Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

Annales UMCS Informatica AI 7 (2007) 171-180

Annales UMCS
Informatica
Lublin-Polonia
Sectio AI

http://www.annales.umcs.lublin.pl/

The self-learning active problem in dynamic systems

Tomasz Hordjewicz*, Edward Kozłowski

Department of Quantitative Methods, Lublin University of Technology,
Nadbystrzycka 38, 20-618 Lublin

Abstract

The work concerns the theory of adaptive control. By using available tools of contemporary mathematics as well as computer science it is possible in the credible way to introduce the laws of control of a dynamic system as well as to simulate such process in the real world. Therefore we can learn the working of system by simulation, i.e. we can get precise knowledge about the parameters of the system, which were unknown at the beginning of the control process.

1. Introduction

In this paper, the problems relating to control of the learning process of an arbitrary dynamic system in the real time were introduced. Despite numerous works relating to the field of adaptive control, the search of better solutions still inspires the scientists. The above opinion is supported by the huge number of papers dedicated to this subject whose authors, among others, are: Bellman [1], Feldbaum [2], Kulikowski [3], Rishel [4], Harris, Rishel [5], Beneš, Karatzas, Rishel [6]. Some technical aspects connected with the self-learning process were considered in Banek, Kozłowski [7,8,9]. By the above mentioned term „better solution" one should understand a way or an algorithm which allows to obtain more exact results in the considerably shorter time as well as the opportunity of wider applicability in the real world. There exists a certain type of the optimization problems for which one can formulate control laws, but obtaining the explicit solution in the form of the vector of optimal controls is quite difficult. The reason is very large computational complexity of the mathematical algorithms which are too big challenge for the human mind, yet they are necessary to obtain a correct solution. The algorithms which were mentioned above are based on the technique of the backward induction, that is why they

*Corresponding author: *e-mail address*: t.hordjewicz@pollub.pl

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

involve multi-level recurrent processes, where the equation including expressions of multiple integration is solved on each level.

The above mentioned statement is the reason for considerations over the construction of a computer algorithm on the base of the mathematical one, and the use of the simulation in order to solve problems of the theory of adaptive control. To do this, the parallel programming technique and using computers joint in a computational cluster turned out to be helpful.

In this paper, the solution of optimization problems for which the process of learning unknown parameters of the system is the key element was introduced. As a tool characterizing the state of knowledge about the parameters of the system, conditional entropy was chosen, which is undoubtedly one of the best tools used for the measurement of knowledge (see e.g. Banek, Kozłowski [8,9], Saridis [10]). The energetic expenditures of controlling the system in the conditions of complete and incomplete information about the system were compared additionally.

2. Adaptive control

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a complete probability space. Let $\omega_1,\ldots,\omega_N$ be a sequence of independent m-dimensional random vectors on Ω with normal distribution $N(0,I_m)$, y_0 be an initial state with distribution $P(dy_0)$ and let ξ be k-dimensional vector with a priori distribution $P(d\xi)$. All these objects are assumed to be stochastically independent. Define $\mathcal{F}_k \triangleq \sigma(y_0)\vee \sigma(\xi)\vee\sigma\{\omega i: i = 1,2,\ldots,k\}$ and set $\mathcal{F} = \mathcal{F}_N$.

We will consider the adaptive control problem for the system with state equation:

$$y_{i+1} = f\left(\xi, y_i, u_i\right) + \sigma\left(\xi, y_i\right)\omega_{i+1}, \tag{1}$$

where $i = 0,1,\ldots,N-1$, $y_i \in \mathbb{R}_N$, $f:\mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^n, \sigma:\mathbb{R}^k \times \mathbb{R}^n \to \mathbb{M}(n,m)$, where $\mathbb{M}(n,m)$ is the set of $n \times m$ matrices. The functions f, σ are assumed to be continuous in all their variables.

