Modeling Mixed Groups of Humans and Robots with Reflexive Game Theory

Sergey Tarasenko

Department of Intelligence Science and Technology, Graduate School of Informatics
Kyoto University, Yoshida honmachi, Kyoto 606-8501, Japan
infra.core@gmail.com

Abstract. The Reflexive Game Theory is based on decision-making principles similar to the ones used by humans. This theory considers groups of subjects and allows to predict which action from the set each subject in the group will choose. It is possible to influence subject's decision in a way that he will make a particular choice. The purpose of this study is to illustrate how robots can refrain humans from risky actions. To determine the risky actions, the Asimov's Three Laws of robotics are employed. By fusing the RGT's power to convince humans on the mental level with Asimov's Laws' safety, we illustrate how robots in the mixed groups of humans and robots can influence on human subjects in order to refrain humans from risky actions. We suggest that this fusion has a potential to device human-like motor behaving and looking robots with the human-like decision-making algorithms.

Keywords: Reflexive Game Theory (RGT), Asimov's Laws of Robotics, mixed groups of humans and robots, human-robot societies.

1 Introduction

Now days robots have become an essential part of our life. One of the purposes robots serve to is to substitute human beings in dangerous situations, like defuse a bomb etc. However human nature shows strong inclinations towards the risky behavior, which can cause not only injuries, but even threaten the human life. The list of these reasons includes a wide range starting from irresponsible kids' behavior to necessity to find solution in a critical situation. In such a situation, a robot should full-fill a function of refraining humans from doing risky actions and perform the risky action itself.

However, robot should not physically force people, but must convince people on the mental level to refrain from doing an action. This method is more effective rather than a simple physical compulsion, because humans make the decisions themselves and treat these decisions as their own. This approach is called a *reflexive control* [1].

We consider the mixed groups of humans and robots. To be able to interact with humans on the mental level robot should posses an ability to "think" and make decisions in a way similar to the one that humans have.

M.H. Lamers and F.J. Verbeek (Eds.): HRPR 2010, LNICST 59, pp. 108–117, 2011.

The principles explaining human decision-making are the basis of the Reflexive Game Theory (RGT) proposed and developed by Lefebvre [1,2,3]. By using the RGT, it is possible to predict choices made by each individual in the group and influence on their decision-making. In particular, the RGT can be used to predict terrorists' behavior [4].

The purpose of the present study is to apply RGT for analysis of individual's behavior in the mixed groups of humans and robots and illustrate how RGT can be used by robots to refrain humans from doing risky actions. We start with brief description of the RGT and illustrate its application with a simple example. Then we formalize the definition of robots, distinguishing them from humans. Finally, we consider two examples, in which humans tend to do risky actions, and show how robots, using RGT, can refrain humans from doing these actions.

2 Brief Overview of the Reflexive Game Theory (RGT)

The RGT deals with groups of individuals (subjects, agents etc). Any group of subjects is represented in the form of *fully connected graph*. In the present study, a subject can be either a human or a robot. Each subject is assigned a unique variable (*subject variable*), which is a vertex of the graph. The RGT uses the set theory and Boolean algebra as the basis for calculus. Therefore the values of subject variables are elements of Boolean algebra.

All the subjects in the group can have either alliance or conflict relationship. The relationships are identified as a result of group macroanalysis. It is suggested that the installed relationships can be changed. The relationships are illustrated with graph ribs. The solid-line ribs correspond to alliance, while dashed ones are considered as conflict. For mathematical analysis alliance is considered to be conjunction (multiplication) operation $(\cdot)$, and conflict is defined as disjunction (summation) operation $(+)$.

The graph presented in Fig. 1a or any graph containing any sub-graph isomorphic to this graph are not decomposable. In this case, the subjects are excluded from the group one by one, until the graph becomes decomposable. The exclusion is done according to importance of the other subjects for a particular one. Any other fully connected graphs are decomposable and can be presented in an analytic form of a corresponding *polynomial*. Any relationship graph of three subjects is decomposable (see [3,4]).

Consider three subjects a, b and c. Let subject a is in alliance with other subjects, while subjects b and c are in conflict (Fig. 1b). The polynomial corresponding to this graph is $a(b + c)$.

Regarding the relationship, the polynomial can be stratified (decomposed) into *sub-polynomials* [2,3,4,5]. Each sub-polynomial belongs to a particular level of stratification. If the stratification regarding alliance was first built, then the stratification regarding the conflict is implemented on the next step. The stratification procedure finalizes, when the *elementary polynomials*, containing a single variable, are obtained after a certain stratification.

