PhD THESES

KAPOSVAR UNIVERSITY
FACULTY OF ECONOMIC SCIENCES
Department of Information Technology

Head of Doctoral School:
PROF. DR. GÁBOR UDOVECZ
Doctor of Hungarian Academy of Sciences

Supervisor:
DR. HABIL. BÉLA CSUKÁS CSc
Associate Professor

DEVELOPMENT AND APPLICATION OF GENETIC ALGORITM FOR MULTICRITERIA ECONOMIC DECISION MAKING

Author:
Sándor Balogh

KAPOSVÁR
2009

1 BACKGROUND AND OBJECTIVES

In the conventional optimization methodologies the model of the investigated problem used to be simplified to the formalism of a mathematical construct that makes possible the determination of the exact optimum.

Considering the importance of the details in the engineering problem solving, in the past decades increasing effort has been made for the possibly most detailed model based optimization. The dynamic simulation tools for the various processes developed rapidly, while the optimization of more and more complex hybrid (continuous/discrete) models became intractable in the development of in-parallel optimization methods. Another actual challenge is the optimal solution of the large scale, long term processes of changing structure with increasing complexity.

Having recognized these difficulties, the inexact, heuristic and/or evolutionary methods of the artificial and computational intelligence became more and more important. In the case of an inexact approach there is no guarantee for the determination of the absolute optimum. This disadvantage is compensated by the fact, that good enough solutions can be determined on the basis of the necessarily most detailed models. One of the heuristic optimization methods of the Computational Intelligence is the Genetic Algorithm. I met these methods in my MSc thesis in the early 90-th, and have been dealing with it in my work since that time. This work is characterized and motivated by the demand on optimize detailed and/or large scale hybrid problems, coming from various field of application, which could not be solved with the available tools. In addition, in the case of an economic objective we had to optimize according to a number of natural criteria, or we had to combine the economic and natural objectives. The in-parallel developing generic simulation method more and more tolerates the arbitrary discrete and continuous changes. In addition, in accordance with the general tendency continuously increased the computational demand for the simulation of the possible solutions.

The fundamental objective of this work is the comparison of the published methods with my results came from the continuous development of the genetic algorithm, motivated by the solution of the various practical problems in the

past decades.

Based on this comparison, a combined genetic algorithm has to be developed for supporting the multicriteria economic decisions.

The detailed objectives of my research are the followings:

1. Considering the practical needs, a complex genetic coding is to be elaborated that makes possible the unified description of the optionally hierarchical implementation of genes, consisting of continuous/discrete and permutational subsequences. Moreover the coding has to support the declaration of the incompatible genes by means of the so called structure lattice. The planed complex coding has to describe the conventional kins of code as special cases to avoid the application of special specific methods.

2. In accordance with this unified coding, utilizing the set of available genetic operators and optionally introducing modified operators a general framework is to be developed for the automatic selection of the respective operators to the actual coding.

3. Another important objective is to elaborate a parametrizable method with the respective parametrizable genetic operators that support the execution evolutionary process with small generation number and population size. Accordingly, taking into account the experiences from the literature and from the successfully solved practical problems a new, so called grid method is to be developed.

4. The multi criteria problem solving is to be improved by an algorithm, calculating the combined Pareto-dominance with the respective parametrizable genetic operators.

5. Finally, I am to implement the elaborated algorithms for the evolutionary optimization of less number of variance with grid computational demand. Accordingly, I try to utilize the possibility to store all of the generated genetic codes, on the one hand. In addition, my objective is the automatically controlled parallel simulation based evaluation, utilizing

the possibilities of the macro-granularly parallelizable (e.g. PC cluster) architectures.

The application of the developed methods for the solution of decision making, economic optimization and logistical problem solving will be illustrated by simple examples, followed by a couple of test problems elaborated for the evaluation of genetic algorithms.

2 MATERIALS AND METHODS

In the present research and development work many open source code software development tools, as well as the collaborating generic simulator, developed by our research team were applied. For the realization of the macro granularly parallel evolutionary development, a computer cluster was built and configured. Accordingly, the methods applied for the development and testing of the elaborated genetic algorithm were the followings:

- software tools, applied for the development of the genetic algorithm;
- the generic simulator, collaborating with the genetic algorithm;
- hardware and software tools, applied for the realization of the macro granularly parallel operation.

The most important open source code software tools, used for the development of the genetic algorithm were the followings:

- fox toolkit: (http://www.fox-toolkit.org);
- plplot: (http://www.plplot.org);
- tclap: (http://tclap.sourceforge.net/);
- c++ compilers: gcc mingw32.

For the macro granularly parallel execution of the evolutionary simulations, a PC cluster, containing 16 units was built. The operation of the cluster was solved by the adaptation of the Open SSI (http//www.openssi.org) software.

Regarding the demonstrated example applications:

– the programs of the benchmark test tasks has been written by myself;

– the simulation of the detailed example applications has been solved by the generic simulator based on the direct computer mapping of processes, developed in the research school of process informatics, using the version running under Windows® with EXCEL® interface.

In the solution of the various practical problems and experimental case studies from the fields of identification, optimal control, optimal design and optimal scheduling in the past decades, the continuously developing genetic algorithm collaborated also with the various adaptations of the generic process simulator, based on Direct Computer Mapping.

3 RESULTS AND DISCUSSION

The methods, applied for the economic optimization of complex systems in practical problem solving, have to satisfy many criteria. One of the two most important demands is supporting of the multicriteria evaluation in decision making. The other is, the capability for the representation of the complex possibility spaces, characterizing the economic and/or technological processes.

In the development of the genetic algorithm, prepared for the multicriteria economic optimization, was motivated by the above criteria.

