \propto A \qquad (1).$$

The sign $\propto$ stands for similarity operation. But what is similarity? In its discussion of fallability of modern logic mathematicians state that the only operator lacking in logic is the similarity operator. Thus if we postulate similarity as the fourth logical symbol $(\vee, \wedge, \neg, \propto)$ and correspondently two-element algebra consisting of the "truth" and "false" values as T and $\perp$ it is worth to introduce first order logic with the similarity operation.

It is clear how to define Boolean operations on limited and unlimited qualifiers; the Boolean operations are denoted using the usual symbols $(\vee, \wedge, \neg, \propto)$ for each propositional function *F* i.e. $Q1 \vee Q2$ is a function such that $Q(F) = Q1(F) \vee Q2(F)$. Since Mostowski gave the theory of the propositional functions of qualifiers earlier for $(\vee,\wedge,\neg)$ [4], we shall only briefly discuss the same features for similarity operation. Here two qualifiears are disposable: existential qualifier $\exists$ and general qualifier $\forall$. Introducing a limited qualifier *Qs* and *I* an arbitrary set with *I* = I x I x I…* its Cartesian power i.e. the set of infinite sequences *(x1, x2, …)* with $x_j \in I$ for *j* = 1, 2, … .

A mapping *F* of *I** into $<\mathrm{T},\perp>$ is called a propositional function on *I* provided that it satisfies the following condition: there is a finite *K* of integers such that if *x = (x1, x2, …)* ∈ *I**, *y = (y1, y2, …)* ∈ *I* and* $x_j = y_j$ *for* $j \in K$, *then F(x) = F(y).*

Let us define the rank operator $\Re$ as a limited qualifier on *I*. It assigns a rank value to each element x_i of *I* as an individual variable *S* ranging from {1,2,…,*k*} to each functional variable of the degree *k*. The case of individual variable *S* possessing the same rank of its particular elements *i* and *j* can be dropped off by adding a small amount of noise to each particular variable *S*. Each element of two *S* variables will be called an *I*-valuation consisting of rank comparison of correspondent elements. Thus

$$Val_p\{S_i \mid S_j\} = \begin{cases} \mathrm{T}\ iff\ (\Re_p S_i = \Re_p S_j)\ for\ all\ p \in 1,2,...k \\ \perp\ otherwise \end{cases} \quad (2),$$

presents the formula that is true in *I* if (2) exhibits logical T and satisfiable in *I* if it exhibits true value for some parts of *p*. The degree of satisfiability can be measured as the rank correlation coefficient.

A latent similarity model *C* of the process variable *A* is a model that satisfies (1) i.e. it is similar to the variable *B* as $C = Model(B) \propto B$. The modeling operation is not a process transitive operation and thus not firmly bounded to *A*. It leads in the extreme to the effects that if increasing *B* means increasing a certain process value x then *C* means decreasing the same value of *x*. Such models are process contradictory.

In order to evaluate (2) for *k* = 9 it is worth to say that coincidence of having all ranks equal evaluates to 1/9! or approximately 2 in a million.

Qualitative model is complete when it covers all essential process features regarding function *A* that is being modeled. Qualitative model is completely acceptable when it is convertible into a appropriate quantitative form.

III. QUACOL algebra and its simple structural extension

QUACOL algebra is a circular evolutional algorithm using alternatively quantitative and qualitative aspects of the same data [3]. It consists basically of algebraic operations on process data series. Criteria of algebra usability demand fulfillment of the valuation given in (2). Algebra calculates also with *S* variable and its inverse, thus pushing forward positive correlation of the variable with given goal function or a set of goal functions. Thus if there is no direct model such as given in (1) two or more *S* variables are included (with their corresponding quantitative parts as well) and a general form is obtainable as

$$Model(A) \propto B\ op\ \{+,-,*,/\}\kappa C \quad (3),$$

where disposable algebraic operators are indicated in brackets and κ is influence factor of the variable *C* in the *B op C* composition. When for instance there is $\kappa = 0.3$ and algebraic operator is addition {+} then model variable *B* is complemented with 0,3 part of the model variable *C* in each of the *k* steps of the qualitative model. A broker can say that in order to folow dollars (*A*) one has to take a mix of 70% yen (*B*) and 30% euro (*C*), meaning that what was in a time point not enough in *B* was compensated by *C* and vice versa. Subtraction operator balances the difference between two model variables meaning that the amount of yen value to be bought should be diminished by the 30% of the euro value in order to follow dollar value.

Simplicity, precision and stability are criteria for using algebraic operators when aiming at obtaining the modeling criterium given in (2).

Numeric stability depends on determination of quantitative values for equal ranks – which is an unstable situation and on controlling marginal values for minimum and maximum value in quantitative part of the model.

The case of equal ranks usually follows in the second step of the Medusa algorithm method, i.e where the difference between the model $\Re(B)$ and $\Re(\Re(A)-\Re(B))$ enters into calculation. This feature has been treated in the Medusa T algorithm as a simple quantitative data series imposing a lower hit rate in the algorithm then expected. The solution in Medusa H algorithm is proposed as a qualitative differential function of the expanded data series [6] permitting more direct comparison of the interpolation data series C in the model synthesis. In this way patterns are used and not quantitative values.

