A frame-based approach for building an acceptable college training schedule based on the artificial bee colony algorithm

O Ponomarenko

College of Engineering and Management of the National Aviation University,
5a Metrobudivska st., Kyiv 03065, Ukraine

E-mail: PonomarenkoS_200@ukr.net

Abstract. Using the example of solving the problem of drawing up an acceptable schedule of college classes, the article explores ways that increase the efficiency of the honeybee algorithm, which belong to the class of heuristic iterative algorithms for nondeterministic search, in particular, game methods for finding solutions in which rational strategies are found that reduce the tree enumeration of possible solutions. The article proposes an approach that uses the concepts of the theory of frames, leading to a "conscious" process of finding solutions, taking into account the experience (acquired "knowledge") obtained in previous iterations of the honeybee algorithm, which consists in modernizing the search scenario, which contributes not only to the optimization evolution of the result, but also the effective evolution of the search itself. Supporting the concept of the classical honeybee algorithm, on the basis of a convincing argument for the naturalness of using the classical theory of random search, the article proposes a frame scenario for an adventurous and reckless random search for "scout bees" and a frame scenario with a calculatingly careful search that, in fact, imitates the method trial and error for foraging bees. For the considered scenarios of algorithms, the boundaries of their correct applicability are indicated. For the indicated frame-scenarios, the parametric adaptability of the algorithm is considered. The theoretical aspects of constructing frame scenarios from the point of view of the general theory of frames are also analyzed.

1. Introduction

Research into the theoretical and applied aspects of bioinspired algorithms has led to the study of one of the new directions, combining multi-agent methods of intelligent search, which are based on the modeling of collective intelligence. It is customary to combine these methods into a class of heuristic iterative methods of non-deterministic search. One of the significant places in this class is occupied by the methods of swarm intelligence, the idea of which is to unite a set of simple agents into a certain "family" community and create laws of collective behavior in a space that does not imply the presence of a subject of global control.

One of the modern multi-agent methods of intelligent search is considered to be the artificial bee colony method (ABC), which assumes, in particular, the presence of self-organization and adaptation [1, 2]. As a bench problem, on which the specific features of the artificial honeybee algorithm (ABC) were studied and continue to be investigated, the problem of scheduling training sessions in colleges was chosen, which, in addition to its relevance, is also interesting for its characteristics of multi-criteria and exponential time complexity, that is, the relation to the NP class - complex problems. A diagram of the process for solving this problem (Figure 1).

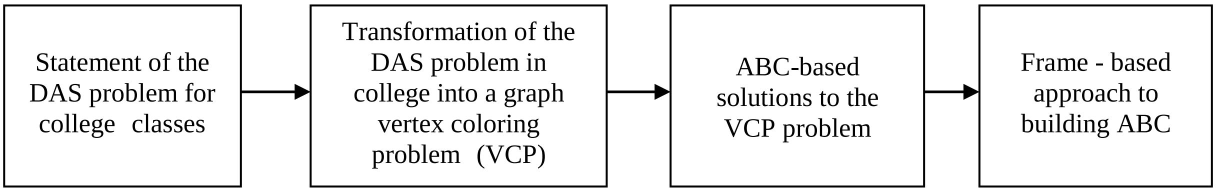


Figure 1. Diagram of the process for solving the problem of drawing up an admissible schedule of college training sessions (DAS).

The admissibility of the lesson schedule assumes its consistency with the previously formulated systems of restrictions and criteria.