On $(\Omega,\mathcal{F},\mathbb{P})$ we define a family of σ-subfields $\mathbb{Y}_j = \sigma\{y_i: i = 0,1,\ldots,j\}$. A vector $u_j \in \mathbb{R}^l$ measurable with respect to $\mathbb{Y}_j$ is called a control action and $u = (u_0,u_1,\ldots,u_{N-1})$ an admissible control. The class of admissible controls is denoted by U.

To specify the aim of control, we introduce loss functions g_i, $i = 0,1,\ldots,N-1$. We assume that $g_i:\mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$ and heredity function $h:\mathbb{R}^k \times \mathbb{R}^{n\times(N+1)} \times \mathbb{R}^{l\times N} \to \mathbb{R}$ is continuous and bounded. The task is to find:

$$\inf_{u\in U} J\left(u\right), \tag{2}$$

where

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

$$J(u)=E\left[\sum_{i=0}^{N-1} g_i\left(\xi,y_i,u_i\right)+h\left(\xi,y_0,u_0,\ldots y_{N-1},u_{N-1},y_N\right)\right]. \tag{3}$$

Theorem 1. *Suppose that the functions* g_j, $j = 0,1,\ldots,N-1$ *and* h *are continuous and bounded*, f, h and g_j, $j = 0,1,\ldots,N-1$ *are continuously differentiable in* u *and* $\det\Sigma(\xi,y) \neq 0$ *for* $(\xi,y) \in \mathbb{R}^k \times \mathbb{R}^n$, where $\det\Sigma(\xi,y) = \sigma(\xi,y)\sigma^T(\xi,y)$. If u^* is an optimal control then

$$\begin{aligned}&E\Big\{\nabla_u g_j\left(\xi,y_j,u_j^*\right)+\nabla_u h\left(\xi,y_0,u_0,\ldots y_{j-1},u_{j-1},y_j,u_j^*,\ldots,y_{N-1},u_{N-1}^*,y_N\right)\\&+\left(\sum_{i=j+1}^{N-1} g_j\left(\xi,y_i,u_i^*\right)+h\left(\xi,y_0,u_0,\ldots y_{j-1},u_{j-1},y_j,u_j^*,\ldots,y_{N-1},u_{N-1}^*,y_N\right)\right)\\&\times\left(y_{j+1}-f\left(\xi,y_i,u_i^*\right)\right)^T\Sigma^{-1}\left(\xi,y_j\right)\nabla_u f\left(\xi,y_j,u_j^*\right)\Big|\mathbb{Y}_j\Big\}=0\end{aligned} \tag{4}$$

for all $j \in \{0,1,\ldots,N-1\}$.

The proof of this theorem we can see in Banek, Kozłowski [9].

3. Conditional entropy

Let $p(\cdot)$ and $p_0(\cdot)$ be the a priori distributions of the random vector ξ and the state vector y_0 respectively and suppose that the density of the joint distribution of (ξ,y_0) is

$$\mu_0\left(\xi,y_0\right)=p(\xi)\,p_0\left(y_0\right).$$

By induction it is easy to obtain (see e.g. Banek, Kozłowski [8,9]) the following recurrence formula for the density of the joint distribution of $\mu_i(\xi,y_0,y_1,\ldots,y_i)$:

$$\mu_i\left(\xi,y_0,y_1,\ldots,y_i\right)=\mu_{i-1}\left(\xi,y_0,y_1,\ldots,y_{i-1}\right)\gamma\left(y_i-f\left(\xi,y_{j-1},u_{j-1}\right),\Sigma\left(\xi,y_{i-1}\right)\right),$$

where

$$\Sigma(\xi,y)=\sigma(\xi,y)\sigma^T(\xi,y)$$

and

$$\mu_N\left(\xi,y_0,\ldots,y_N\right)=p(\xi)\,p_0\left(y_0\right)\prod_{j=0}^{N-1}\gamma\left(y_{j+1}-f\left(\xi,y_j,u_j\right),\Sigma\left(\xi,y_j\right)\right),$$

where

$$\gamma(x,\mathrm{m},\mathrm{Q})=\frac{1}{\sqrt{(2\pi)^k\det(\mathrm{Q})}}\exp\left(-\frac{1}{2}[x-\mathrm{m}]^T Q^{-1}[x-\mathrm{m}]\right).$$

