Electronic training polygon for artificial intelligence systems

A.A. Karandeev[1,2], V.I. Baluta[1,2], V.P. Osipov[1]
karalex755@gmail.com | vbaluta@keldysh.ru | osipov@keldysh.ru
[1]Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Moscow, Russia
[2]Plekhanov Russian University of Economics

In the framework of the concept of neoconflictology, the possibilities and methods of mathematical modelling of conflict dynamics under uncertainty are considered. Complex contradictory situations, when it is necessary to quickly respond to changes in the situation and perform some actions in conditions of uncertainty, are often found in different spheres of activity. In this case, the uncertainty may be due to incomplete knowledge of the situation, the inability to quickly understand and evaluate possible options for its development, the influence on it from other participants in the process, etc. The result of further developments depends on how the decision will be to the situation and what actions will follow on its basis. Usually, the effectiveness of behavior in difficult situations is determined by the level of special training and experience of the decision maker. The trend of using artificial intelligence systems to support decision-making, including complex conflict situations with a high level of uncertainty, is now more frequent. Focusing on the use of such systems requires the creation and improvement of specialized tools and technologies, thanks to which the artificial intelligence systems themselves can form an information base, which is a necessary factor for developing rational solutions in a complex environment. The creation of a virtual polygons is one of the ways for producing such information base.

Keywords: *neoconflictology, conflict, training polygon, intelligent agent, Artificial intelligence, modelling.*

1. Introduction

In an antagonistic conflict, the effectiveness of achieving goals depends on the adopted strategy of behaviour, the speed of decision-making in a changing environment and the level of their optimality, or in other words, the degree of their compliance with the current situation.

The best way to analyze situations of this kind that demonstrate emerging phenomena or generate unforeseen patterns is to model and simulate them [1]. The problematic complexity of solving these issues is directly related to the level of uncertainty, which is often due to a lack of time to obtain the information necessary for decision-making in a conflict situation and many other factors.

Given the modern development of computer methods and tools, it is numerical modelling that assumes the role of the tool by which you can explore a variety of conflicts to identify the role and significance of certain behavioural strategies, determine the list of effective actions in various situations. One of the methods that greatly simplify the modelling process is the creation of an electronic polygon. Despite the complexity of developing such kind of software systems, they make it possible to visually simulate various situations and see what these or those solutions lead to.

The main participants in a computer polygon are intelligent agents [2-3] who are trying to achieve their goals. They can have both common goals and completely opposite. It is worth considering that for each of the participants in the conflict, their own space of probable states can be formed, due to its capabilities and ideas about the "world in which it operates." Moreover, the state of the agent can be changed not only due to the actions taken by him, but also under the influence of other agents who are most often the adversaries. The agent’s task is to construct an optimal trajectory for transition from its initial state to a given target state.

Usually, the set of actions that an agent can perform is limited, but each of these actions must be taken into account not only affects its own state, but also the state of other agents in their phase space. The purpose of these repeated experiments is to train the agent for rational action in the given conditions. To achieve this effect, a mathematical model is designed.

2. Basic provisions of the model approach

The developed research technology is based on numerical modelling of possible ways to actualize an antagonistic conflict based on the following considerations. In the simulation model, each of the subjects of the conflict can be represented as an “intelligent” agent. The theory of agents or multi-agent systems [4] is a computer theory which seeks to apprehend the coordination of competing independent process. An agent is thus a computer process [5], which can be considered as autonomous since it is capable of adapting when its environment changes. The basic characteristics of such an agent are the formation of an internal model of the surrounding world and the presentation of its place in it, a system of rules for creating and changing this model, decision-making methods for taking actions to respond adequately in response to a changing situation. Intelligent agents have targeted behaviour to achieve a certain goal in the most optimal way.

This process comes down to setting optimization problems and finding their solutions. Based on the described representations, each of the subjects of the confrontation can be represented in the form of a complex system that has certain resources and has its own goals. It is worth considering that many types of conflicts, such as social and political, cannot be strictly defined [6]. Because of this, it may be necessary to discard some non-significant parameters.

The subject's capabilities depend on his position in the phase state space at each current point in time, which is characterized by a certain set of particular characteristics. Otherwise, a change in its state over time depends on the position of the subject in the phase space of his states, which also include the physical coordinates of his location. Moreover, each subject has its own phase space, only partial intersection with the opponent’s space is possible. The transition of the subject from the initial state to the target is carried out through a set of intermediate states, the trajectory of motion in phase space is a result of their change.