The convenience of presenting the correctly formulated initial data of the DAS problem of training sessions in colleges by the network model and the determinism of the educational process in it argue the naturalness of the graph-theoretic approach in the process of solving this problem. Its transformation into an VCP problem involves the consideration of an undirected graph with single edges without loops, where each non-fictitious vertex of the graph G (B, P) corresponds to a lesson in a specific discipline, and the edges incident to some vertices determine the limiting conditions. Different coloring f:

$B \rightarrow N$ sets the ordering of the vertices by time (the time of a particular lesson). The problem of coloring the vertices of a graph, being a satisfactory model of the problem of composing the optimal schedule, preserves the NP-complexity. Indeed, on the one hand, the successful formulation of the VCP problem allows you to preserve even some specific details of the process of scheduling training sessions (for example, the specifics of an educational institution, its traditions), on the other hand, the exponential expansion of the search space for solutions dependent on the input data does not motivate the construction of exact and even approximate solution algorithms for the general case, especially taking into account the fact that no effective methods for determining the chromatic number of arbitrary graphs have been found to date, but in practice, when compiling a specific schedule, the order $|B|$ of the graph G (B, P) turns out to be extremely large for an admissible coloring for "reasonably perceived" time. The foregoing explains the prospect of studying both existing heuristic algorithms, in particular, swarm ones, and the development of new ones.

2. Materials and Methods

In the process of studying and researching the use of ABC in solving the VCP problem, it was found that the main idea of the algorithm that simulates the behavior of bees in the process of searching and collecting nectar in the natural environment is to determine the main segments of the search space and their surroundings, contributing to the diversity of populations of solutions at the next stages of the search. The foregoing explains the naturalness of the iterative nature of this algorithm. The efficiency of the bee algorithm can be increased by game methods of finding solutions, in which there are rational strategies that reduce the search tree, estimate the terminal vertices and contribute to the development of a reasonable move. But the simple application of various methods of constructing hierarchies of symbolic goals does not lead to a "conscious" process of finding solutions and does not improve the quality of information presentation. This, in turn, means that each new iteration of the search for solutions does not take into account the experience (knowledge) gained at the previous iteration, the accumulated experience “how to search”, and therefore makes it impossible for the algorithm to evolve in terms of search efficiency, leaving only the evolution of the result.

Following the ideas of the famous American scientist Marvin Minsky [3], the dominant goal of solving problems with the involvement of artificial intelligence (AI) is the process of "comprehending" the space of the problem and defining the concepts within which this problem is most easily solved. Wherein, the purpose of the search is to obtain information that contributes to the formation of these ideas, and not to find a solution.

Minsky's rule for the bee algorithm can be interpreted as follows: the use of ABC becomes valuable if the algorithm provides opportunities for improving the strategy at subsequent iterations. Thus, the minimax strategy used in the classical bee search algorithm is modified into a strategy for studying possible moves from each node Ui of the decision tree, which is formalized, for example, as follows:

$$OO(U_i) = F\left(O(U_k); O(U_{k+1}); \dots O(U_{k+1})\right),$$

where *OO* (U_i) is general assessment of the node U_i;

$U_k, U_{k+1} \dots U_{k+l}$ is possible moves from the node U_i, $k, l \in N$;

$O(U_t)$, $t = k, \dots k+l$ is the values of the estimates of the corresponding vertices.

Function *F* summarizes the search in the entire tree below the node U_i and determines the estimate of the position U_i.

To train the bee algorithm not to repeat the previous mistakes, the *OO* estimate (U_i) may contain summary information on the reason for the unsatisfactory position U_i obtained in the search process, called Minsky's summary. In turn, due to the tree-like structure of the enumeration, it is necessary to recursively summarize the results obtained. One of the concepts that satisfy these considerations is the concept of frames.

Adhering to the opinion of the author of this concept, M. Minsky, the frame for training the bee algorithm to remember the mistakes made and not repeat them in the future is the minimum necessary structured information in the form of a summary, that is, the data structure for representing the situation "awareness of the search", which in the process of implementing the algorithm is stereotypical. From what has been said, it follows that it is advisable to include a set of “resume” - frames in the ABC for each specific case of unsatisfactory position U_i. The corresponding frame is called when its terminals match the descriptions of lower-level situations, and the markers agree with the current goals. This ensures that a specific frame is activated only if it matches the current situation.