3.1 Development of a complex genetic coding

The representation of complex systems by structure lattices is based on the works of Blickle. The essential feature of the corresponding theory is that an actual set of complex objects (systems) can be described by properties, classified into equivalency classes. The individual property classes correspond to the fundamental properties, characterizing the objects. The lattice is determined by the following principle: the objects can be determined by property combinations, consisting of one, and only one property from each equivalency class. Of course, a well defined part of the pair wise combination is forbidden. Having defined this pair wise incompatibility relation, we can generate all of the compatible

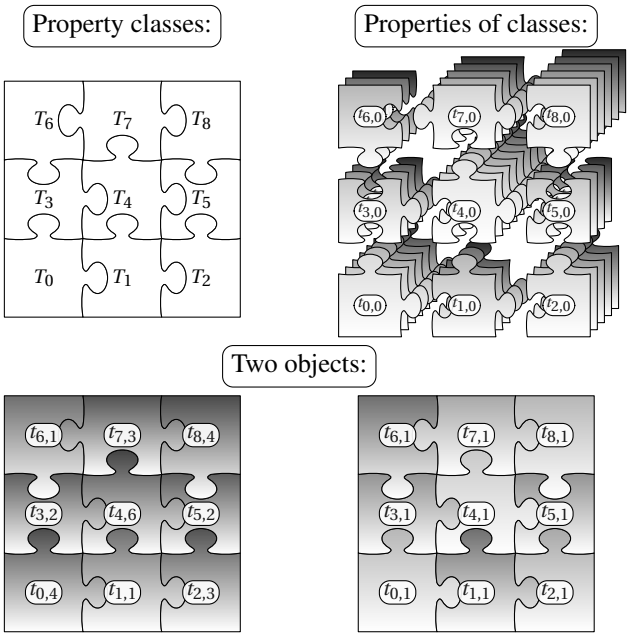


Figure 1. Representation of complex systems by structure lattices.The form of incompatibility relation is $I = \{(t_{0,1}, t_{3,2}), (t_{2,0}, t_{5,3}), (t_{8,0}, t_{7,4})\}$, where $t_{i,j}$ is the j-th property of the i-th class.

property combinations automatically. The essential features of the structure lattice are illustrated in Fig. 1.

From the viewpoint of genetic coding, the possible codes are determined by the ensemble property classes. The individual genes of the chromosome correspond to the respective property classes. The locus of the gene refers to the index of the corresponding property class. The possible alleles of the genes can be described by the properties of the given property class. At the same time, we can declare all of the incompatible pairs of the alleles. The respective genetic coding can be seen in Fig. 2.

When introducing the genetic algorithm, I emphasized that they can be applied for the coding of discrete and continuous characteristics, as well as of full permutations. In describing the complex system, the genes (i.e. the property classes) can be discrete (e.g. type of process units), continuous (e.g. economic or technical parameters) or permuted discrete variables (e.g. production sequences). Accordingly, the description of these complex systems needs the simultaneous

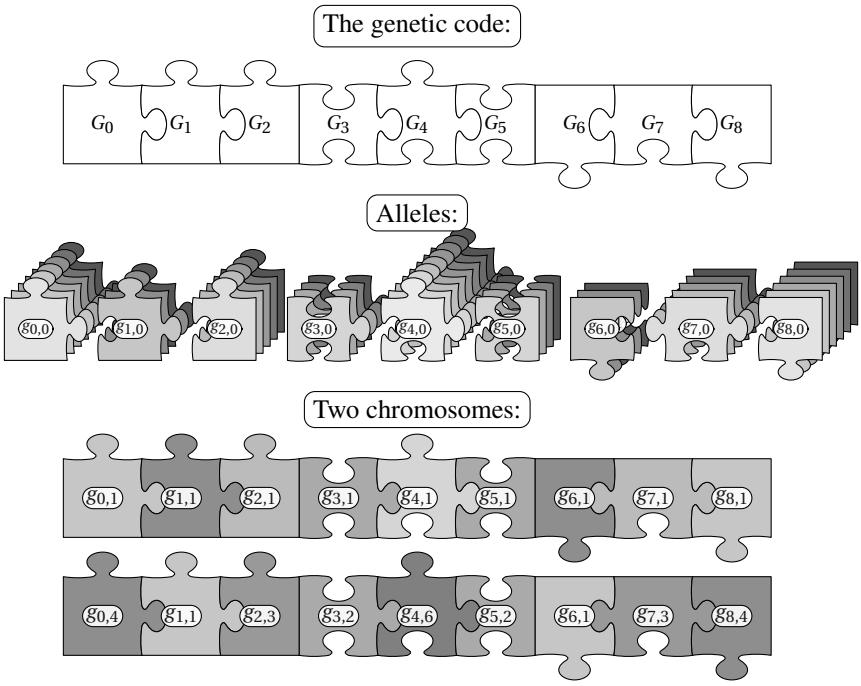


Figure 2. Genetic code of the complex system using property classes. The incompatible alleles of the $g_{i,j}$ allele characterized by the $I_{i,j} = \{g_{l,m}, g_{k,n}\}$.

use of these various gene types. Considering the representation of real numbers in the computer, all of the continuous properties are transformed to discrete values. In addition, the continuous parameters, characterizing the practical problems, have also a given accuracy (e.g. granularity). Accordingly, all of the gene types can be described by discretized values in genetic coding. The essential difference between the continuous and discrete genetic operators is whether they can follow the Euclidean or the Hammingian metrics. Finally, the continuous and discrete cases can be distinguished by the appropriate metrics.