Conventionally, the reflection of the model of the intelligent agent in the form of a graph, according to [7-10], can be displayed in the form of the circuit shown in Fig. 1. The initial position is conditionally shown in a blue circle, and the target in yellow. Choosing a path from one state to another involves a certain sequence of actions (steps), which is determined by the subject's ideas about the cost of resources for each of them and an assessment of their proximity to the target state. The diagram shows that each individual action is associated with a certain resource of the real world, the availability of which (in the representation of the object) is a factor determining the possibility of its implementation.

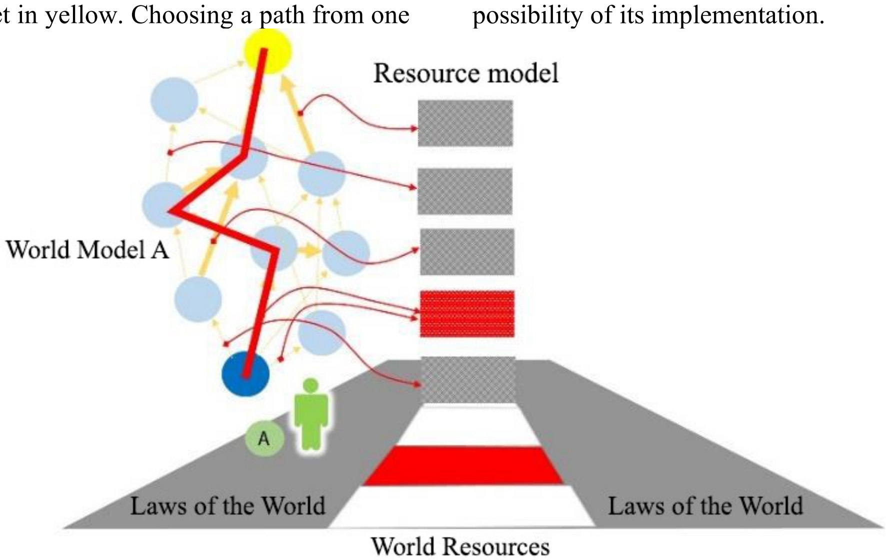

Fig. 1. Intelligent agent action planning

In other words, any change in the parameter that describes some of the characteristics of the subject's position in the phase space can be associated with a change in some scalar quantity — the transition price, which is understood as the conditional cost of the action. Each of the parameters has its own "price scale", which dynamically depends on the current state of the subject and is determined by the value of this parameter to solve the problem.

3. Adapting rules of conduct

In the phase space or in the space of states we define some arbitrary point, which will be considered as the initial position of the subject. We also set the second point denoting the position of the target. The state of the subject changes as a result of some elementary actions in the state space, while the set of possible actions is limited and is described by a set of rules of behavior, each of which is displayed by a vector in the phase space. The sequential step-by-step execution of actions forms a certain trajectory of the subject in the state space. In a two-dimensional interpretation, the task is shown schematically in Fig. 2.

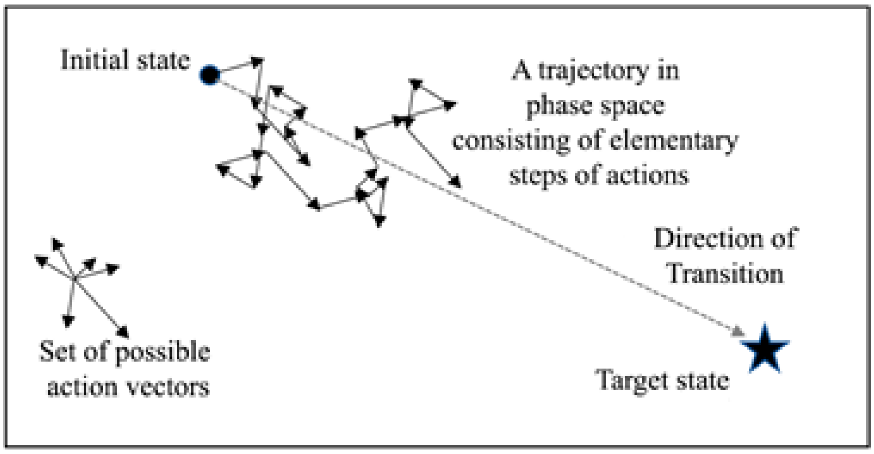

Fig. 2. Schematic representation of the problem

The shortest direction or trajectory from the source to the target state is shown by a dashed line. It is obvious that the result of the implementation of one or another rule of behavior, which is displayed in the phase space by some action vector, is unlikely to coincide with the shortest direction. Clearly, a rational trajectory of movement should be as close to this line as possible.