The basis of the classical honey bee algorithm can be considered the theory of random search in optimization problems, that is, in the problems of minimizing or maximizing a criterion formulated in a certain way, which characterizes the quality or efficiency of the solution obtained [4, 5, 6]. If we denote such a criterion by the symbols *K*, and the vector characterizing the solution to the problem $r = (r_1, r_2, \dots r_m)$, then the scalar function

$$\Delta K_n \geq 0 \;\; K(r) = K(r_1, r_2, \dots r_m)$$

will express the dependence of the criterion on the obtained solution.

In its turn, the solution r is always limited by external conditions that determine the set of feasible solutions *R*.In these terms, the optimization problem is formalized as follows:

$$K(r) \to m, \text{ where}$$

$$r \in R$$

m is interpreted as min or max depending on the problem statement. If the optimal solution r_{omn} exists and is achievable, then it has the following property:

$$K(r_{omn}) = m\{K(r)\}$$

$$r \in R$$

If it is not attainable, then its search is carried out in some vicinity $\delta\ (r_{omn}) \subset R$. The method (algorithm) for solving the optimization problem is a method for finding r_{omn} (or $r \in \delta\ (r_{omn})$). Naturally, this method has a recurrent nature, that is, it determines the transition from one solution to another, better one, which forms a procedure for the sequential improvement of the solution

$$r_0 \to r_1 \dots \to r_n \to r_{n+1} \to \dots$$

From what has been said it follows that the algorithm for finding the optimal solution establishes a connection between the following solutions. In the simplest case, the formalization of this fact is as follows:

$$r_{n+1} = P(r_n).$$

The search algorithm P is designed to indicate the operations that need to be done at node r_i in order to find the next, more preferred node. In order to call algorithm P a search one (from the point of view of the frame-based approach), algorithm P, in particular, must contain:

– subalgorithm for collecting information about $K(r)$ and R;
– subalgorithm for choosing state r as preceding the next stage of the search;
– subalgorithm for collecting information for resume.

Convenience of representing the search algorithm in terms of increments (steps) at each stage of the search:

$$\Delta r_n = r_n - r_{n-1}$$

leads to relation

$$\Delta r_n = P^*(r_{n-1}),$$

where P^* is also an algorithm (instruction, indication) for finding a displacement Δr_n from node r_{n-1} to a preferred node r_n, which, as in the case of a resume, generally speaking, describes a stereotypical situation.

Formally, a rather rigid theory of algorithms requires the fulfillment of the uniqueness property, the meaning of which lies in the uniqueness of the result of the implementation of the algorithm under the same initial conditions. But this property is possessed only by the class of deterministic, regular search algorithms $\{P\}$.

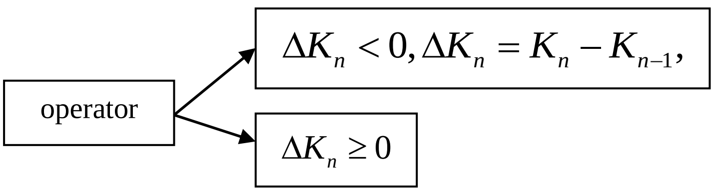
The emergence of the theory of random search has expanded the concept of an algorithm in the sense of assuming the ambiguity of the result of its implementations under the same initial conditions. The emergence of a class of non-deterministic random search algorithms $\{P^{\xi}\}$ contributed to the emergence of the possibility of giving the result of the implementation of the algorithm a statistical (or some other natural) character. The obvious fact is $\{P\} \subset \{P^{\xi}\}$. [7].