Having determined the gene types, we can declare the possible alleles of them. In the case of the discrete genes, it can be solved by an enumeration, while for the continuous characteristics we can give at least the lower and upper bounds, as well as the refinement of the discrete steps. In this work, these properties were supplied by the lower and upper limits, awaited from the experts. In this genetic coding, the elements of the ith gene type, are characterized by the $D_i = \{(g_{i,1}, I_{i,1}, K_{i,1}), (g_{i,2}, I_{i,2}, K_{i,1}), \ldots, (g_{i,n}, I_{i,n}, K_{i,n})\}$

triplet, where $g_{i,j} \in \mathbb{N}$ that in the majority of the cases corresponds to the ordinal number of the element in the "catalogue" of the algorithm, counting of the evaluating point of views. The $I_{i,j} = \{g_{l,m}, \ldots, g_{k,v}\}$ corresponds to the set of the incompatible alleles, while $K_{i,j}$ determines the later defined gene type of the chromosome, corresponding to the further elements of the "catalogue" (e.g. the parameters or changeable parts of the process unit). The continuous gene types are determined by the $C_i = \left(l_i, h_i, e_i^l, e_i^h, d_i^{max}, d_i^{min}\right)$ sextuplets, where l_i, h_i and e_i^l, e_i^h determine the lower and upper bounds of the domain, A and B describe the bounds of the awaited interval, while $d_i^{max}$and d_i^{min} determine the maximal and minimal discretization of the interval.

The two above gene types are supplemented by three combined gene type. This corresponds to the discrete gene sequences, to permutable gene sequences, as well as to the chromosome. The gene sequences are defined as the ordered sets of the discrete or discrete permutable genes, while the elements of the chromosome may be all of the above defined simple or complex genes. The introduction of complex genes is motivated of the fact, that the initialization, recombination and the mutation of the full permutations can be treated only with the knowledge of the whole ordered permutable gene sequences. On the other hand, the complex genes, having recursively embedded, can form also hierarchical data structures. The solution is conform with the later described efficiency increasing methods. The above defined genetic elements are visualized in Fig. 3.

The extension of genetic coding makes possible the respective extension of the algorithm, too. The elements of the algorithm, elaborating the code, are the initialization, as well as the recombination and mutation operators. These activities are defined for all of the gene types, according to their properties. Accordingly, the continuous, discrete or permutable types can use the corresponding recombination and mutation operators, described in the literature, while the initialization is carried out mostly by random generation.

In the solution of a given task, the genetic coding corresponds one of the above defined gene types. Having generated the alleles, we have to describe the incompatibility relations for the discrete genes. All of the gene types have

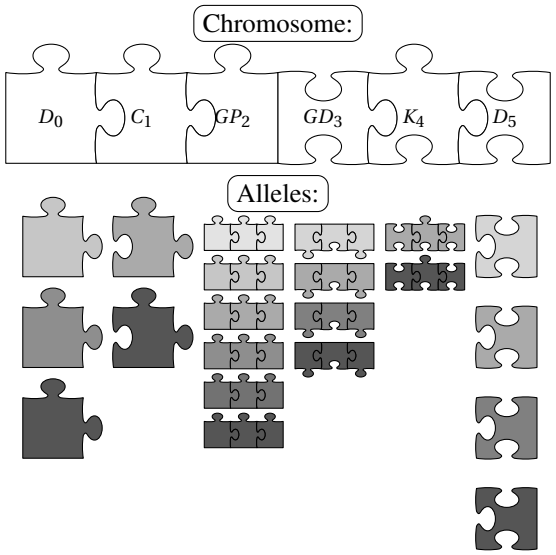


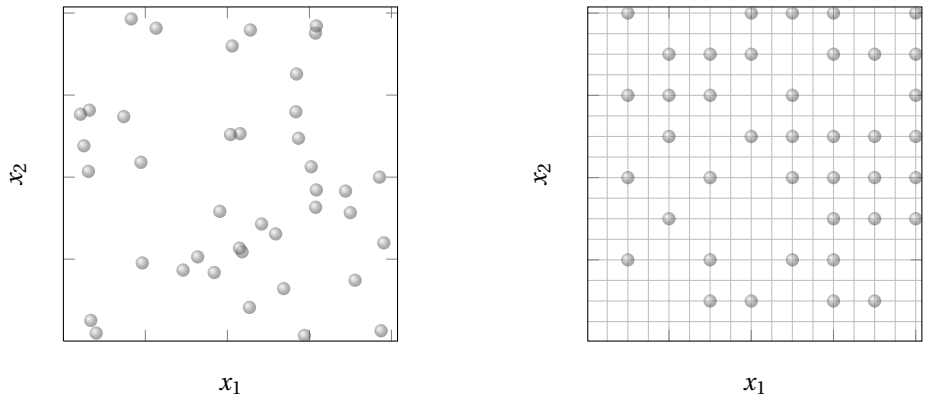
Figure 3. The complex and the simple gene types.

default initialization, recombination and mutation operators, however, we can replace them for another embedded one.

3.2 Elaboration of a grid method

The necessity for the use of small population size causes many problems to be solved. One of them is that the solution is very sensitive for the "quality" of the initial population. The other is that the diversity of the population can decrease rapidly. The random number generators produce really uniform random numbers only for larger population size, while in the case of small populations the distribution of the variance is not equal in the different segment of the search space. In the case of the continuous gene types, the generation of random numbers can be controlled by the awaited limits e_i^l, e_i^h, estimated by the expert. In addition, with the knowledge of the (d_i^{max}) maximal discretization of the genes (given in the initialization), the values, suggested by the random number generator, can be rounded for the respective preciosity. This grid, generated in the search space, supports the more uniform distribution of the initial population. The method is illustrated in Fig. 4.

Another useful application of the grid is the filtering of the identical variants.



a) Random population without grid method. b) Random population with grid method.

Figure 4. The modified initialization.

Without the grid the equality of two variants cannot be identified simply because of the floating point number representation. However, the grid supports the control of the minimal distance between a given variant, and the neighboring ones.