Conditional density is

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

$$\mu_N\left(\xi \middle| y_0,u_0,\dots,y_{N-1},u_{N-1},y_N\right)=\frac{\mu_N\left(\xi,y_0,\dots,y_N\right)}{\int \mu_N\left(x,y_0,\dots,y_N\right)dx}.$$

Hence the conditional entropy of parameter ξ is

$$H\left(\xi \middle| y_0,u_0,\dots,y_{N-1},u_{N-1},y_N\right)=E\left[-\ln \mu_N\left(\xi \middle| y_0,\dots,y_N\right)\right]$$

$$=-\int \ln\left(\frac{\mu_N\left(\xi,y_0,\dots,y_N\right)}{\int \mu_N\left(x,y_0,\dots,y_N\right)dx}\right)\frac{\mu_N\left(\xi,y_0,\dots,y_N\right)}{\int \mu_N\left(x,y_0,\dots,y_N\right)dx}d\xi \tag{5}$$

$$=\ln\left(\int \mu_N\left(x,y_0,\dots,y_N\right)dx\right)-\frac{\int \ln\left(\mu_N\left(\xi,y_0,\dots,y_N\right)\right)\mu_N\left(\xi,y_0,\dots,y_N\right)d\xi}{\int \mu_N\left(x,y_0,\dots,y_N\right)dx}.$$

4. Optimal adaptive control of system with switch

Let us consider the system described by the state equation

$$y_{i+1}=\varphi\left(\xi,y_i\right)+u_i+\omega_{i+1}, \tag{6}$$

where

$$\varphi\left(\xi,y\right)=\begin{cases} a & y<\xi \\ b & y\ge\xi \end{cases}$$

is a function which accepts values *a* or *b* depending on the parameter *y*.

The problem, except for calculation of optimal controls in the next time steps, depends on the identification of the parameter (switching border) whose value influences the state of system (6).

Let ξ has a normal distribution $N(m,\sigma)$. Therefore the joint distribution is

$$\mu_N\left(\xi,y_0,\dots,y_N\right)=\frac{1}{\left(\sqrt{2\pi}\right)^{N+1}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{\left(\xi-m\right)^2}{\sigma^2}+\sum_{i=0}^{N-1}\left(y_{i+1}-\varphi\left(\xi,y_i\right)-u_i\right)^2\right)\right\}. \tag{7}$$

4.1. Control of system in the conditions of incomplete information

Let us consider the following state. We possess a priori knowledge about the parameter ξ which is the border of switching. We control system (6) and in the N steps, we want to get to know most exactly the principle of working of this system (more exactly border of switching ξ) and direct the system to the point A. As a function of losses at time *j* we choose (we bear a certain energetistic expenditures during the control)

$$g\left(\xi,y_j,u_j\right)=u_j^2$$

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

however, the heredity function is the sum of the square of the distance between point A and the final state of the system as well as the conditional entropy of the parameter ξ

$$h\left(\xi, y_0, u_0, \ldots, y_{N-1}, u_{N-1}, y_N\right) = \left(y_N - A\right)^2 + H\left(\xi \middle| y_0, u_0, \ldots, y_{N-1}, u_{N-1}, y_N\right).$$

Then, the task is to find

$$\inf_{u \in U} E\left[\sum_{i=0}^{N-1} u_i^2 + h\left(\xi, y_0, u_0, \ldots, y_{N-1}, u_{N-1}, y_N\right)\right]. \tag{8}$$

The necessary condition for an optimal control of system (6) in the conditions of incomplete information with performance criterion (8) has the following corollary.