But in conditions of inaccurate information about the properties of each of the vectors, it is difficult to construct. For the two-dimensional problem shown in figure 2, such an operation can be done "manually", based on visual estimates. One of the possible results may be, for example, the path options shown in Fig. 3.

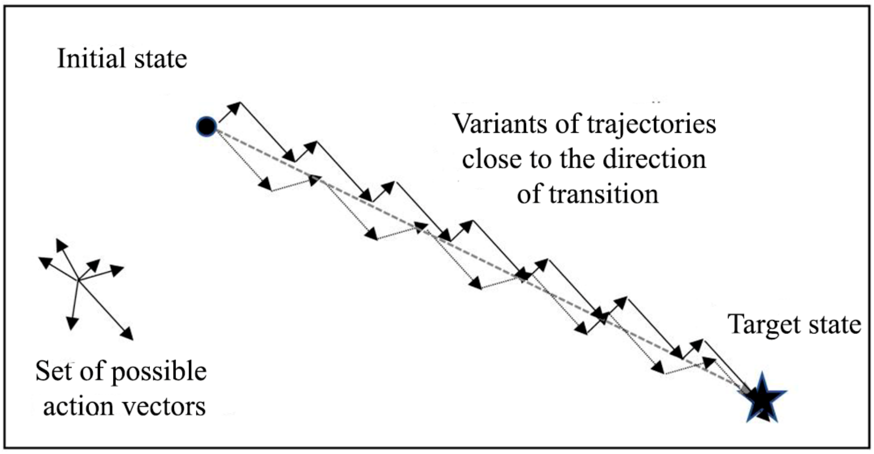

Fig. 3. A schematic example of the construction of rational trajectories

In the presented schematic example, based on a visual assessment of the set of action vectors, such trajectories can be constructed "manually". There are at least two rational options for the trajectories of the transition from the initial state to the target, and it is enough to use a combination of only two of all possible actions. When comparing the obtained trajectories, it is seen that that the upper one is slightly more preferable in terms of the degree of approximation to the target position, but its implementation requires one more step. To determine the best of them, additional evaluation criteria are needed, for example, a comparison of criteria for accuracy and resource costs. Based on the above examples, it can be argued that in a multidimensional phase space, the construction of the optimal path is possible only with full knowledge of the situation: where, by what means, with what intermediate results certain actions arising from the rules of behaviour of agents can be realized.

The following scheme was proposed and tested. Multidimensional phase space and an unordered set of permissible actions are set. At the same time, the number of actions and the dimension of space are determined by the conditions of a specific task. It is assumed that the basis for decision-making by the agent in the current situation to choose a specific action is some a priori assessment of the value of certain actions. Initially, it can be set by experts or randomly generated.

Further refinement of the assessment of the actions that the agent must carry out is the result of a wide series of numerical experiments, in each of which the trajectory of motion of the agent states in the phase space from the source to the target is constructed. In the process of performing calculations, a certain rating is assigned to each perfect action. The rating of each action is determined by the degree of approximation of the new state to the target state obtained as a result of this action from this position. With a positive effect (approaching the goal), incentive is introduced, with a negative effect (moving away from the goal) - a penalty. At the beginning of the calculation, all ratings are set close to zero.

The next move step is randomly selected from the list of all possible actions, but taking into account their rating. The probability of choosing an action increases in proportion to the rating. As a result of the calculations, the initially uniform distribution of the rating of actions is gradually deformed towards growing ratings. At each iteration, the current state in which the subject was located is remembered. Thus, a model of stereotypical situations is formed in which the effects of the same actions can manifest themselves in different ways. Each iteration stops after reaching the target state or after completing a given number of steps. At each subsequent iteration, the ratings accumulated in previous experiments are saved and used when choosing actions. Therefore, the knowledge base is accumulated and the conditional agent is "trained" in the rational choice of actions depending on the position in the state space in which it is located.

4. Results

Testing of the methodology was carried out on an arbitrary example in a space of 12 measurements with a total number of rules of behavior equal to 50 (Fig. 4).