The developed complex genetic coding effectively supports the description of the search space for the genetic algorithm. However, comparing with the simple string-like and homogeneous chromosome coding, it is disadvantageous, that the evaluation of the similarity measures is more difficult. By means of the Euclidean and Hammingian distances, it is not possible to define a consistent similarity measures for the optionally hierarchical combination of continuous and discrete gene sequences. This situation limits the application of the methods, used for the conservation of the diversity in the population. The introduced grid method makes possible another new solution to converse the diversity. Accordingly, the new variants are to be generated only into the grid points, while having modified the grid size, we can control the equilibrium between the exploration and exploitation that controls also the diversity. For this solution, I have modified the recombination and mutation operators appropriately.

The modification of the operators for the continuous genes is similar to the method, applied for the initialization, i.e. the values produced by the conventional operators, are rounded to the nearest grid points. In the case of

discrete genes, a simplified grid method has been realized. Accordingly the gene sequences are associated with an integer, that less then the half of the sequence length that determines the awaited minimal Hammingian distance, between the parent and child gene sequences in the execution of the crossover and mutation. This solution is similar to the consideration of precision for continuous genes. The algorithm of the operator should not be modified, only the Hammingian distance between the gene sequences is compared with the awaited value, and the operators are repeatedly applied until either the appropriate child will have appeared, or the number of repetitions reaches the prescribed upper bound. This upper bound is the function of the length of the given sequence, and its default value is the half of the sequence.

It is advantageous, that the introduced development does not change the conventional algorithm of the operators, only the results are modified. Accordingly, in a marginal case the precision and/or the awaited minimal distance can be set to zero, and in such a way the operators behave according to the conventional situation. Starting with a rough grid and refining the grid along the subsequent generations, we can avoid the too early convergence. The critical element of the method is the scheduling of the changing decomposition of the grid. In the case of a single objective function, the adaptive scheduling has been proposed that follows the change of the actually best solution. In the case of the multicriteria evaluation, only the 'a priori', deterministic scheduling can be applied. According to the experiences, with the use of an exponential decrease, to get the awaited preciosity at the 75% of the planned generations, seems to be an effective and safe scheduling strategy.

3.3 Management of the multicriteria evaluation

Supporting of the multicriteria decisions has a keynote role in the elaboration of the new algorithm. In the solution I made possible fitting the preferences of the decision maker by each of the 'a priori', interactive and 'a posteriori' methods. In the declaration of the evaluating criteria, we can define the optional properties. These properties are the priority, as well as the apparent lower bounds of the given objective. By means of these properties the decision maker can guide the

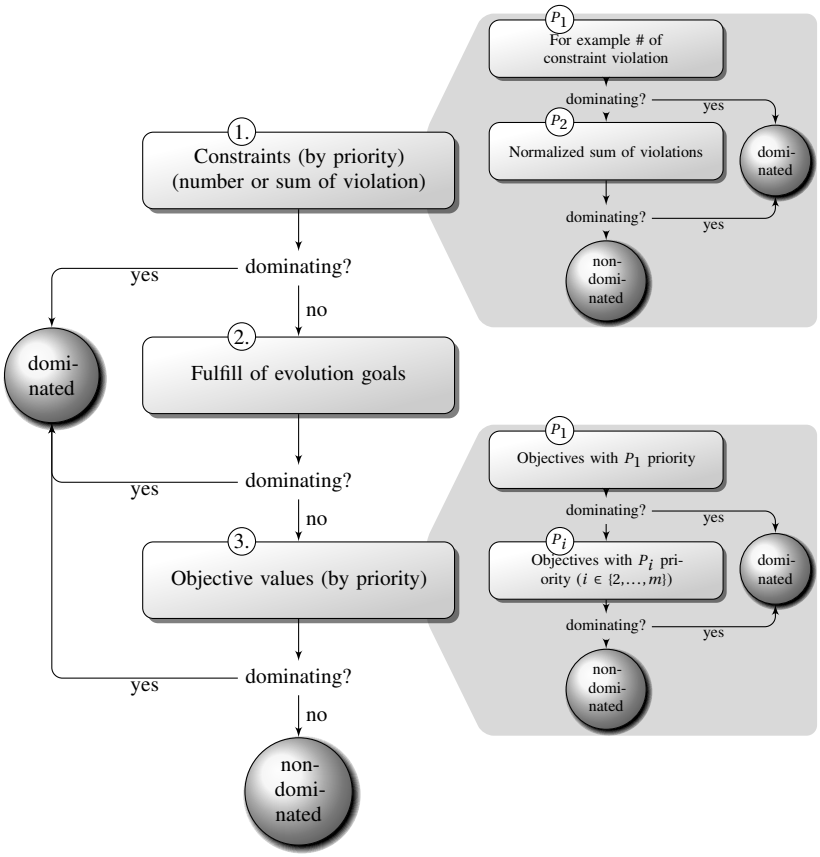


Figure 5. The calculation of modified Pareto-dominance.

optimization process to the preferred regions of the Pareto-front.

The flexible possibility of the unified consideration of constraints, priorities and objectives, makes possible the application of a ranked fitness calculation, according to a modified Pareto-dominance in the multicriteria genetic algorithm. The ranking can be used both for a single objective and in the multicriteria evaluation with the possibility of the treatment of conditions. There are various strategies for the fitness calculation, and the majority of the methods, published in the literature can be applied. For example we can use the number of the dominated variants or the number of the variants, dominating a given variant, as well as the depth of the Pareto-front. In the calculation of the dominance, first the conditions are evaluated one after the other. Optionally, a priority ranking can be defined for the conditions. This makes possible the combined use of

the condition violations (i.e. how many conditions are violated by the given variant), followed by the consideration of the summarized or maximal measure of constrain violations (i.e. how long is the distance from the awaited range).