The result of stratification is the *polynomial stratification tree (PST).* It has been proved that each non-elementary polynomial can be stratified in an unique

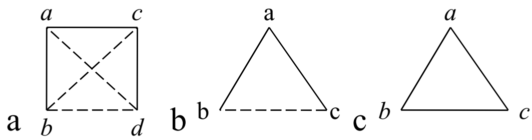

Fig. 1. The relationship graphs

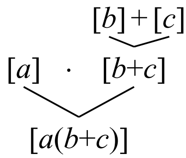

Fig. 2. Polynomial Stratification Tree. Polynomials $[a], [b]$ and $[c]$ are elementary polynomials.

way, i.e., each non-elementary polynomial has only one corresponding PST [5]. Each higher level of the tree contains polynomials simpler than the ones on the lower level. For the purpose of stratification the polynomials are written in square brackets. The PST for polynomial $a(b+c)$ is presented in Fig. 2.

Next, we omit the branches of the PST and from each non-elementary polynomial write in top right corner its sub-polynomials. The resulting tree-like structure is called a *diagonal form*. Consider the diagonal form corresponding to the PST in Fig. 2:

$$[a(b+c)]^{[a][b+c]^{[b]+[c]}} .$$

Hereafter the diagonal form is considered as a function defined on the set of all subsets of the *universal set*. The universal set contains the *elementary actions*. For example, these actions are actions α and β. The *Boolean algebra* of the universal set includes four elements: $1 = \{\alpha, \beta\}, \{\alpha\}, \{\beta\}$ and the empty set $0 = \{\}$. These elements are all the possible subsets of universal set and considered as alternatives that each subject can choose. The alternative $0 = \{\}$ is interpreted as an inactive or idle state. In general, Boolean algebra consists of 2^n alternatives, if universal set contains n actions.

Formula $P^W = P + \overline{W}$, where $\overline{W}$ stands for negation of W [6], is used further to fold the diagonal form. During the folding, round and square brackets are considered to be interchangeable. The following equalities are also considered to be true: $x + \overline{x} = 1, x + 0 = x$ and $x + 1 = 1$. Next we implement folding of diagonal form of polynomial $a(b+c)$:

$$[a(b+c)]^{[a][b+c]^{[b]+[c]}} = [a(b+c)]^{[a]([b+c]+\overline{[b]+[c]})} = a(b+c) + \overline{a} \, .$$

The goal of each subject in the group is to choose an alternative from the set of alternatives under consideration. To obtain choice of each subject, we consider the *decision equations*, which contain subject variable in the left-hand side and the result of diagonal form folding in the right-hand side: $a = (b + c)a + \overline{a}, b = (b + c)a + \overline{a}$ and $c = (b + c)a + \overline{a}$.

To find solution of the decision equations, we consider the following equation, which is a *canonical form of the decision equation*:

$$x = Ax + B\overline{x} \ , \tag{1}$$

where x is the subject variable, and A and B are some sets. This equation has solution if and only if the set B is contained in set A: $A \supseteq B$. If this requirement is satisfied, then equation has at least one solution from the interval $A \supseteq x \supseteq B$ [6]. Otherwise, the decision equation (1) has no solution, and it is considered that subject cannot make a decision. Thus, he is considered to be in frustration state.

Therefore, to find solutions of equation, one should first transform it into the *canonical form*. Out of three presented equations only the decision equation for subject a is in the canonical form, while other two should be transformed into. We consider explicit transformation only of decision equation for subject b: $a(b+c)+\overline{a} = ab+ac+\overline{a} = ab+(ac+\overline{a})b+(ac+\overline{a})\overline{b} = (a+\overline{a}+ac)b+(ac+\overline{a})\overline{b} = (1 + ac)b + (ac + \overline{a})\overline{b} = b + (ac + \overline{a})\overline{b}$. Therefore, $b = b + (ac + \overline{a})\overline{b}$.

The transformation of equation for subject c be can be easily derived by analogy: $c = c + (ab + \overline{a})\overline{c}$.

