In Game-Theoretic Models in Mathematical Ecology / Editors: V. Mazalov, D. Novikov, G. Ougolnitsky, L. Petrosyan. – New York: Nova Science Publishers, 2015. – P. 167 – 174. ISBN 978-1-63483-489-6.

MODELS OF REFLEXIVE GAMES IN CONTROL PROBLEMS OF ECOLOGICAL-ECONOMIC SYSTEMS

D.A. Novikov and A.G. Chkhartishvili
(Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences, Moscow, Russia)

We substantiate the feasibility and reasonability of employing the framework of reflexive games for describing decision-making and control problems of ecological-economic systems.

Keywords: ecological-economic system, reflexive game, awareness structure, informational equilibrium.

Introduction

From the game-theoretic viewpoint, the control problem of an ecological-economic system [14, 15] whose elements can demonstrate a purposeful behavior [3] consists in the following. A control subject (a Principal) has to design a game of controlled subjects (agents) with some rules [5, 8] so that its outcome appears most beneficial to the former [7, 13]. Therefore, a necessary step of such control problems concerns game-theoretic analysis allowing a Principal to forecast the response of a controlled system to certain control actions.

Reflexive games [12] represent a method of game-theoretic modeling with due consideration of agents' complex awareness (particularly, their mutual awareness [1, 2, 6]). Nowadays, reflexive games have found wide application in the description of awareness, joint decision-making of agents and solution of associated informational control problems (control of agents' awareness structures) in different fields, namely, corporate management, economics, marketing, political science, etc. [4, 9, 10, 11, 12].

Agents' awareness in a reflexive game is defined by a structure comprising their beliefs about essential parameters of a current situation and the beliefs of their opponents (other agents). The solution of a reflexive game is an informational equilibrium, *viz*., a set of actions chosen by real and phantom agents (the ones existing in the minds of real agents), where each agent maximizes a goal function based on its awareness.

This paper focuses on several models of ecological-economic systems, exploring how the outcome of agents' interaction depends on their awareness structure (an informational equilibrium). If agent's awareness is false (i.e., the agent possesses wrong beliefs about the game conditions), then the result observed by it can be either unexpected or meet the expectations. The latter being the case for all agents, we obtain a stable informational equilibrium [12] whose existence conditions are examined below.

1. "The number of agents on a market"

Consider n homogeneous (identical) agents numbered by the elements of the set $N=\{1, \ldots, n\}$. Agents choose nonnegative production outputs $x_i \geq 0$ and have the goal functions

(1) $f_i(x) = x_i - (x_i)^2/2 - \frac{\chi}{n}\sum_{j \in N} x_j$,

where $x = (x_1, x_2, \ldots, x_n)$ and $\chi \geq 0$ means a penalty coefficient. The first summand in the expression (1) corresponds to the agent's proceeds from product sales at unit price. The second summand answers for the agent's costs, whereas the third summand plays the role of penalties for environmental pollution (we believe that the penalty for the total pollution proportional to the total production output is equally shared by all agents). Suppose that the expression (1) forms the common knowledge of all agents and the system adopts the following sequence of moves. Agents simultaneously and independently choose their production outputs, and then a Principal reports to each of them the penalty imposed.

If the number of agents is common knowledge, then under such awareness each agent chooses the action

(2) $x_i^* = 1 - \chi / n$,

which maximizes its goal function (1).

Consider possible cases of agents' awareness about their number n. If each agent believes that the number makes up $\hat{n}$ and this is common knowledge, then each agent expects to get the following penalty:

(3) $\frac{\chi}{\hat{n}}(1-\frac{\chi}{\hat{n}})\hat{n} = \chi(1-\frac{\chi}{\hat{n}})$.

By observing the actual value of its penalty

(4) $\frac{\chi}{n}(1-\frac{\chi}{\hat{n}})n = \chi(1-\frac{\chi}{\hat{n}})$,

none of the agents doubts the correctness of its beliefs (as far as the right-hand sides of the expressions (3) and (4) do coincide). Hence, the informational equilibrium (2) is stable under any (particularly, false) beliefs $\hat{n}$ of the agents (their common knowledge) about the number n. The stated property directly follows from the fact that the penalty is proportional to the mean action of the agents. As a result, the influence of their number gets "compensated" by the total action.

Concluding this section, let us emphasize an important aspect. The conclusion regarding the stability of any (particularly, false) equilibria does not depend on the parameter χ of the penalty scheme. In other words, the model under consideration admits no transition to a true informational equilibrium by varying the penalty scheme: it is necessary to apply informational impact on agents that modify their individual and/or mutual awareness.