The evaluating criteria have a priority ranking, too. In a given priority group, first the fulfillment of the objectives are compared (i.e. the variants that fulfill the objectives are better). Next, the variants that fulfill the objectives are compared according to the value of the criteria. The procedure of the calculation is illustrated in Fig. 5.

The necessary conditions for the good estimation of the Pareto-front are that the proposed solutions have a uniform distribution along the front. Therefore, in addition to the fitness values, the variants are characterized by a crowding parameter. There are many options for these crowding parameters. In the present work, the multiplication of the distance from the k nearest neighbors was used as a crowding parameter, where k was the number of the evaluating criteria.

In the applied selection algorithm, this parameter was used for the selection and replacement of the comparison of the parents for the variant, having the identical fitness value. This helps the uniform distributed identification of the Pareto-front.

Because of the small population size, the number of the non-dominated solutions in the last population is also small. Consequently, it is advantageous to save all of the known Pareto-optimal solutions in outer archive storage. The size of the archives can be configured freely. Having reached the maximal size of the archives, the saved Pareto-optimal variants are deleted, according to their crowding parameters.

3.4 Implementation properties

Because of the high computational demand of the evaluations, usually only a couple of thousand or ten thousand variants can be calculated in problem solving. Nowadays, the usually applied computers make possible the fast enough storage of genetic code with a rapid enough treating of them, without slowing down the execution of the algorithm.

Consequently, I made possible the storage for all of the evaluated individuals. The role of this storage is similar to the well known tabu list. The storage helps to avoid the repeated evaluation of the variants. Accordingly, the modules of the algorithm, responsible for the generation of the new variants (initialization, recombination and mutation), are prepared for the investigation, whether the new genetic code participate in the storage of the evaluated ones. In the case of a possible repetition, the operators try to generate a new genetic code. The number of the repeated generations is limited, while the default value is the half of the number of simple genes in the genetic coding. Having arranged this limit value, the usual recombination and mutation are replaced by an extra mutation. This extra mutation is executed for each gene with the probability of 0.5, according to the respective gene type.

Constrains are treated in two levels in the method. In the first level the incompatibility relations declared for the discrete gene types, are considered in the initialization, recombination and mutation. The forbidden combinations are managed by the method, describe for the tabu list. In the second level, constrains are taken into consideration in the calculation of the extended Pareto-dominance.

The developed genetic coding is conform with the object oriented programming paradigm. The object oriented design and implementation of the program made possible the optional runtime change of the program components. This characteristic helps in the adaptation of the algorithm for the solutions of various practical problems, as well as in the testing of the operators, fitness calculations, and crowding parameters.

The highest class of the object hierarchy is responsible for the calculation of the evaluating criteria. The children of this class are the most frequently used classes, communicating with the collaborating simulator, as well as the classes executing the various test and benchmark problems. In the former version, the EXCEL®, as a DDE server was used for the communication with the simulator, running under Windows. The new version uses a platform independent communication protocol. This protocol supports the parallel work with multiple simulators in sense of server/client model. The main functionalities of this protocol are the configuration of the source space, the configuration of

1. Algorithm: The pseudocode of the new algorithm.

Input: cmp_F: the modified Pareto-dominance operator
Input: toGrid: the round to grid function
Input: s, a: the population and archive size
Input: r, m: the recombination and mutation factor
Data: t: the generation counter
Data: Pop: the population
Data: Arc: the archive with the actual best individuals
Data: Par: container with the actual parents
Data: Off: container with the new children
Data: v: fitness access function
Output: $X^\star$: the optimum set

```
1  begin
2      X ⟵ createDecisionSpace(𝕏)                    //create the decision space
3      Y ⟵ createObjectiveSpace(𝕐)                   //create the objective space
4      C ⟵ createConstraintSpace(ℂ)                  //create the constraint space
5      t ⟵ 0
6      Arc ⟵ ∅
7      Pop ⟵ toGrid(createPop(s, X))                 //initialize the population
8      Pop ⟵ evaluateIndividuals(Pop, X, Y, C)                          //evaluate
9      v ⟵ assignFitness(Pop, Arc, cmp_F)            //the Pareto ranking function
10     while ¬terminationCriterion() do
11         Arc ⟵ updateOptimalSetN(Arc, Pop)          //update the archive
12         Arc ⟵ pruneOptimalSet(Arc, a)              //pruning the archive
13         X ⟵ updateGrid(X, t)                       //update the grid
14         Par ⟵ select(Pop, Arc, v, s)
15         for i ⟵ 0 up to |Par| − 2 do
16             if rand_u() ≤ r then Off[i,i+1] ⟵ toGrid(recombine(Par[i], Par[i+1]))
17             if rand_u() ≤ m then Off[i,i+1] ⟵ toGrid(mutate(Off[i,i+1]))
18         Off ⟵ evaluateIndividuals(Off, X, Y, C)                          //evaluate
19         v ⟵ assignFitness(Pop, Arc, cmp_F)        //the Pareto ranking function
20         Pop ⟵ reproducePop(Par, Off, Pop, v)      // reproduce the population
21         t ⟵ t + 1
22     return extractOptimalSet(Pop ∪ Arc)
23 end
```

the evaluating criteria, the sending of the code of variants to be evaluated, as well as the receiving of the evaluations.

The simple pseudo-code of the above describe genetic algorithm is describe in algorithm 1.

4 CONCLUSIONS

Considering the above described research results, as well as the experiences, obtained from the application of the continuously developing genetic algorithm, the most important conclusions are the followings:

The elaborated complex genetic coding can effectively be applied for the common representation of discrete and continuous genes, while the ‘a priory determination of the incompatibility relations helps to avid the elaboration specific individual genetic coding for the various complex tasks.