Table 1. Influence Matrix

	a	b	c
a	a	$\{\alpha\}$	$\{\beta\}$
b	$\{\beta\}$	b	$\{\beta\}$
c	$\{\beta\}$	$\{\beta\}$	c

The variable in the left-hand side of decision equation in the canonical form is the variable of the equations, while other variables are considered as influences on the subject from the other subjects. All the influences are presented in the *Influence matrix* (Table 1). The main diagonal of the Influence matrix contains the subject variables. The rows of the matrix represent influences of the given subject on other subjects, while columns represent the influences of other subjects on the given one. The influence values are used in decision equations.

For subject a: $a = (\{\beta\} + \{\beta\})a + \overline{a} \Rightarrow a = \{\beta\}a + \overline{a}$.
For subject b: $b = b + (\{\alpha\}\{\beta\} + \overline{\{\alpha\}})\overline{b} \Rightarrow b = b + \{\beta\}\overline{b}$.
For subject c: $c = c + (\{\beta\}\{\beta\} + \overline{\{\beta\}})\overline{c} \Rightarrow c = c + (\{\beta\} + \{\alpha\})\overline{c} \Rightarrow c = 1$.

Equation for subject a does not have any solutions, since set $A = \{\beta\}$ is contained in set $B = 1$: $A \subset B$. Therefore, subject a cannot make any decision and is in frustration state.

Equation for subject b has at least one solution, since $A = 1 = \{\alpha, \beta\} \supseteq B = \{\beta\}$. The solution belongs to the interval $1 \supseteq b \supseteq \{\beta\}$. Therefore subject b can choose any alternative from Boolean algebra, which contains alternative $\{\beta\}$. Thus, only alternative $\{\beta\}$ can be implemented.

Equation for subject c turns into equality $c = 1$. This is possible only in the case, when $A = B$. Here $A = B = 1$. Subject c can implement any alternative except for alternative $0 = \{\}$. However, he does not have absolute freedom of choice, since this implies ability to choose inactive alternative $0 = \{\}$, as well.

This concludes a brief overview of the RGT. Next we consider formalization of robotic subjects.

3 Defining Robots in RGT

It is considered by default that robot follows the program of behavior generated by the control system. This control system consists of at least three modules. The Module 1 implements robot's ability of human-like decision-making. The Module 2 contains the rules, which refrain robot from making a harm to human beings. The Module 3 predicts the choice of each human subject and suggests the possible strategies of reflexive control.

We suggest to apply Asimov's Three Laws of robotics [7], which formulate the basics of the Module 2:

1) a robot may not injure a human being or, through inaction, allow a human being to come to harm;
2) a robot must obey any orders given to it by human beings, except where such orders would conflict with the First Law;
3) a robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

We consider that these Laws are intrinsic part of robots "mind", which cannot be erased or corrupted by any means.

The interaction of Modules 1, 2 and 3 in robot's control system is presented in Fig. 3. First the information from environment is formalized as the Boolean algebra of possible alternatives. Then the human-like decision-making system is implemented in Modules 1 and 3. The robot's decision-making based on the RGT is implemented in Module 1. The output of Module 1 is set D, which contains solution of robot's decision equation.

The Boolean algebra is filtered according to Asimov's Laws in Module 2. The output of Module 2 is set U of approved alternatives. Then the conjunction of sets D and U is performed: $D \cap U = DU$. If DU is the empty set ($DU = \{\}$), then a robot chooses alternative from the set U. If set DU is not empty, then a robot selects the actions from the set DU.

To achieve the goal of refraining human subjects from risky action, robot predicts choice of each human subject in Module 3. The decision-making system similar to the one in Module 1 is employed. The output set D_h corresponds to output set D. If robot predicts that choice of some human subject is risky alternative, the robot analyzes all the possible scenarios which succeed in not

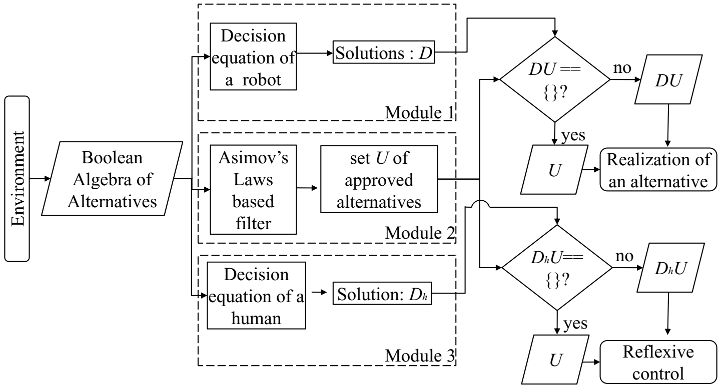

Fig. 3. Schematic representation of robot's control system

choosing the risky alternative and applies reflexive control according to this scenario.