2. "Joint production"

Consider n enterprises (agents) operating in a region and manufacturing homogeneous products. The goal function of enterprise i has the form

(5) $f_i(x) = \lambda x_i - \frac{x_i^2}{2(r_i + \beta X_{-i})} - \chi x_i$,

where $X_{-i} = \sum_{j \neq i} x_i$, $x_i \geq 0$ denotes the admissible action of agent i, $r_i > 0$ means its type, $\lambda > 0$ is the unit price of the products, $\chi \geq 0$ indicates a penalty coefficient, and β corresponds to a nonnegative parameter.

According to the expression (5), the costs of each agent depend on the actions of other agents (e.g., via technology transfer).

Suppose that formula (5) and the values of all incorporated parameters form the common knowledge of the agents. Set $X = \sum_{j=1}^{n} x_j$ and $R = \sum_{j=1}^{n} r_j$. Recall that a Nash equilibrium is a set of agents' actions such that the action of each agent maximizes its goal function (under fixed actions of the rest agents). In the present case, we find a Nash equilibrium using the first-order necessary optimality conditions. Construct the derivative of the goal function (5) with respect to the agent's action x_i. Trivial transformations bring to the following formula:

(6) $$x_i = \frac{(r_i + \beta X)(\lambda - \chi)}{1 + \beta(\lambda - \chi)}.$$

Next, sum up the expressions (6) over all agents and get the total action:

(7) $$X = \frac{R(\lambda - \chi)}{1 - \beta(\lambda - \chi)(n-1)}.$$

By substituting (7) into (6), we finally obtain that

(8) $$x_i^* = \frac{\lambda - \chi}{1 + \beta(\lambda - \chi)}\left(r_i + \frac{\beta R(\lambda - \chi)}{1 - \beta(\lambda - \chi)(n-1)} \right).$$

Obviously, increasing the price and/or the number of agents enlarges the total equilibrium production output; on the other hand, higher "penalty strength" χ reduces the output (see (7)).

Now, study the case when agents' beliefs about the number n and the quantities $r = \{r_i\}$ can be false. Since agent i knows its type and action, it easily calculates

$$\beta X_{-i} = \beta(X - x_i) = \frac{r_i \gamma}{1+\gamma} + \frac{\gamma}{1+\gamma} \cdot \frac{R}{1 - \gamma(n-1)}$$

(see (5)). Here $\gamma = \beta(\lambda - \chi)$ for the sake of compactness.

The last formula shows the following. If an agent treats the quantities $\hat{n}$ and $\hat{R}$ as the real values of the number of agents and the sum of their types, then the equilibrium is stable [12] provided that

(9) $$\frac{\hat{R}}{1-\gamma(\hat{n}-1)} = \frac{R}{1-\gamma(n-1)}.$$

The condition (9) guarantees that the value of the goal function coincides with the agent's expectations.

3. "Threshold penalties"

This section is dedicated to a model reflecting the following situation. Agents are penalized if their total result $\sum_{i \in N} x_i$ (the joint efforts on environmental protection against industrial waste pollution) is smaller than a given threshold. Penalties are sufficiently large to make production unprofitable; therefore, agents have to avoid penalization (or suspend production). In the absence of penalization, agent i receives the income V_i, $i \in N$.

The result represents an increasing function of the total effort applied by all agents. Assume that agents can have nonidentical beliefs about the parameters of this function, *ergo* about θ such that no penalization takes place under $\sum_{i \in N} x_i \geq \theta$.

For agent $i \in N$, implementing an action $x_i \geq 0$ incurs the costs $c_i(x_i, r_i)$, where $r_i > 0$ designates the agent's type (a parameter describing its individual characteristics).

Suppose that the cost functions $c_i(x_i,, r_i)$ enjoy continuity, increase in x_i, decrease in r_i and $c_i(0, r_i) = 0$, $i \in N$.

Let X' be the set of all admissible action profiles $(x_1, \ldots, x_n)$ and define the set of individually rational actions of the agents:

$$IR = \{x \in X' \mid \forall\ i \in N\ \ V_i \geq c_i(x_i, r_i)\}.$$

Readers can easily observe that $IR = \prod_{i \in N} [0; x_i^+]$, where

$$x_i^+ = \max \{x_i \geq 0 \mid c_i(x_i, r_i) \leq V_i\},\ i \in N.$$

Introduce the notation

$$X(\theta) = \{x \in X' \mid \sum_{i \in N} x_i = \theta\}.$$

Consider different variants of agents' awareness about the parameter $\theta \in \Theta$.