The flexibility and efficiency of the complex genetic coding and of the respective extended set of operators are proven by illustrating examples and test problems. In addition, the elaborated methods were successfully applied for the identification and model based optimal design, control and scheduling of various complex practical problems with hybrid models.

The synergic application of the elaborated complex coding and the grid method can effectively be applied for description of the genes for the continuous property classes. The developed methodology gives possibility to utilize the heuristic knowledge of the expert in the description of the search space. Simultaneously, another advantages feature is, that having recognize the solutions in the vicinity of the prescribe bounds, the algorithm automatically extent the search space.

The elaborated grid method supports the adequate initialization and effective run of the genetic algorithm, also in the case of small population size and of small generation number. This makes possible to increase the computational demand of the evaluation, consequently the application of the possibly most detailed models in the solution of many practical problems. This is very important, because the applicability of the economic and technical optimizations is determined

by the exhaustiveness. Accordingly, the less number of calculations with a more detailed model is preferred to the more number of evaluations with a simplified one. The new method elaborated for the treatment of Pareto-dominance, contributes to the correct and powerful solution of the multicriteria problems. The new developments help the adequate choice and ranking of the constraints and evaluations, as well as the evolution of the multicriteria good enough solution. In the solution of the practical problems, the priority ranking of the constraints and evaluations combined with the new grid method, help to focus on the very part of the Pareto-front, where the good solutions are awaited. In practical applications to determine the evaluations, the expert often tends intuitively or consciously to define objectives of approximately identical importance. Consequently, the solutions with almost commeasurable values used to be preferred.

One of the lessons, coming from the practical resolved optimization and identification problems was that it is not possible or it does not worth to aggregate the evaluation into a single objective function. In optimization, almost everybody wants to make an economic evaluation (i.e. minimizing the cost or maximizing the profit), however the data for the calculation of the economic goal function are not known. It is a typical case, when we have to optimize one, by-itself also complex part of a technological process, consisting of many steps. The economic parameters of the input and output materials are often not known. Consequently, the study ought to be extended to a greater system consisting of this part. On the other hand, the field experts can declare very good natural objective functions. Nevertheless, on ongoing methodological development tends to bridge the existing gap between the technological and economical processes, and this makes possible the more and more correct economic evaluation. In accordance with the results, obtained from the logistical example of the present work, the combined application of the economic and natural evaluations seems to be a feasible method, temporarily.

The experiences, obtained with the continuously developing and presently further developed, integrated genetic algorithm, proved that the applied coding and operators, as well as the archived storage of the investigated variants support

the optimization process with small population and generation number.

The optimization of the practical tasks with great computational demand for the evaluation can be solved by the macro-granularly parallel simulation and evaluation of the variants. The method, implemented in a PC cluster, can accelerate the genetically controlled evolutionary process almost proportionally with the number of CPUs of the cluster.

5 NEW SCIENTIFIC RESULTS

The new scientific results of my work are the followings:

1. I have been developed a new complex genetic coding, based on the structure lattice for the combination of optionally hierarchic, discrete/-continuous and permutable gene sequences. The structure lattice makes possible the 'a priory' definition of the optional incompatibility relations between the discrete properties, classified into equivalency classes. The developed coding supports the uniform treatment of the discrete and continuous property classes. The alleles of the continuous property classes are described by automatically generated discrete elements, determined by the prescribed composition of the given domain. Within a domain, the user can define the lower and upper bounds of the awaited subinterval. The new coding supports the hierarchic coding, determined by the tree structure of the combined genes. Also the method makes possible the application, automatic recognition and treatment of the full permutations. I have introduced new, extended genetic operators for the initialization, recombination and mutation, which automatically consider the type of the given gene sequence within the scope of the previously described coding. I have extended the genetic algorithm with a new, global operator, which with the knowledge of the evaluations automatically decreases the great decomposition of the properties for both the continuous and discrete gene sequences.

2. I have elaborated a user-friendly, complex method for the calculation of the Pareto-dominance that contains the published methods for the

treatment of the constraints and evaluations, as full compatible special cases of the Pareto evaluation. The implementation of constraints and evaluations supports the interchangeable equal use of them. In the case of the goal determining constraints, the increased efficiency is supported also by the integrated description and execution of the constraints, together with the evaluations. In addition, optional priority can be defined for both the constraints and the evaluating objectives.

3. I have developed a new, platform independent macro-granularly parallelizable solution for the organization of the communication between the genetic algorithm and an optional simulator, which calculates and evaluates the proposed solution, according to the optionally multiple objectives.

4. I have successfully applied the genetic algorithm of above 1-3. theses (optionally combine with an appropriate implemented generic simulator) for the solution of 11 industrial economical and technological problem, 8 economical and technological problem in experimental phase. The algorithm was applied in 8 MSc theses, in one successfully defended and two ongoing PhD theses, as well as in the preparation of many educational demonstration programs (e.g. identification of metabolic networks, planning of cultivation process, design, planning and scheduling a multi-product batch plant.

6 RECOMMENDATIONS

Based on discussion of the results my recommendations are the followings.

The consolidated software implementation of the genetic algorithm, developed in the present research work, ought to be applied in the ongoing and planned solution of the various practical problems. Accordingly, the methodology has to be utilized in the research project of the Department of Information Technology, as well as of the collaborating University environment. Considering

the experiences, I plan to refine the embedded methodology. Simultaneously, the application will intensify the publication activity.

The development of the methodology has been motivated by the theoretical and practical demands for the solution of various technological and scientific problems. One of the most important and actual challenge is the utilization of the results in the model based economic optimization of the various complex processes. Another important task is to combine the methodology with other software tools, applied for the design and control of economic processes. Considering the activity of Faculty of Economic Sciences and Doctoral School of Economic and Regional Sciences at the Kaposvár University, I shall make additional efforts toward further collaboration in economic applications.