3. Results and Discussion

Due to the fact that in the classical ABC the main areas of the possible location of "nectar sources" are found and processed by "scouting bees", according to the general theory of random search algorithms, it becomes useful to distinguish between local and global search. To solve the problem of coloring the vertices of the ABC graph, algorithms of local random search with adventurous gambling and prudently cautious scenarios were investigated. The adventurous-gambling scenario of a random search for the main areas of the possible location of the nectar source turned out to be effective at the first iterations of the AMP, that is, at the stage of rough “localization” of sources far from extrema and at quasilinear areas of the function $K(r)$. Thus, the "scout bees" are primarily supplied with an adventurous search scenario. Its simplest content includes two main operators:

– random step operator;
– operator of repetition the previous step.

Since the problem of coloring the vertices of a graph involves finding the minimum number of coloring colors, the organized search consists in minimizing the function $K(r)$. The actions of each of the operators of the gambling scenario can lead to one of two results:



where, $K_n = K(r_n)$ – the value of the function to be minimized at the n-th search step. Depending on the result obtained, one of the indicated operators acts at the next step. The classical random search algorithm with an adventurous gambling scenario is a consequence of the "optimistic" hypothesis that the probability of success $\left(\Delta K_n < 0\right)$ in a previously successful direction more than in a random one, which motivates to repeat successful steps, and in case of failure, to take another random step.
In this case, the frame script can contain the following recurrent form:

$$\Delta r_n = \begin{cases} \Delta r_{n-1}, \Delta K_{n-1} < 0 \\ \nu\xi, \Delta K_{n-1} \geq 0, \end{cases}$$

where

$\nu = |\Delta r|$ is random step size;

ξ is unit random vector, evenly distributed in all directions of the search space, which makes the directions of the random vector, ξ equally probable. The adventurousness and gambling of the scenario lies in the adventurous behavior of the seeker - the repetition of a successful step that is inherent, according to experts, to the collective behavior of living organisms. Prudently scripted random search practically simulates trial and error. The classical algorithm with a prudently cautious scenario is implemented by two main operators:

- operator of a random step (as well as an adventurous one);
- return operator;

Its idea is that only successful steps are used in the search, and unsuccessful ones are guided by a return statement. The script frame of this algorithm can contain a recurrent form of the following form:

$$\Delta r_n = \begin{cases} -\Delta r_{n-1}, \Delta K_{n-1} \geq 0 \\ \nu\xi, \Delta K_{n-1} < 0. \end{cases}$$

A random search algorithm with a prudently cautious scenario is natural in situations in the vicinity of extrema or at the bottom of the "ravines" of the function to be minimized, where repeating successful steps is impractical due to the very low probability of repeated success. Foraging bees are supplied with this search algorithm. It is known that a random search algorithm with a prudently-prudent scenario is effective in optimizing multiparameter problems. A decrease in its efficiency with an increase in the dimension of the problem is comparable to $\frac{1}{\sqrt{n}}$, in contrast to regular algorithms, in which the decrease in efficiency is commensurate with $\frac{1}{n}$. This fact convinces of the reliability of the random search algorithm with a prudently cautious scenario when solving problems of high complexity.

The considered scenarios of random search algorithms without additional adjustments are applicable only to solving problems of unconstrained optimization, that is, for problems in which the boundaries of the region of constraints (region of feasible solutions) do not play a significant role. In optimization problems, where the area of limitations is rather narrow and significant, in each considered scenario it is necessary to have an indication of going beyond the boundary of the area of feasible solutions in the search process and the corresponding reaction (return to the specified area). In classical random search algorithms, the procedure for taking into account constraints is formulated on the perception of violation of the constraint as an unsuccessful step and forms a response to it by returning to the previous state. Therefore, in an adventurous and gambling scenario, it is natural to use the return operator, the purpose of which is not a direct search, but retention in the area of task constraints.

In a prudent scenario, a return statement can perform both of these functions. In the algorithm of honey bees, to solve the problem, the situation of a successful step was formulated as follows:

$$SSS:\left(r\in R\right)\&\left(\Delta K<0\right),$$

and the situation of an unsuccessful step, respectively

$$SUS:\left(r\notin R\right)\vee\left(\Delta K\geq 0\right).$$

The initial stage of the process of drawing up an acceptable timetable for college studies cannot but contain the hypothesis that the desired acceptable timetable exists [8].