The fig. 4 shows 33 of 50 actions in a 12-dimensional state space. The positions of the initial and target states were randomly selected as vectors in the phase space. In addition, in the form of increments of vectors, action vectors were arbitrarily specified.

Name	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12
1	0,07	-0,45	0,29	-0,27	0,26	-0,04	0,47	0,3	0,2	0,12	-0,22	0,15
2	0,11	-0,38	0,52	0,82	-0,5	-0,83	0,17	0,63	-0,66	0,29	-0,2	0,04
3	0,55	0,13	-0,5	-0,07	-0,14	-0,09	0,19	0,71	0,62	-0,44	-0,64	0,41
4	0,32	-0,06	-0,33	-0,3	-0,04	-0,24	-0,34	0,18	-0,09	-0,23	0,3	0,61
5	-0,19	0,23	0,35	0,48	0,74	-0,73	-0,08	0,27	-0,29	-0,18	-0,6	-0,37
6	-0,71	0,06	0,07	0,03	0,37	0,03	0,42	0,03	0,47	-0,78	0,21	-0,55
7	-0,5	-0,46	0,43	-0,34	0,39	0,15	0,28	0,04	0,05	-0,08	-0,06	0,39
8	0,17	-0,1	0,23	-0,37	0,2	-0,31	0,04	0,97	0,03	-0,36	-0,14	-0,07
9	0,21	-0,05	0,22	-0,07	0,51	-0,2	0,01	0,37	0,31	0,47	0,64	0,12
10	0,21	-0,56	-0,66	-0,49	-0,38	0,55	-0,64	0,09	0,31	0,23	-0,01	-0,19
11	-0,67	-0,25	0,15	-0,56	-0,29	0,2	0,89	0,15	-0,29	0,53	-0,41	0
12	0,16	-0,21	-0,02	-0,2	0,18	0,27	-0,46	0,34	0,17	0,5	0,8	-0,17
13	0,26	-0,18	0,26	-0,53	0,24	-0,08	-0,5	0,11	-0,64	0,41	-0,1	-0,55
14	-0,3	-0,04	0,61	0,4	0,09	-0,31	-0,08	0,25	-0,04	0,02	-0,35	0,01
15	-0,14	-0,52	0,09	-0,38	0,74	0,11	-0,1	0,95	0,21	-0,48	0,1	0,48
16	-0,14	0,03	-0,47	0,05	0,42	0,3	0	-0,2	0,5	0,09	0,18	0,24
17	-0,38	-0,17	0,56	0,38	-0,05	0,48	-0,92	0,08	0,26	-0,35	-0,12	-0,11
18	0,8	0,47	0,12	-0,47	0,76	-0,41	0,35	0,13	0,15	-0,31	-0,35	0,55
19	0,16	0,26	0,56	-0,2	-0,49	0,4	0,38	-0,14	-0,61	0,04	-0,33	0,06
20	-0,42	-0,39	-0,69	0,06	-0,6	-0,1	-0,13	-0,54	-0,12	-0,36	0,06	0,41
21	-0,56	-0,46	0,33	0,88	-0,63	0,19	0,04	-0,11	0,01	-0,42	-0,4	-0,99
22	0,01	-0,42	0,57	-0,85	0,45	0,21	-0,09	-0,12	-0,35	0,01	-0,72	0,62
23	-0,14	-0,22	0,35	-0,16	0,51	-0,26	-0,02	0,32	0	-0,65	-0,12	-0,22
24	0,13	-0,77	0,7	-0,2	0,53	0,34	-0,1	-0,01	-0,91	0	-0,14	0,64
25	0,1	-0,36	0,36	0,91	-0,2	-0,46	0,14	0,6	0,65	-0,1	0,44	0,53
26	-0,22	-0,2	0,53	0,51	-0,36	-0,13	0,32	0,5	0	0,19	-0,59	-0,11
27	0,38	-0,07	-0,2	-0,7	0,17	0,3	0,13	-0,55	-0,39	-0,5	-0,56	-0,08
28	0,01	0,54	-0,32	0,75	-0,7	-0,81	-0,42	0,64	-0,25	0,35	-0,03	0,09
29	0,07	-0,45	-0,13	0,14	0,56	0,69	-0,28	0,13	0,14	-0,75	-0,3	0,89
30	0,11	-0,15	0,53	0,03	0,08	0,64	0,14	0,42	0,76	0,47	-0,36	0,02
31	-0,31	-0,32	-0,05	0,21	0,04	-0,43	0,15	0,18	0,5	-0,81	-0,16	-0,51
32	0,41	0,31	-0,48	0,78	0,12	0,01	0,12	0,38	-0,4	-0,69	-0,82	-0,31
33	0,59	0,19	0,05	0,48	0,12	-0,8	-0,62	0,23	0,1	0,91	-0,19	-0,07