The result of interaction between Modules 1 and 2 includes robot's choices, which are harmless for human subjects, while result of interaction between Modules 3 and 2 is the set of possible scenarios of reflexive control. This is important difference between the outputs of Modules 1 and 2 and Modules 3 and 2.

4 Sample Analysis of Mixed Groups

Here we consider two examples of how robots in the mixed groups can make humans refrain from risky actions. The first example considers robot baby-sitters. The second one illustrates the critical situation with mountain-climbers and rescue robot. The goal of robots in both situations is to refrain humans from doing risky actions.

4.1 Robot Baby-Sitters

Suppose robots have to play a part of baby-sitters by looking after the kids. We consider a mixed group of two kids and two robots. Each robot is looking after a particular kid. Having finished the game, kids are considering what to do next. They choose between to compete climbing the high tree (action α) and to play with a ball (action β). Together actions α and β represent the active state 1=$\{\alpha, \beta\} = \{\alpha\} + \{\beta\}$. Therefore the Boolean algebra of alternatives consists of four elements: 1) the alternative $\{\alpha\}$ is to climb the tree; 2) the alternative $\{\beta\}$ is to play with a ball; 3) the alternative $1 = \{\alpha, \beta\}$ means that a kid is hesitating what to do; and 4) the alternative $0 = \{\}$ means to take a rest.

We consider that each kid considers his robot as ally and another kid and his robot as the competitors. The kids are subjects a and c, while robots are subjects b and d. The relationship graph is presented in Fig. 4.

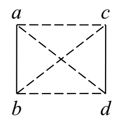

Fig. 4. The relationship graph for robot baby-sitters example

Next we calculate the diagonal form and fold it in order to obtain decision equation for each subject:

$$[ab+cd]^{[ab]^{[a][b]}+[cd]^{[c][d]}} = ab + cd\ .$$

From two actions α and β, action α is a risky action, since a kid can fall from the tree and this is real threat for his health or even life. Therefore according to Asimov's Laws, robots cannot allow kids to start the competition. Thus, robots have to convince kids not to choose alternative $\{\alpha\}$. In terms of alternatives, the Asimov's Laws serve like filters, which excludes the risky alternatives. The remaining alternatives are included into set U. In this case, $U = \{\{\beta\}, \{\}\}$.

Next we discuss some possible scenarios.

Scenario 1. Let every robot tries to convince the kid, it is looking after, to play with a ball, i.e., $b = d = \{\beta\}$: $a = a\{\beta\} + c\{\beta\}$ and $c = a\{\beta\} + c\{\beta\}$.

Then there are the following solutions for kid a.

If $c\{\beta\} = 0$, then $a = a\{\beta\}$. This equation has two solutions 0 and $\{\beta\}$. The equality $c\{\beta\} = 0$ holds, if c equals 0 or $\{\alpha\}$. Therefore, if both robots influence $\{\beta\}$ and kid c influences either 0 or $\{\alpha\}$ on kid a, kid a can choose either to have a rest or to play a ball.

If $c\{\beta\} \neq 0$, then $c\{\beta\} = \{\beta\}$ and $a = a\{\beta\} + \{\beta\} = (a+1)\{\beta\} = \{\beta\}$. The equality $c\{\beta\} = \{\beta\}$ holds, if c equals either 1 or $\{\beta\}$. Therefore, if both robots influence $\{\beta\}$ and kid c influences either 1 or $\{\beta\}$ on kid a, kid a can only choose to play a ball.

The kid c will behave in the same way, if we exchange the roles of kid a and kid c.

Scenario 2. Both robots influence $0 = \{\}$ on both kids. In this case, to have the rest is the only option for the kids: $ab + cd = a0 + c0 = 0 \Rightarrow a = 0$ and $c = 0$.

In the presented example, robots can successfully control kids' behavior by refraining them from doing risky actions.

4.2 Mountain-Climbers and Rescue Robot

We consider that there are two climbers and rescue robot in the mountains. The climbers and robot are communicating via radio. One of the climbers (subject b) got into difficult situation and needs help. Suggest, he fell into the rift because the edge of the rift was covered with ice. The rift is not too deep and there is a thick layer of snow on the bottom, therefore climber is not hurt, but he cannot get out of the rift himself. The second climber (subject a) wants to rescue his friend himself (action α), which is risky action. The second option is that robot will perform rescue mission (action β). In this case, the set U of approved alternatives for robot includes only alternative $\{\beta\}$, since inaction is inappropriate solution according to the First Law. The goal of the robot is to refrain the climber a from choosing alternative $\{\alpha\}$ and perform rescue mission itself. We suggest that from the beginning all subjects are in alliance. The corresponding graph is presented in Fig. 1c and its polynomial is abc.

