A COLLECTIVE OF ALGORITHMS

L.A. Rastrigln
Institute of Electronics and Computing Technique,
Academy of Sciences,
R.H. Erenshtein
Riga Medical Institute
Riga,Latvian SSR,USSR

A concept of a collective of algorithm is proposed,which solve problems from the described set of problems. In this case the efficiency of the collective solution over the whole set appears to be higher than the best algorithm of the collective. Examples of the work of the collective of algorithms are given when solving problems of pattern recognition root determination.

In the typical problems of the behaviour synthesis, of the choice for the optimal solution, of the selecting the best plan ect., the difficulty as a rule arises in the choice of the best algorithm (in the definite sense) for the problem solution from the available set of algorithms. In such situation the invention of a new algorithm is not expedient and in any case it is not economical.

Living beings can effectively solve this problem and in typical situation they behave by combining algorithms already known to them. Let us formalise a problem.

Consider the following rather widely spread situation. Let some set (finite or infinite) of problems of one type $\{X_i\}$, i=1,2,... is given. These problems are coded by n parameters $x_1,\ldots,x_n$ so, that eaoh problem in the space R^n is represented by the point $X=(x_1,\ldots,x_n)$, and the whole set of problemsforms tt region Q $(X_i\in\Omega, i=1,2,\ldots)$.

Let also a finite set of algorithms $\{A\}=\{A_1,\ldots,A_L\}$ for solving problems $I\pounds Q$ be givin. Eacn of these algorithms A_1 (I-*1,...,L) solves or does not solve the problem X^*. The efficiency of solution is determinated by some given criterion Q_t which to the pair $I._x$ and A-. relates the q number:

$$q=Q(X_i,A_l) \qquad (1)$$

which characterizes the performance of the algorithm A-, when solving the problem

Let us give examples of the problems of such kind. The pattern recognition problem, where I is a particular pattern recognition problem, {A} is a set of recognition algorithms, e.g., Bayes method, method of potential functions, method of perceptron etc., q is a probability that recognition is correct.The root determination problem where I is a particular function the root of which is to be determined. {A} is a set of algorithms for determining the root, e.g,, methods of tangents, of chords, of dichotomy etc.,q is the number of iteration for convergence into £-point of a root.The problem of the search optimisation where X is a situation set in the course of optimisation, {A} is a set of the search algorithms, e.g., gradient method,random search method etc., q is mean increment of a function to be optimized or the number of unsuccessful steps.

The choice of the optimal algorithm for solving the problem X_{it}thus, reduses to solving the extremum problem:

$$Q(X_i,A_l) \xrightarrow[l=1,\ldots,L]{} \text{extr} \Longrightarrow A_j, \qquad (2)$$

where A_j is the best from the available algorithms for solving the problem X_i.

It is expedient for solving the set problem to consider the use of the methods of collective decision making [l,2]. Usually we mean by thiB or that form of the voting algorithm. To realise the collective solution we ought to solve the problem by using each *of* the available algorithms $A_1, A_2,\ldots,A_L$ (i.e.to>make enumeration), and then by means of some procedure to generalize the individual solutions in the collective. The main shortcoming of the voting algorithms is first of all the potential possibility of arising an Arrow paradox p] ,which is known in the collective solution theory, or nontransitlvity paradox. This paradox arises due to contradition of individual preferences and then it is impossible to work out a generalized solution. Another shortcoming of the voting procedures is that we are not sure that the solution made by the algorithm, separately taken from the {A}, in the sense of criterion Q, better then the collective solution.

On the other hand the solving of the optimization problem (2) is associated with solving the problem I, only once, which makes senseless futher choice of the optimum algorithm because the problem is already solved. That is why the problem of constructing the function j=f(X), determining the number of the optimum algorithm j for the solution of the problem I, Bhould be solved without full enumeration.

The approach proposed in the paper allows to eliminate the mentioned shortcomings of the voting and the enumeration algorithms, although it belongs to algorithms for the collective decision making,but in the more wide sense.