In this case, the hypothesis can be considered justified if at each moment of time of this process the values of all its initial conditions are known.

Not knowing at least one of the initial conditions makes the process unpredictable. Despite the fact that the concepts of unpredictability and randomness are not identical, the nature of the possible algorithm for solving the problem becomes explainable. In contrast to unpredictability, the idea of a random approach assumes the presence of certain stable statistical properties. The honeybee algorithm, having acquired the necessary stable statistical properties and, together with them, the "ability" to solve the problem, at the same time acquires the ability to be controlled (change in the course of work) when the external search conditions and certain search situations change. The development of the ability to adapt to changing initial conditions and the search situation can be carried out both in the parametric and structural directions.

Representing the search algorithm *P* as a pair: *P* = <*S*, *V*>, where *S* is a structure, *V* is a set of parameters, the ways and mechanisms of this development become clear. The parametric adaptability of the algorithm is associated with control, in particular, with the main parameters of the random search algorithm - the step size *v* and the law of the probability distribution of the random step ξ. The search step size can be controlled, for example, using eristics, implying a decrease in the step in case of an unsuccessful random step (hypothesis about the proximity of the search target and the need to search in smaller steps) and increase the step with a successful random step (hypothesis about the distance of the target). Obviously, this heuristic does not give any guarantees, but it presupposes a completely understandable well-known theoretical formalization:

$$v_n=\begin{cases}\alpha v_{n-1},\Delta K_{n-1}<0\\ \dfrac{1}{\alpha^{\frac{p}{1-p}}}v_{n-1},\Delta K_{n-1}\geq 0\end{cases}$$

the control parameter $\alpha > 1$, and it is obvious that as $\alpha \to 1$, the step size will differ less and less from the previous one, *p* is the probability of a successful random step at its (current) value. In this case, the amount of step reduction, in case of failure $\dfrac{1}{\alpha^{\frac{p}{1-p}}}<1$.

The well-known theoretical formalization of control over the size of a random step corresponds only to the situation when *K (r)* is a spherical field around the optimum r_{onm} . Each specific problem assuming the use of ABC for the purpose of finding δ (r_{omn}) implies the determination of the parameter α (for $\Delta K_{n-1}<0$) and the parameter $\beta=\dfrac{1}{\alpha^{\frac{p}{1-p}}}$ (for $\Delta K_{n-1}\geq 0$) in the process of direct implementation of the search, which, in fact, is the parametric adaptability of ABC. The adaptability of the probability distribution law of the random step ξ in the simplest case presupposes the compulsion to a systematic drift of the random walk in the search process in the direction of the search target r_{omn} . Under the assumption that the mathematical expectation of a random vector is ξ_{sb}:

$$M\left[\xi_{sb}\right]=0,$$

and the direction of the target is known, then the addition of the vector d (determining the systematic drift to ξ_{sb}:

$$\xi=\xi_{sb}+d,$$

will lead to the fact that $M\left[\xi\right]=d.$

Following the theory of random search, the vector d, in turn, is a "participant" in the history of the search and an indicator of its promising direction. A heuristic approach to the formation of the direction of the vector d in the search process can be given by the following simple recurrent form:

$$d_n=id_{n-1}-\partial\Delta r_{n-1}\Delta K_{n-1},$$

where

$0 \leq i \leq 1$ is coefficient of ignoring previous information, $\partial > 0$ is coefficient of dominance of new information. In particular, with complete disregard for the past ($i = 0$):

$$d_n=-\partial\Delta r_{n-1}\Delta K_{n-1},$$

the direction of the drift coincides with the directions of the previous successful step or opposite to the failed one.