Fig. 4. List of available actions

Let us dwell in more detail on the data given in the table. As already mentioned above, the agent's training technology is demonstrated using a hypothetical example, in which a certain zone of the phase space is conventionally defined as a range of changes in the coordinate values of the parameters, and the initial and target position of the agent are determined inside this zone. Variants of the agent's possible actions are randomly generated by coordinatewise specifying a certain set of increment vectors. When generating incremental vectors, the only condition is to limit the length of each of these vectors to a value significantly less than the distance between the initial and target position of the agent in the phase space.

The data given in the table reflect the obtained picture of the learning outcomes at some intermediate stage of the experiments. For clarity, a visual analysis introduced the color of the cells. In this case, color means the directivity of the effect, and the intensity of the color is its immediate meaning. In other words, the cells highlighted in red indicate that the choice of this action will lead to the loss of a parameter or resource. Moreover, the more intense the color, the greater the loss. Positive effects highlighted in green.

Fig. 5 shows the result of a computational procedure for finding a rational trajectory.

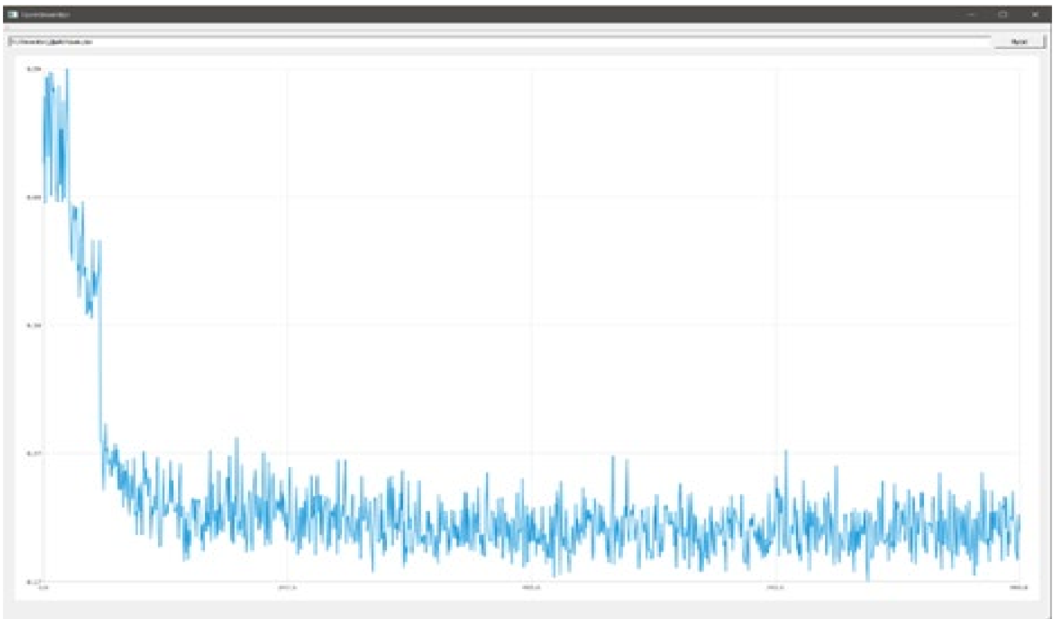

Fig. 5. Schematic representation of the problem

The process of qualitatively improving the procedure for choosing rational actions from a given set is visible in the fig. 5. The x-axis shows the number of experiments, in each of which a motion pattern limited in the number of steps is plotted in the phase space, and the y-axis is the distance from the target state in the adopted metric. It can be seen that the agent training process leads to the achievement of some non-improved level of performance, which, however, allows to get close enough to the target state.

5. Conclusions

Proposed technological solution for modeling the behavior of intelligent agents provides a fairly broad basis for studying the features of constructing trajectories of achieving goals in phase space and the adaptive behavior of agents in a conflict situation.