From the field specific professional point of views, the most exciting research goal is to develop powerful methods for the interactive and automatic analysis and control of the Pareto-fronts. In addition to the optimization task, there may be interesting problems to be solved also from the field of the decision support systems.

7 PUBLICATIONS IN THE FIELD OF THE DISSERTATION

Scientific publications in foreign language

Aranyi, A., Temesvári, K., Csukás, B. and Balogh, S. (1998). Computer assisted design of industrial scale chromatographic separation. *SPICA Strassbourg*, 152–157.

Balogh, S. (2009). Multicriteria decision support by genetic algorithm. *Regional and Business Studies (under submitting)*.

Csukas, B. and Balogh, S. (2001). Evolutionary synthesis of almost closed conservational processes. *Gani and S. B. Jorgensen Eds., European Symposium on Computer Aided Process Engineering , Computer Aided Process Engineering, Elsevier 9*, 381–386.

Csukás, B. and Balogh, S. (1996). Combining genetic programming with generic simulation models in evolutionary synthesis. *In, Bertrand, Jafari, Fransoo, Rutten Eds*, 158–172.

Csukás, B. and Balogh, S. (1998). Combining genetic programming with generic simulation models in evolutionary synthesis. *Computers in Industry 36*, 181–197.

Csukás, B., Balogh, S. and Bánkuti, G. (2005). Generic bi-layered net model – general software for simulation of hybrid processes, in, daoliang li and Íbaoji wang eds., artificial intelligence applications and innovations ii. 2nd ifip conference of tc12 wg 12.5, springer. *pp.*, 700–710.

Csukás, B., Balogh, S., Kováts, S., Aranyi, A., Kocsis, Z. and Bartha, L. (1999). Process design by controlled simulation of the executable structural models. *Comput. Chem. Engng. 23*, 569–572.

Csukás, B., Balogh, S. and Structures, E. (1997). Evaluation feedback between the generic simulation and the genetic synthesis. *Joint Conf. of Information Systems*, 1–5.

Csukás, B., Lakner, R., Varga, K. and Balogh, S. (1996). Combining genetic programming with generic simulation models in evolutionary synthesis. *Comput. Chem. Engng. 20*, 61–66.

Temesvári, K., Aranyi, A., Balogh, S., Bánkuti, G. and Csukás, B. (2005). Computer-aided process design of the separation of a two-component steroid mixture by simulated moving bed technique. *J. Ind. Chem. Hung. 32*, 5–12.

Temesvári, K., Aranyi, A., Balogh, S. and Csukás, B. (2004). Simulated moving bed separation of a two components steroid mixture.

Scientific publications in Hungarian language

Balogh, S., Csukás, B., Bartha, L., Kocsis, Z. and Kis, G. (2000). Szakaszos polimerizációs recept fejlesztése genetikus algoritmussal összekapcsolt dinamikus szimulátorral. *Műszaki Kémiai Napok' Veszprém*, 36–40.

Balogh, S., Csukás, B., Sógor, A., Budai, M. and Miklósi, M. (2004). Egy soktermékes üzemcsarnok oldószer szennyeződésének szimulációs vizsgálata. *Acta Agraria Kaposvariensis 8*(3).

Balogh, S., Csukás, B., Takátsy, T. and Tari, C. (2000). Többérdekű logisztikai láncok fejlesztése genetikus algoritmussal összekapcsolt dinamikus szimulátorral. *Műszaki Kémiai Napok' Veszprém*, 41–45.

Balogh, S., Négyesi, G., Budai, M., Sógor, A. and Csukás, B. (2003). Szakaszos üzemcsarnok légtér szennyezettségének mérése és dinamikus szimulációja. *Műszaki Kémiai Napok' Veszprém*, 284–290.

Boity, O., Gudlin, G., Tari, C., Balogh, S., Csukás, B. and Takátsy, T. (2001). Kisérlet egy farmgazdálkodást segítő genetikus algoritmussal fejlesztett szimulátor kialakítására. *Műszaki Kémiai Napok' Veszprém*, 242–247.

Csukás, B. and Balogh, S. (2001). Egy konfigurálható, generikus, dinamikus szimulátor és újabb alkalmazási lehetőségei. *Műszaki Kémiai Napok' Veszprém*, 45–50.

Kis, G., Csukás, B., Bartha, L., Kocsis, Z. and Balogh, S. (2000). Szakaszos polimerizációs művelet számítógéppel segített recept fejlesztése. *Műszaki Kémiai Napok' Veszprém*, 29–51.

Lukács, A., Takátsy, T., Csukás, B. and Balogh, S. (2001). Baromfiistálló energetikai és makroszintű metabolikus szimulációjának tapasztalatai. *Műszaki Kémiai Napok' Veszprém*, 248–253.

Takátsy, T., Csukás, B. and Balogh, S. (2000). Az állat és környezete kapcsolatának dinamikus szimulációja. *Műszaki Kémiai Napok' Veszprém*, 22–28.

Temesvári, K., Aranyi, A., Balogh, S., Bánkuti, G. and Csukás, B. (2004). Kétkomponensű szteroid elegy szimulált mozgó ágyas (smb) elválasztásának számítógéppel segített tervezése. *Acta Agraria Kaposvariensis 8*(3).

Presentations

Aranyi, A., Csukás, B., Temesvári, K. and Balogh, S. (1997). Structural model based dynamic simulation of preparative hplc. *International Symposium on Chromatography, Balatonszéplak, September*, 3–5.

Balogh, S. (2006). A genetikus algoritmussal kapcsolt generikus szimulátor szoftver implementációjának fejlesztése. *V. Alkalmazott Informatika Konferencia*.