In the complete absence of new information ($\Delta K_{n-1} = 0$), with $i << 1$ $d_n \rightarrow 0$, the search becomes equally probable, that is, unbiased. The ratio of the directions of the vectors d_n and Δr_n provides a turn of d_{n+1} in the direction of a successful step and in the direction opposite to a failed one. This kind of adaptability mimics the self-learning of an algorithm. It does not necessarily have a positive effect on the search process (there is a significant dependence on the formulation of the problem), but the facts of negative influence are unknown. Structural adaptability of the algorithm implies changing the tactics (and maybe even the strategy) of the search in order to maintain the level of efficiency. If the scenarios for the search for the main areas of possible location of the "nectar sources" are designated $C_1, C_2 \dots C_n$, then the established sequence of the scenarios $\{C_i\}_{i=1,\dots n}$ can be considered a search structure. It is natural to assume that each specific problem (with specific goals and a system of restrictions) has its own structure for finding a solution and the process of its evolution. There are many methods of structural adaptability of the search algorithm, in particular, the situation identification method, the multi-choice method, etc. etc.

In this case, naturally, structural adaptability must be supported by parametric adaptability. The questions of the structural adaptability of ABC in these studies are still open. The questions of the structural adaptability of AMP in these studies are still open. The foregoing increases interest in the theory of frames, in particular, in the theoretical aspects of frame-scenarios as typical structures for certain actions, concepts, events, including their characteristic elements. Frame script terminal markers contain specific recommendations for filling terminals with a job, while terminals contain recommendations for assessing a situation corresponding to a job. Therefore, according to M Minsky , it is useful to represent each situation at a certain stage of solving the problem with a system of frame scenarios, where each frame scenario represents one of the possible points of view on the situation. The framing system then corresponds to many different ways of using the information available on the common terminals. From the point of view of R Schenk and R Abelson , a scenario is understood as a sequence of actions in a frequently occurring situation, in which the principle of casual communication is used, implying the result of each action as conditions leading to the next action.

Conclusion

Thus, the frame-scenario approach of organizing the bee algorithm allows:

- significantly save memory;
- to speed up the process of perception of the situation that has developed in the process of solving the problem;

- to unify the algorithm due to the experience in solving similar or similar problems.

Interesting, relevant and promising issues for further research on AMP are:

- exploring the benefits of "natural" and "artificial" selection in the process of tuning and training the honeybee algorithm;
- selection criteria for “artificial” selection at different stages of AMP work;
- study of the share of "independence" and the share of "collectivity" for each simple agent in the process of "searching" and "collecting nectar".

References

[1] Kurejchik V V and Zaporozhec D Y 2010 Roevoj algoritm v zadachakh optimizacii. *Izvestie YFU. Tekhnicheskie nauki* vol 7(108) pp 28-32

[2] Kurejchik V V and Kazharov A A 2010 Primenenie pchelinogo algoritma dlya raskraski grafov. *Izvestie YFU. Tekhnicheskie nauki* vol 12(113) pp 30-36

[3] Minsky M 1975 A framework for representing knowledge. P H Winston (Ed.) *The psychology of computer vision* (New York: McGraw-Hill Book).

[4] Rastrigin L A 1968 Statisticheskiye metody poiska (Moscow: Nauka) p 376

[5] Rastrigin L A 1969 Teoriya I primenenie sluchajnogo poiska (Riga: Zinatne) p 309

[6] Rastrigin L A 1974 Sistemy ehkstremalnogo upravleniya. *Glavnaya redakciya fiziko-matematicheskoj literatury izdatelstva* (Moscow: Nauka) p 632

[7] Ripa K K and Tarasenko G S 1978 Adaptaciya sluchajnogo poiska (Riga: Zinatne)

[8] Ponomarenko O 2020 Building an acceptable training schedule based on the artificial bee colony algorithm. Humanitarian approaches to the Periodic Law *Science and society Methods and problems of practical application. Proceedings of the 9th International conference. Accent graphics communications & Publishing.* Vancouver, Canada. Pp 10-15