Next we calculate diagonal form and perform folding procedure:

$$[abc]^{[a][b][c]} = [abc] + \overline{[a][b][c]} = 1 \ .$$

Thus, any subject in the group is in active state. Therefore, group is uncontrollable. In this case, robot makes decision to change his relationship with the climber b from alliance to conflict. Robot can do that, for instance, by not responding to climber's orders. Then the relationship graph transforms into the graph depicted on Fig. 1b and the decisions of subjects are defined by the following equations: $a = (b + c)a + \overline{a}; b = b + (ac + \overline{a})\overline{b}$ and $c = c + (ab + \overline{a})\overline{c}$ (see section 2).

The choice of climber a is defined by the interval $(b + c) \supseteq a \supseteq 1$. Therefore, climber a is capable of choosing only alternative $1 = \{\alpha, \beta\}$, if the condition $(b+c) = 1$ is satisfied. Thus, next climber a can realize one of the alternatives $\{\alpha\}$ or $\{\beta\}$. This case is not acceptable, since climber a can realize risky alternative $\{\alpha\}$. On the other hand, if $(b+c) \subset 1$, then climber a is in frustration and cannot make any choice.

Therefore the only way to refrain climber a from choosing alternative $\{\alpha\}$ is to put him into frustration state.

Next we consider various options of climber's b influence on climber a. Let at first he makes influence to send the robot (alternative $\{\beta\}$). In this case, if $(b + c) \subset 1$, then climber a gets into frustration and cannot make any decision. Therefore the influence of robot c on the climber a should be $\{\beta\}$, as well. Then $(b + c) = (\{\beta\} + \{\beta\}) = \{\beta\} \subset 1$, and climber a cannot make any decision.

If climber b makes influence $\{\alpha\}$ on climber a, then robot has to make influence $\{\alpha\}$ on climber a, as well. Then $(b + c) = (\{\alpha\} + \{\alpha\}) = \{\alpha\} \subset 1$, and climber a cannot make a decision.

Next, we illustrate that regardless of climbers' simultaneous (joint) influences on the robot, it can realize alternative $\{\beta\}$, thus, completing the rescue mission.

Here four scenarios of climbers' joint influences on the robot c are considered.

Scenario 1. Climbers a and b make influences $\{\alpha\}$ and $\{\beta\}$, respectively. Then $a = \{\alpha\}, b = \{\beta\}$: $1 \supseteq c \supseteq \{\alpha\}\{\beta\} + \overline{\{\alpha\}} \Rightarrow 1 \supseteq c \supseteq \{\beta\}$. Therefore robot can choose any alternative, which includes alternative $\{\beta\}$. In this case, $D = \{\{\alpha, \beta\}, \{\beta\}\}$ and $U = \{\{\beta\}\}$, consequently, $DU = \{\{\beta\}\}$. Therefore robot will choose alternative $\{\beta\}$.

Scenario 2. Climbers a and b make influences $\{\beta\}$ and $\{\alpha\}$, respectively. Then $a = \{\beta\}, b = \{\alpha\}$: $1 \supseteq c \supseteq \{\beta\}\{\alpha\} + \overline{\{\beta\}} \Rightarrow 1 \supseteq c \supseteq \{\alpha\}$. Therefore robot can choose any alternative, which includes alternative $\{\alpha\}$: $D = \{\{\alpha, \beta\}, \{\alpha\}\}$. Since $U = \{\{\beta\}\}$, $DU = \{\}$. According to the control schema in Fig. 3, robot will be choosing from alternatives in set U. Therefore, robot will choose alternative $\{\beta\}$.

Scenario 3. Both climbers make influences $\{\alpha\}$. Then for robot c, $a = b = \{\alpha\}$: $1 \supseteq c \supseteq \{\alpha\}\{\alpha\} + \overline{\{\alpha\}} \Rightarrow c = 1$ and $D = \{\{\alpha, \beta\}\}$. Since $U = \{\{\beta\}\}$, $DU = \{\}$. Thus as in the previous scenario, robot will choose alternative $\{\beta\}$.