Corollary 2. *If u* is an optimal control, then*

$$2u_j^* + E\Big\{\nabla_u H\left(\xi \middle| y_0, u_0, \ldots, y_{j-1}, u_{j-1}, y_j, u_j^*, \ldots, y_{N-1}, u_{N-1}^*, y_N\right)$$
$$+\left(\sum_{i=j+1}^{N-1}\left(u_i^*\right)^2 + \left(y_N - A\right)^2 + H\left(\xi \middle| y_0, u_0, \ldots, y_{j-1}, u_{j-1}, y_j, u_j^*, \ldots, y_{N-1}, u_{N-1}^*, y_N\right)\right)$$
$$\times\left(y_{j+1} - f\left(\xi, y_j\right) - u_j^*\right)^T \Big| \mathbb{Y}_j\Big\} = 0$$

for all $j \in \{0,1,\ldots,N-1\}$, *where*

$$H\left(\xi \middle| y_0, u_0, \ldots, y_{j-1}, u_{j-1}, y_j, u_j, \ldots, y_{N-1}, u_{N-1}, y_N\right)$$
$$= \ln\left(\int \mu_N\left(x, y_0, \ldots, y_N\right) dx\right) - \frac{\int \ln\left(\mu_N\left(\xi, y_0, \ldots, y_N\right)\right) \mu_N\left(\xi, y_0, \ldots, y_N\right) d\xi}{\int \mu_N\left(x, y_0, \ldots, y_N\right) dx}$$

and

$$\nabla_{uj} H\left(\xi \middle| y_0, u_0, \ldots, y_{j-1}, u_{j-1}, y_j, u_j, \ldots, y_{N-1}, u_{N-1}, y_N\right)$$
$$= \frac{\int \ln\left(\mu_N\left(\xi, y_0, \ldots, y_N\right)\right) \mu_N\left(\xi, y_0, \ldots, y_N\right) d\xi \int \nabla_{uj} \mu_N\left(\xi, y_0, \ldots, y_N\right) d\xi}{\left(\int \mu_N\left(x, y_0, \ldots, y_N\right) dx\right)^2}$$
$$- \frac{\int \ln\left(\mu_N\left(\xi, y_0, \ldots, y_N\right)\right) \nabla_{uj} \mu_N\left(\xi, y_0, \ldots, y_N\right) d\xi}{\int \mu_N\left(x, y_0, \ldots, y_N\right) dx}$$

We receive **Corollary 2** from **Theorem 1** directly.

4.2. Control of system in the conditions of complete information

Let us consider the following situation. We possess exact knowledge about the parameter ξ which is the border of switching. We control system (6) and in

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

the N steps, we want to direct the system to the point A. As a function of losses at time j we choose (we bear a certain energetistic expenditures during the control)

$$g(\xi, y_j, u_j) = u_j^2$$

however, the heredity function is the square of the distance between point A and the final state of the system

$$h(\xi, y_0, u_0, \ldots, y_{N-1}, u_{N-1}, y_N) = (y_N - A)^2 .$$

Then, the task is to find

$$\inf_{u \in U} E\left[\sum_{i=0}^{N-1} u_i^2 + (y_N - A)^2\right]. \tag{9}$$

The necessary condition for an optimal control of system (6) in the conditions of complete information with performance criterion (9) has the following corollary.

Corollary 3. *If u* is an optimal control, then*

$$2u_j^* + E\left\{\left(\sum_{i=0}^{N-1}(u_i^*)^2 + (y_N - A)^2\right) \times \left(y_{j+1} - f(\xi, y_j) - u_j^*\right) \middle| \mathbb{Y}_j \right\} = 0$$

for all $j \in \{0, 1, \ldots, N-1\}$.