The basis of the approach proposed is the suggestion that in the space of problems R^n there are compact preferences regions of algorithms $A_1, A_2, \ldots, A_L$. The preference region B_i of the algorithm A_i consists of those problems X_i ,in the course of the solution of which the given algorithm extremes the given efficiency criterion (1):

$$X_i \in B_j, \text{if } Q(X_i, A_j) = \underset{l=1,\ldots,L}{\text{extr}} Q(X_i, A_l), \quad (3)$$
$$j=1,\ldots,L.$$

In the course of solving the particular problem X preference is given to that algorithm from the set A, to the preference region of which this problem in the space R^n belongs [3-5]:

$$A_j \longrightarrow X, \text{ if } X \in B_j. \quad (4)$$

Such approach reminds very much of a "dictatorial" one, which is known in the theory of the collective decision making and which, as it is known, has no nontransitivity paradox [2]. Each preference region, however, has its own "dictator", that is why the approach remains collective. It was shown experimentally [3-5], that such way of choice, on the average,allows to solve the problem better than the best of the competitive algorithms of the set {A}•

It is evident that to realise such approach it is necessary to learn to construct the estimates of the algorithm preference regions and assign the new problem X to one of these regions.For this it is expedient to use the known learning algorithms [6,7].

In the paper as the learning algorithms modifications of the algorithms of the potential function [6]and the Fix-Hodges algorithm [7]are proposed.The basis of any learning algorithm "with the teacher" is the training sequence, which should consist of problems which are representatives of each preference region (3). To form it one should use the already solved problems, i.e. "the experience" of the algorithms from the set {A}• Let us consider in brief each of the learning algorithms.

To realise the method of potential functions 6 one should in each point of the training sequence,i.e. for each problem $X_k, k=1,\ldots,K$, a charge g_{lk} is set, which form around it the potential surface of the following type:

$$\varphi_l(X, X_k) = g_{lk} / [1 + \alpha_{lk} \rho^2(X, X_k)], \quad (5)$$

where

$$g_{lk} = \begin{cases} +1, \text{if } Q(X_k, A_l) = \underset{p}{\text{extr}}\, Q(X_k, A_p), \\ -1, \text{if } Q(X_k, A_l) \neq \underset{p}{\text{extr}}\, Q(X_k, A_p), \end{cases} \quad (6)$$

$p=1,\ldots,L,$

α_{lk} is the attenuation factor of the potential function; $\rho(X, X_k)$ is a distance (in the wide sence) between the problem X and the problem X_k of the training sequence in the space R^n.

Futher for the given particular problem X and each algorithm $A_l, A_l \in \{A\}$, a summary potential is constructed:

$$\Phi_l(X) = \sum_{k=1}^{K} \varphi_l(X, X_k) \quad (7)$$

and for solving the problem X an algorithm with the number j is selected,for which

$$\Phi_j(X) = \max_l \Phi_l(X), \; l=1,\ldots,L. \quad (8)$$

The meaning of the Fix-Hodges algorithm is as follows. For each newly set problem X in the space of the problems R^n there are m the most similar problems from the ones already solved. The "similiarity" can be evaluated by the value of the Euclidean distance:

$$\rho(X, X_k) = |X - X_k| \;, \; k=1,\ldots,K. \quad (9)$$

The preference in the solving of the problem X is given to the algorithm which extremizes criterion Q in the most of these m problems.

To illustrate the above statements we consider a series of experiments.

The pattern recognition problem

This problem appears to be the most relevant for the use of the approach proposed. The case is that there is no necessity to specially form the problem space R ,such space is the property space. Each object in this space represents a narrow recognition problem.As for criterion Q,it was said above that it means a correct recognition of the given objects.

Experiments were done by using both the model and the real recognition problems. As the model problem, the problem of recognizing two classes,the deviding surface between which is nonlinear, was considered. The dimensionality of the space R was equal to two. As the competitive algorithms two linear decision rules A_1 and A_2 were chosen. The volume of the training sequence was equal to 100,for the checking 10 000 realizations were examined. As the learning algorithm in all the recognition problems solved the method for potential functions was chosen.

The table 1 gives the values of the probabilities of the recognition error by using each of the rules,the collective of the rules and the method of potential functions.

Table 1.

A_1	A_2	$\langle A_1, A_2 \rangle$	The method of potential functions.
0,3156	0,4429	0,0393	0,0826

It follows from the table that the proposed approach is advantageous both as to each of the algorithms and as to the known algorithms of potential functions. The analogous conclusion can be done also from the following experiment.