Scenario 4. Both climbers make influences $\{\beta\}$. Then for robot c, $a = b = \{\beta\}$: $1 \supseteq c \supseteq \{\beta\}\{\beta\} + \overline{\{\beta\}} \Rightarrow c = 1$ and $D = \{\{\alpha, \beta\}\}$. Since $U = \{\{\beta\}\}$, $DU = \{\}$ and robot will choose alternative $\{\beta\}$.

The discussed example illustrates how robot can transform uncontrollable group into controllable one by manipulating the relationships in the group. In the controllable group by its influence on the human subjects, robot can refrain the climber a from risky action to rescue climber b. Robot achieves its goal by putting climber a into frustration state, in which climber a cannot make any decision. On the other hand, set U of approved alternatives guarantees that robot itself will choose the option with no risk for humans and implement it regardless of climber's influence.

5 Discussion and Conclusion

In the present study we have shown how the Reflexive Game Theory merged with Asimov's Laws of robotics can enable robots to refrain the human beings from doing risky actions. The beauty of this approach is that subjects make decisions themselves and consider these decisions as their own. Therefore the RGT fused with Asimov's Laws play a part of social buffer, providing safer life for people with no extensive psychological pressure on human subjects. To our knowledge up to date there has been no similar approach proposed.

The first example illustrates how robots can filter out the risky action and choose both active and inactive alternatives. Here robots are not required to perform any actions. The second example describes critical conditions. In this case, the inactive alternative cannot be chosen and robot has to take the burden of performing risky action. The example with kids illustrates less intense conditions of robot's inference application, while the second example requires more sophisticated multistage strategies: 1) to change the group's configuration; 2) to refrain a climber from trying to go on a rescue mission himself; and 3) to perform the rescue mission by robot itself. The first example plays a role of introductory passage from the kids' yard games to the real life situation in the mountains, which requires more powerful calculus than the first one. The proposed approach based on fusion of the RGT with Asimov's Laws shows its capability of managing successfully the robots' behavior in either situation.

This allows to make the next step in human and robot integration. The RGT provides human-like decision-making system, thus enabling robots to track the human decisions and influence on the people in the way humans are perceptive to. The RGT presents the formal and general model of the group. This model enables to compute the choices of human subjects in the form of feasible algorithms and apply the reflexive control accordingly. Therefore, the RGT is the core part enabling robots to "think" like humans. The Asimov's Laws are the basis for filtering out the risky actions from the existing alternative. Thus, robots programmed with fusion of the RGT and Asimov's Laws are the tools to create risk free psychologically friendly environments for human beings. It opens prospectives of creating the robotic agent, capable of behaving itself like human beings on both motor and mental levels. The primary goal the robots thought to be used for was to substitute the human beings in deadly environments or in the case, when human abilities are not enough, for instance, when extremely high precision is needed. The capability of "thinking" like humans opens new frontiers to the outer limits of the human nature.

The core idea of this study is to show how human-like robots can refrain human beings from doing the risky actions by using the RGT and Asimov's Laws. Therefore the questions of development of required interfaces to extract Boolean algebra directly from the environment are not discussed. The present study also is not answering the technical issues as software and hardware implementation. We consider these questions as the future research trends. This study only shows how human mental world can be represented in the form of feasible RGT algorithms, then fused with Asimov's Laws and implanted into robots' mind. We hope the result of this fusion is a one small step towards making our life safer and our world a better place.

References

1. Lefebvre, V.A.: The basic ideas of reflexive game's logic. Problems of systems and structures research, 73–79 (1965) (in Russian)
2. Lefebvre, V.A.: Lectures on Reflexive Game Theory. Cogito-Centre, Moscow, (2009) (in Russian)
3. Lefebvre, V.A.: Lectures on Reflexive Game Theory. Leaf & Oaks, Los Angeles (2010)
4. Lefebvre, V.A.: Reflexive analysis of groups. In: Argamon, S., Howard, N. (eds.) Computational models for counterterrorism, pp. 173–210. Springer, Heidelberg (2009)
5. Batchelder, W.H., Lefebvre, V.A.: A mathematical analysis of a natural class of partitions of a graph. J. Math. Psy. 26, 124–148 (1982)
6. Lefebvre, V.A.: Algebra of Conscience. D. Reidel, Holland (1982)
7. Asimov, I.: Runaround. Astounding Science Fiction, 94–103 (March 1942)