5. Simulations

Analysing the system state equation, the conclusion arises that because of the nonlinearity, the possibility of the application of the control rights does not exist directly. To prove functionality of the above introduced theory and finding the control rights, the simulations of the work of the system were conducted. The process of the simulation consisted of 3 time steps during which the controls and the values of performance criteria were marked. In the case of control of system with incomplete information, the parameters of the probability distributions were introduced additionally.

For the number of time steps N = 3, simulations were performed on the basis of the following equations in the support of the technique of the backwards dynamic programming:

in step $i = 2$

$$2u_2 + \iint \left[\nabla_{u_2} H(\xi, \mathbb{Y}_2) + \left[(y_3 - A)^2 + H(\xi, \mathbb{Y}_2)\right]\left(y_3 - f(\xi, y_2) - u_2\right)\right] P(dy_3) P(d\xi | \mathbb{Y}_2) = 0$$

in step $i = 1$

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

$$2u_1 + \iiint \Big[\nabla_{u_1} H(\xi, \mathbb{Y}_2) + \big[u_2^2 + (y_3 - A)^2 + H(\xi, \mathbb{Y}_2) \big] \\ \big(y_2 - f(\xi, y_1) - u_1\big) \Big] P(dy_2) P(dy_3) P\big(d\xi \big| \mathbb{Y}_1\big) = 0$$

in step i = 0

$$2u_0 + \int \iiint \Big[\nabla_{u_0} H(\xi, \mathbb{Y}_2) + \big[u_1^2 + u_2^2 + (y_3 - A)^2 + H(\xi, \mathbb{Y}_2) \big] \\ \big(y_1 - f(\xi, y_1) - u_0\big) \Big] P(dy_1) P(dy_2) P(dy_3) P\big(d\xi \big| \mathbb{Y}_1\big) = 0$$

The simulation of control system was conducted in the relationship with the large complexity of the algorithm, which came into being to solve above mentioned equations on the computational cluster.

In order to shorten the time of calculations one should, execute parallelization of algorithm and also write the code of the program in the programming language C++ with the use of the library of the function of programming of parallel MPICH as the implementation of the standard MPI.

On the basis of simulations, the results are presented below.

Table 1. Values of the control parameters and conditional entropy

	ξ = -1	ξ~N(-3,2)	ξ~N(-1,2)	ξ~N(1,2)	ξ~N(3,2)
U[0]	-0.42437	-0.22674	-0.55939	-0.3573	-0.29933
U[1]	1.11231	0.758018	1.28788	1.30166	1.15385
U[2]	0.999673	0.999673	0.999673	0.999673	0.999673
H($\xi \mid \mathbb{Y}_2$)	---	1.75904	1.58654	1.07328	0.783729

Table 1 presents the values of the control parameter on the path $y_0 = 1.5$, $y_1 = 0.8$, $y_2 = 0.4$, for well-known still before the control, the parameter ξ (in the case of control system in the conditions of complete information) and for various a priori distributions of the parameter ξ (in the case of control system in the conditions of incomplete information). Additionally, the values of the system conditional entropy were given (concerning the system in the conditions of incomplete information).

In Fig. 1 the values of conditional entropy were introduced in the aspect of bar chart.

According to the aim of control, which is a minimization of performance criteria, the values obtained as the result of simulation were introduced. The results in Fig. 2 show the cost of self-learning parameters of the system during the control process. The first bar (on the left) represents the value of performance criteria in the system with complete information (at the beginning

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

of the settled parameter ξ), the next one shows the values of performance criteria in the conditions when we do not possess full knowledge about the system before control but we acquire it during the control process.