Variant I. The parameter $\theta \in \Theta$ is common knowledge. Then the agents' game has a parametric Nash equilibrium belonging to the set $E_N(\theta) = IR \cap X(\theta)$ (i.e., this equilibrium depends on the parameter θ).

Variant II. Agents' beliefs about the threshold are pairwise different, but the set $\{\theta_i\}$ forms common knowledge (the so-called asymmetrical common knowledge).

Without loss of generality, number agents so that their beliefs increase: $\theta_1 < \ldots < \theta_n$. Here the structure of admissible equilibria is described by

Assertion 1. Suppose that $\theta_i \neq \theta_j$ under $i \neq j$. Depending on the relationship of the parameters, an admissible informational equilibrium is the following $n + 1$ action profiles: $\{ x^* \mid x_i^* = 0,\ i \in N\}$; $\{x^* \mid x_k^* = \theta_k$, $x_i^* = 0,\ i \in N,\ i \neq k\}$, $k \in N$. In a practical interpretation, either all agents do nothing, or only agent k applies its efforts by choosing the action θ_k.

In the general case, we have $\theta_1 \leq \ldots \leq \theta_n$ (i.e., agents' beliefs can coincide). This possibly leads to a similar equilibrium domain as in Variant I. In other words, in an equilibrium the efforts are applied by agents with an identical belief about the threshold.

Variant III. Agents' beliefs about the threshold differ, but each agent considers the game with the asymmetrical common knowledge (generally speaking, its beliefs about the opponents' beliefs are false). Here the set of admissible equilibrium action profiles becomes the largest possible one: $\prod_{i \in N} [0; x_i^+]$. Moreover, we easily establish

Assertion 2. For any action profile $x^* \in \prod_{i \in N} [0; x_i^+]$, there exists an awareness structure such that each agent subjectively plays the game with the asymmetrical knowledge and the vector x^* is a unique equilibrium.

Assertions 1 and 2 are argued by analogy to the assertions in [11, Section 4.10].

4. “Principals’ interests coordination”

Consider an ecological-economic system comprising one enterprise (agent) and two Principals. As its strategy, the agent chooses a production output $x \geq 0$ and a safety level $y \geq 0$, which incur the costs $x^2 / 2r$ and $y^2 / 2w$, respectively ($r > 0, w > 0$). Each Principal gains some “income” from agent’s activity (described by a function $H_i(u, y)$) and pays some “incentive” $\sigma_i(x, y)$ to the agent, $i = 1, 2$. Therefore, the goal function of Principal i takes the form

(10) $\Phi_i(\sigma_i(\cdot), x, y) = H_i(x, y) - \sigma_i(x, y)$,

whereas the goal function of the agent is defined by

(11) $f(\{\sigma_i(\cdot)\}, x, y) = \lambda x - x^2 / 2r - y^2 / 2w + \sigma_1(x, y) + \sigma_2(x, y)$.

This ecological-economic system possesses the following sequence of moves. The Principals simultaneously and independently choose the incentive functions and report them to the agent. Next, the latter chooses its action. Further analysis gets confined to the set of Pareto efficient Nash equilibria in the game of the Principals. As shown in [13], their strategies are

(12) $\sigma_i(x', x, y', y) = \begin{cases} V_i, & x = x',\ y = y' \\ 0, & \text{otherwise} \end{cases}, \; i = 1, 2.$

In a practical interpretation, the Principals agree about agent’s joint stimulation for choosing the production output x' and achieving the safety level y'. Such interaction of the Principals is called the cooperation mode [13].

According to the goal function (11), the agent chooses zero safety level in the absence of incentives. Find the optimal production output $x^* = \arg\max_{x \geq 0} [\lambda x - x^2 / 2r] = \lambda r$. The Pareto optimality conditions dictate that the total incentive of the agent from the Principals (in the case of satisfying their recommendations) is defined by

(13) $V_1 + V_2 = \lambda (x^* - x') - [(x^*)^2 - (x')^2] / 2r + y'^2 / 2w$.

The beneficial cooperation condition for each Principal can be stated as follows. In the cooperation mode, each Principal gains a utility not smaller than under its independent stimulation of the agent. The utility of Principal i from “independent” interaction with the agent is

(14) $\Phi^*_i = \max_{x, y \geq 0} [H_i(x, y) - \lambda (x^* - x) + [(x^*)^2 - (x)^2] / 2r - y^2 / 2w]$.

Let

(15) $S = \{x \geq 0, y \geq 0 \mid \exists\ (V_1; V_2) \in \Re^2_+ : H_i(x, y) - V_i \geq \Phi^*_i,\ i = 1, 2;$

$$V_1 + V_2 = \lambda\,(x^* - x) - [(x^*)^2 - (x)^2] / 2\,r + y^2 / 2\,w\}$$

represent the domain of compromise, i.e., a set of agent's actions such that Principals' cooperation is beneficial for their implementation.