Balogh, S. and Csukás, B. (1995). A makroszintű modell, mint a mikro szintű modell genetikus kódja - a nem string típusú genetikus kód lehetőségei és korlátai. *Műszaki Kémiai Napok'*, 89–90.

Balogh, S. and Csukás, B. (1996). Számítógéppel segített folyamattervezés a részletes modellel generált és értékelt genetikus algoritmussal. *Műszaki Kémiai Napok'*, 37–38.

Balogh, S. and Csukás, B. (1997). Esettanulmány dinamikus szimuláció és genetikus algoritmus összekapcsolására. *Műszaki Kémiai Napok'*, 141–142.

Balogh, S. and Csukás, B. (1998). A genetikus programozás lehetőségei a folyamatmérnöki munkában. *Műszaki Kémiai Napok'*, 5–6.

Balogh, S. and Csukás, B. (2001). A generikus szimulátorral visszacsatolt kapcsolatban működő genetikus algoritmus és újabb alkalmazási lehetőségei. *Műszaki Kémiai Napok' Veszprém*, 206–209.

Balogh, S., Lakner, P. R. and Csukás, B. (1994). Többszempontú genetikus algoritmusok vizsgálata. *Műszaki Kémiai Napok'*, 26–28.

Balogh, S., Négyesi, G., Budai, M., Sógor, A. and Csukás, B. (2003). Szakaszos üzemcsarnok légtér szennyezettségének mérése és dinamikus szimulációja. *Műszaki Kémiai Napok' Veszprém*, 284–290.

Barthó, I., Sinkó, B. I., Hantos, G., Balogh, S., Csukás, B. and Varga, M. (2006). *Bioreaktor modelljének identifikálása és egy biokonverziós folyamat modell bázisú fejlesztése*. Kaposvár május 26: V. Alkalmazott Informatika Konferencia.

Csukás, B. and Balogh, S. (1998). Megmaradási folyamatok strukturális modelljének közvetlen leképezése végrehajtható program értékű adatbázisra. *Műszaki Kémiai Napok'*, 54–55.

Csukás, B., Balogh, S., Takátsy, T., Bóity, O., Guldin, G. and Tari, Cs. (2001). Mérnöki logisztika az üzemirányításban. *MTA Agrárműszaki Bizottságának Tanácskozása, Gödöllő január*, 23–25.

Csukás, B., Debelak, K. A., Prokop, A., Balcarcel, R. R., Tanner, R. D., Bánkuti, G. and Balogh, S. (2003). *Generic Bi-layered Net Model Based Discrimination of Chemical and Biological Warfare Agents, AIChE Annual Meeting, San Francisco, November 16-20*. Manuscript 474f.

Csukás, B., Kováts, S., Aranyi, A., Temesvári, T.-K. and Balogh, S. (1997). A valódi és szimulált mozgó ágyas folyamatos üzemű preparatív kromatográfia szimulációjának tapasztalatai. *Műszaki Kémiai Napok'*, 100–101.

Domonkos, D., Könczöl, K., Balogh, S., Csukás, B. and Varga, M. (2006). Rekombináns fehérje szintézis számítógépi modellen alapuló fejlesztésének lehetőségei.

Katona, A., Balogh, S. and Csukás, B. (2006). A generikus kétrétegű háló modell fpga bázisú hardver implementációjának lehetőségei.

Lehőcz, G., Balogh, S., Bánkuti, G. and Csukás, B. (2005). Gazdasági potenciál számításon alapuló lokális döntéseket támogató algoritmusok fejlesztése. *Informatika a Felsőoktatásban'Konferencia, Debrecen*, 24–26.

Nagy, K., Csukás, B., Kis, G., Bartha, L. and Balogh, S. (2001). Study on preparation and properties of olefin-maleic-anhydride copolymers. 40th international petroleum conference. *September*, 17–19.

Nagy, K., Kis, G., Bartha, L., Csukás, B. and Balogh, S. (2001). Olefin - maleinsavanhidrid kopoli-merek előállítási körülményeinek és tulajdonságainak vizsgálata. *Műszaki Kémiai Napok' Veszprém*, 24–26.

Takátsy, T., Csukás, B., Balogh, S. and I., L. A. (2001). Az állati metabolizmus makroszintű dinamikus szimulációja mérnöki alkalmazásokra. *MTA Agrármű-szaki Bizottságának Tanácskozása, Gödöllő január*, 23–25.

Temesvári, K., Aranyi, A., Csukás, B. and Balogh, S. (2001). Kisérletek és szimulációs vizsgálatok egy királis elválasztás szimulált mozgó ágyas megvalósíthatóságának elemzéséhez. *Műszaki Kémiai Napok' Veszprém*, 24–26.

Temesvári, K., Aranyi, A., Csukás, B. and Balogh, S. (2003). Simulated moving bed separation of a two components steroid mixture. *International Symposium on Chromatography, Balatonszéplak, September*, 4–6.

Temesvári, T.-K., Csukás, B. and Aranyi, A. (1997). Determination of the equilibrium, hydrodynamic and kinetic parameters for the structural modeling of preparative hplc. *International Symposium on Chromatography, Balatonszéplak, September*, 3–5.

Varga, M., Bíró, B. B., Kitanics, B. T., Bánkuti, G. and Csukás, B. (2005). Vállalkozók adózási stratégiáinak szimulációja generikus kétrétegű háló modellel. *Informatika a Felsőoktatásban'Konferencia, Debrecen*, 24–26.

Veizer, A., Bánkuti, G., Balogh, S. and Csukás, B. (2005). Metabolikus hálózatok generikus kétrétegű háló modelljének identifikálása. *Informatika a Felsőoktatásban'Konferencia, Debrecen*, 24–26.