Due to the specifics of the problem, approaches and methods for modeling the phase space of states of a conflict environment are proposed, which allow determining strategies for rational trajectories of goal achievement. The technology for modeling agent behavior is based on a detailed description of their behavior based on a scheme of intelligent transitions between states and behavior models.

Such a detailed description of complex behavior models makes it possible to uniformly display in the model of a conflict environment the essential aspects of the behavior of real-world prototype objects. The proposed scheme for ensuring the adaptive behavior of agents in a conflict environment is an integral part of the methodological and algorithmic support of an electronic training ground for studying the conflict interaction of complex systems.

According to the authors, this technology has shown its effectiveness in test tasks and can be recommended for use as part of an electronic training polygon to test its capabilities in developing strategies for agents' behavior under conditions of uncertainty during their interaction.

A similar technology can be applied in automated process control support systems, where an operational assessment of the situation is necessary based on the available incomplete or inaccurate data with the development of recommendations for rational actions. In particular, when creating combat control systems, ensuring the safety of facilities, responding in emergency situations, as well as in control systems for unmanned vehicles.

References

[1] Klügl F., Oechslein C., Puppe F., Dornhaus A., Multi-Agent Modelling in Comparison to Standard Modelling, 2014.

[2] Nechaev Y., Osipov V., Chetverushkin B., Baluta V., Ontological synthesis of management decisions in conditions of antagonistic conflicts, 2018.

[3] Baluta V., Nechaev Y., Osipov V., Chetverushkin B., Conceptual base of a supercomputer platform of application-oriented simulation, prediction and expertizes of conflict interaction, 2018.

[4] Bousquet F., Barreteau O., Le C., Mullon C., Weber J., An environmental modelling approach. The use of multi-agent simulations, 1999.

[5] Ferber J., La kénétique: dés systems multi-agents à une science de l'interaction, 1994.

[6] Petukhov A., Malhanov A., Sandalov V., Petukhov V., Mathematical Modeling of the Ethno-social Conflicts by Non-linear Dynamics. Proceedings of the 7th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, 2017.

[7] A.Kh. Khakimova, O.V. Zolotarev, M.A. Berberova. Visualization of bibliometric networks of scientific publications on the study of the human factor in the operation of nuclear power plants based on the bibliographic database Dimensions. Scientific Visualization, 2020, volume 12, number 2, pages 127-138, DOI: 10.26583/sv.12.2.10, E-ISSN:2079-3537.

[8] M.A.Berberova, K.I.Chernyavskii, «Comparative assessment of the NPP risk (on the example of Rostov and Kalinin NPP). Development of risk indicators atlas for Russian NPPs», GraphiCon 2019 Computer Graphics and Vision. The 29th International Conference on Computer Graphics and Vision. Conference Proceedings (2019), Bryansk, Russia, September 23-26, 2019, Vol-2485, urn:nbn:de:0074-2485-1, ISSN 1613-0073, DOI: 10.30987/graphicon-2019-2-290-294, http://ceur-ws.org/Vol-2485/paper67.pdf, p. 290-294.

[9] M.A.Berberova, A.S.Oboimov, A.Kh.Khakimova, O.V.Zolotarev, «Risk-informed security system. The use of surveillance cameras for the particularly hazardous facilities safety», GraphiCon 2019 Computer Graphics and Vision. The 29th International Conference on Computer Graphics and Vision. Conference Proceedings (2019), Bryansk, Russia, September 23-26, 2019, Vol-2485, urn:nbn:de:0074-2485-1, ISSN 1613-0073, DOI: 10.30987/graphicon-2019-2-321-325, http://ceur-ws.org/Vol-2485/paper74.pdf, p. 321-325.

[10] Baluta V., Osipov V., Chetverushkin B., Yakovenko Y., Conceptual issues of model representation of conflicts. Proceedings of the International Scientific Conference, 2019.

About the authors

Karandeev Alexander A., phd student and Junior Scientific Associate of the Keldysh Institute of Applied Mathematics, Russian Academy of Sciences. E-mail: KarAlex755@gmail.com

Baluta Viktor I., PhD in Technical Science, Docent, Senior Scientific Associate of the Keldysh Institute of Applied Mathematics, Russian Academy of Sciences. E-mail: vbaluta@keldysh.ru

Osipov Vladimir P., PhD in Technical Science, Docent, Lead Scientist of the Keldysh Institute of Applied Mathematics, Russian Academy of Sciences. E-mail: osipov@keldysh.ru