Introduce the notation

(16) $\Phi^*_0 = \max_{x, y \geq 0} \{H_1(x, y) + H_2(x, y) - \lambda\,(x^* - x) +$

$$+ [(x^*)^2 - (x)^2] / 2\,r - y^2 / 2\,w\}.$$

By analogy to [3, 13], we can demonstrate that the domain of compromise is nonempty iff

(17) $\Phi^*_0 \geq \Phi^*_1 + \Phi^*_2$.

Consider the following example: $H_1(x, y) = \alpha\,x + (1 - \alpha)\,y$, $H_2(x, y) = (1 - \alpha)\,x + \alpha\,y$, where $\alpha \in [0; 1]$ is a constant reflecting the degree of Principals' interests coordination, i.e., the "proportion" of economic and ecological indicators in their goal functions. If $\alpha = 0$ or $\alpha = 1$, one Principal is interested in economic indicators only (the production output), whereas the other concerns ecological indicators only (the safety level).

Using (14) and (16), we find

$$\Phi^*_1 = r\,\alpha\,[2\,\lambda + \alpha] / 2 + w\,(1 - \alpha)^2 / 2,$$

$$\Phi^*_2 = r\,[\alpha^2 + 1 + 2\,\lambda - 2\,\alpha - 2\,\alpha\,\lambda] / 2 + w\,\alpha^2 / 2,$$

$$\Phi^*_0 = r\,[2\,\lambda + 1] / 2 + w / 2.$$

Moreover, the condition (17) holds true as an identity for any values of the parameters (α, r, w). Consequently, in the current example the domain of compromise is surely nonempty for any awareness and/or mutual awareness of the Principals!

Conclusion

This paper has studied a series of simple models illustrating the feasibility and reasonability of employing the framework of reflexive games for describing decision-making and control problems of ecological-economic systems.

The conducted analysis indicates that the mutual awareness of the members of ecological-economic systems appreciably affects their decision-making. By exerting control actions (i.e., varying such awareness), one can modify the equilibrium states of these systems.

References

1. Aumann R.J. Agreeing to Disagree // *The Annals of Statistics*. 1976. Vol. 4. No. 6. P. 1236–1239.
2. Aumann R.J. Interactive Epistemology I: Knowledge // *International Journal of Game Theory*. 1999. No. 28. P. 263–300.
3. Burkov V.N., Novikov D.A., and Shchepkin A.V. *Control Mechanisms for Ecological-economic Systems*. – Berlin: Springer, 2015. – 174 p.
4. Chkhartishvili A.G. Concordant Informational Control // *Automation and Remote Control*. 2012. Vol. 73. No. 8. P. 1401–1409.
5. Germeier Yu.B. *Non-antagonistic Games*. - Dordrecht, Boston: D. Reidel Pub. Co., 1986. – 331 p.
6. Lefebvre V.A. *Lectures on the Reflexive Games Theory*. – NY: Leaf & Oaks Publishers, 2010. – 220 p.
7. *Mechanism Design and Management: Mathematical Methods for Smart Organizations* / Ed. by Prof. D. Novikov. – NY: Nova Science Publishers, 2013. – 204 p.
8. Myerson R. *Game Theory: Analysis of Conflict*. – London: Harvard Univ. Press, 1991. – 568 p.
9. Novikov D.A. and Chkhartishvili A.G. Information Equilibrium: Punctual Structures of Information Distribution // *Automation and Remote Control*. 2003. Vol. 64. No. 10. P. 1609–1619.
10. Novikov D.A. and Chkhartishvili A.G. Graph of a Reflexive Game and Belles-Lettres // *Studia Humana*, Vol. 3:3 (2014), P. 11-15.
11. Novikov D.A. and Chkhartishvili A.G. Mathematical Models of Informational and Strategic Reflexion: a Survey // *Advances in Systems Science and Applications*. 2014. No. 3. P. 254–277.
12. Novikov D.A. and Chkhartishvili A.G. *Reflexion and Control: Mathematical Models*. – Leiden: CRC Press, 2014. – 298 p.
13. Novikov D.A. *Theory of Control in Organizations*. – N.Y.: Nova Scientific Publishing, 2013. – 341 p.
14. Ougolnitsky G.A. *Sustainable Management*. – N.Y.: Nova Scientific Publishing, 2012. – 287 p.
15. Whitelaw K. *Environmental Systems Handbook*. – Oxford: Elsevier, 2004. – 237 p.