A problem of recognizing accident situations in the electric and energetic system is set* The accident situation in the electric and energetic system happens most frequently due to redistribution on the power flow over the transmission lines,the accident flow of the active power exceeding the admissible level of the transmiesive capacity of the given transmission line. The property space consists of 6 measurements,the volume of the training sequence is equal to 100, for cheking 330 realizations are examined. A set *{A}* consisted of 6 linear decision rules. In the table 2 the results of this problem solving by a series of conventional algorithms and by the mentioned collective of rules are given.

Table 2.

Method of the "minimum distance" to the middle point.	Method of potential function.	Method of the "nearest value".	Qvadratic approximation.	$\langle A_1, \ldots, A_6 \rangle$
0,325	0,212	0,172	0,203	0,109

The third experiment considers the case when elements of the set *{A}* are not formalized. This case corresponds to the collective of specialists, in this example it is an artificially formed consilium. The organized consilium was solving the problem of recognizing the myocardial infarction using EGG. The table 3 gives the results of the work of each of the specialists as well as the quality of solution obtained by means of the proposed approach and the voting algorithm.

Table 3.

A_1	A_2	A_3	A_4	A_5	A_1-A_5	Voting
0,324	0,304	0,376	0,454	0,412	0,244	0,304

It is evidently seen that the given approach is advantageous compared with the voting algorithm.

The problem of the function root determination

This problem is set in the following way. A set {A} is given, which consists of four elements: A- is a dichotomy method, A„ is a method of chords, A_3 is an iterative method, A_{n4} is a tangentmethod. The problem BpaceR consists of the raeasurments of the function values with the definite values of an argument. The dimensionality of this space was equal to 5. For learning and checking a special programme generated functions of one class. The training sequence consisted of 100 realizations, checking one - of 250 realizations. As the criterion Q the number of search steps was chosen. The decision as to what algorithm to prefer was made by using the Fix-Hodges method. The number m was found which give the possibility to lessen the mean number *of* steps taken by this algorithm for searching roots of the checking functions, i.e.to improve the quality of its performance. Table 4 gives the results of the experiments.

Table 4.

A_1	A_2	A_3	A_4	$\langle A_1, \ldots, A_4 \rangle$
4,552	3,556	8,160	19,984	3,486

The results of experiments allow one to made a conclusion that the approach proposed is universal. It is evidently that the main difficulties associated with its applicability concerns the forming of the space R^n, which is adequate to the problem to be solved.

The approach proposed can be used to solve problems of the pattern recognition, of optimization,of optimum control,etc.On the other hand it can be used as a model for understanding the processes of selecting the behaviour of organisms in the typical situations. In this case the way of investigation is obvious- it is necessary to describe the space of problems R, which are being solved by the organise single out a set of the behaviour algorithms {A} and according to observation of the behaviour in different situations, i.e. in solving problems from the set {x}, to single out "the competence" region of each behaviour algorithm A_j, If these regions do not intersect, then it is reasonable to suppose that the living being in selecting the behaviour realizes the algorithm described above.

Bibliography

I.R.Duncan Luce and Howard Raiffa. Games and Decisions.,New York, Wiley & Sona 1957.
2.G.T.Guibaud.Les theories de l'interest general et le probleme logique de l'aggregation. ,"Economic Appliquoe", N4,Paris Publications de l'Universits Francois,1952.

3. L.A.Rastrigin,R.H.Erenshtein. Learning of a collective of solving rulee. "Adaptive Systems",v.4. ,Riga,Zinatne, 1974,(Bus)
4. L.A.Rastrigin,R.H*Erenshtein. A collective of solving rules as an effective recognition algorithm. Materials of VI Allunion meeting on management problems,v. 1 .,Moscow,Nauka,1974,(Rus)
L,A.Rastrigin,R.H»Erenehtein. On a method of collective decision making.Ma-
5. terials of VII Allunion simpoeium on cybernetics, Tbilisi, 1974.(Rus)
M.A.Iserman,E.ILBraverman,L. I.Rozonoer The method of potential functions in maschine learning theory.Moscow,Nauka» 1970,(Rus)
V.N.Vapnik,A.J.Chervonenkis. The theory of pattern recognition.Moscow,Nauka,
7. 1974,(Rue)