Generally a QUACOL algorithm tends to instability when maximum value runs away from the numeric margins of the computer algorithm, or it demands immense complexity in its numeric application part regarding the same rank situation.

In order to control numerical divergence in expressions of the type (3) a new analytic form for modeling in QUACOL algebra is proposed that is based on the general analytical form as

$$Model(A) \propto B\ op\ \{+,-,*,/\}\kappa C - \\ - mean\{B\ op\ \{+,-,*,/\}\kappa C\} \quad (4),$$

where *mean*{} represents a mean value of the quantitative part of the model in brackets. Such form prevents instability observed in some process model calculations. *B* and *C* are only optionally given since the forms (3) and (4) are recursive goal driven forms and *B* can be in a next step the whole expression (4) and so on.

Thus models with greater correlation can be built that enable appropriate quantitative model conversion.

IV. Quantitative models from QUACOL algebra

When a model is completed in its qualitative form it contains quantitative data part with data that are principally very very far away from the expected quantitative model. Restauration of the quantitative part can be done from calibrated data so that for instance if we know that the third ranked value is 36 and the minimum tenth ranked value is 4 then a linear model can be made that can intra/extrapolate the rest of the model. The slope of the conversion straight line equals to (36 – 4) / (10 – 3) = 4,571… and the mean value is 24,571…

No calibrated quantitative data and only similar quantitative known data can be a problem [7]. Further on no one can claim linearity of the quantitative model. In this case the recalculation of the mean value and amplification factor can be an outcome. Fitting can be done from similar already known data profiles such as known production data profiles of the oil well [7].

In such a way even the nonlinearities can be compensated.

V. Experimental results

Tables 1, 2, and 3 present the cases of different process models. Table 1 shows modeling possibilities on seismic data profiles of the Beničanci crude-oil field. Correlation coefficient and stability of the numerical solution have been tested for cases from (3) and (4). The case of modeling working point of the power laser installation [5] are given in Table 2. with the same algorithm demands.

Similarly data are given for modeling of the vertical profile of the Žutica oil field wel production data in Table 3.

Authors are thankfull to the INA-Naftaplin experts from the Exploration Department and to Prof. Dieter Schuöcker from the Vienna University of Technology, Institute for Power Laser Research, for having the possibility to use their process data files for modeling purposes.

Table 1. The model of Beničanci oil-field seismics data; maximum number of algorithm iterations for particular algebraic operation with and without data equalization

	Without equalization					With equalization				
	+	-	*	/	Opt	+	-	*	/	Opt
Iteration	6	5	6	6	6	8	15	3*	3*	7
Rank co-rrelation	0.901	0.84	0.88	0.88	0.92	0.989	0.934	0.934	0.934	0.989

(*) third iteration data diverge for multiplication and division operations

Table 2. The model of power laser parametrisation data; maximum number of algorithm iterations for particular algebraic operation with and without data equalization; many variables possess equal rank values

	Without equalization					With equalization				
	+	-	*	/	Opt	+	-	*	/	Opt
Iteration	13	10	6	6	6	6	7	6	6	6
Rank co-relation	0.93	0.96	0.87	0.87	0.95	0.92	0.95	0.88	0.88	0.95

Table 3. The model of Žutica oil-well log geology data; maximum number of algorithm iterations for particular algebraic operation with and without data equalization

	Without equalization					With equalization				
	+	-	*	/	Opt	+	-	*	/	Opt
Iteration	5	5	4	3	3	5	5	2*	2*	3
Rank co-rrelation	0.88	0.89	0.89	0.89	0.89	0.88	0.89	0.88	0.87	0.89

(*) second iteration data diverge for multiplication and division operations

REFERENCES

[1] Beckenbach E, Bellman R., "An Introduction to Inequalities", *The Mathematical Association of America*, 1961.

[2] Bogunović N., Jagnjić, Ž. Jović F., "An Efficient Automated Learning of Qualitative Process Models", in Wang, Q.G., Ed., The 4th Asian Control Conference Proceedings, Singapore: Causal Productions, pp2107-2111, September 25-27, 2002.

[3] Jović F., "A Circular Qualitative Algebra", *International Journal of Computing Anticipatory Systems*, Vol.8,2001, pp213-224.

[4] Mostowski A., "On generalization of quantifiers", *Fund. Math*. Vol. 44, pp 12-36, 1957.

[5] Jović F., Rajković V., Lugomer S., Stipančić M., "Precision of Quality Prediction in Laser Welding", *Lasers in Engineering*, 2000, Vol.10, pp193-212.

[6] Bogunović N, Jagnjić Ž., Jović F., "Qualitative Modelling of Complex System based on Similarity Calculus", in Amborski, Meuth, H.,

Eds., Proceedings of the 16th European Simulation Multiconference 2002, ESM2002, Darmstadt, Germany, June 3-5, 2002, pp 312-316.

[7] V. Rajković, "Calculating quantitative data on oil-well production profiles from qualitative models", to be published.