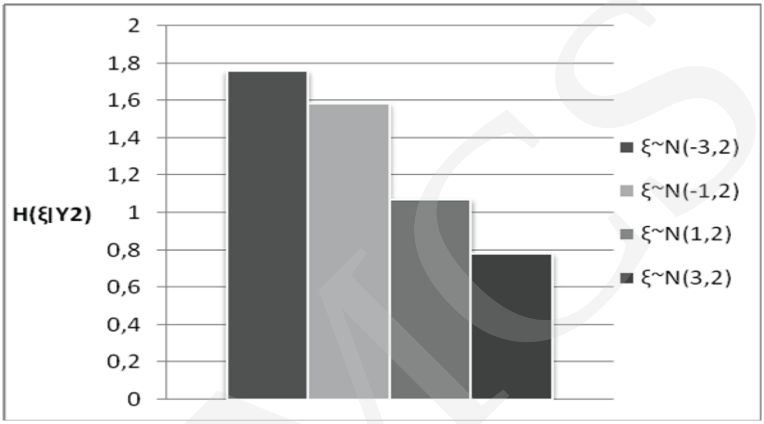

Fig. 1. Values of conditional entropy

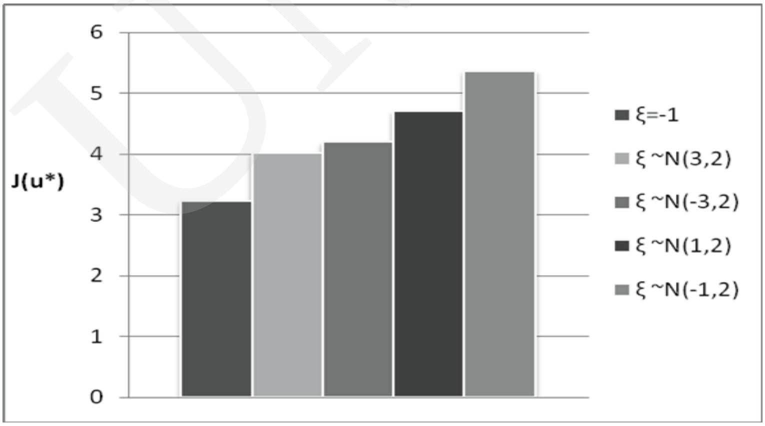

Fig. 2. Values of performance criteria

Below, Fig. 3 present the differences of performance criteria for the system with complete information for $\xi = -1$ and the performance criteria for individual a priori distributions of parameter ξ for the system with incomplete information.

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

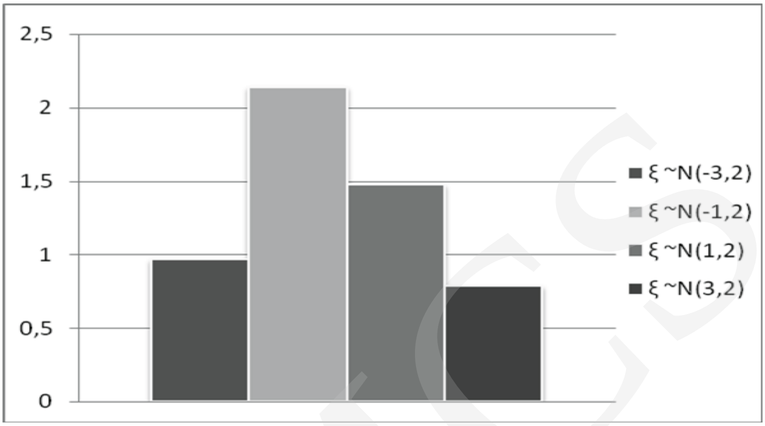

Fig. 3. Differences of values of performance criteria for the systems with complete and incomplete information

Conclusions

In the paper, the optimal control law of a nonlinear dynamic system was formulated and the conditional entropy of the system was also determined. The task which was considered in the paper presents the problem in which, besides determining the control, the cost caused by the losses resulting from the control should be minimized. In the case when the problem has a linear form with respect to the parameter of the control, one can give the solution ‘directly’ in some cases. Unfortunately, in the case when the state equation contains a non-linear expression, the way of solving is to perform simulation of the control process, taking into account the previously determined condition of the solution optimality.

In the described example, in order to perform the simulation, the knowledge from the field of computer science, and, more precisely, programming was indispensable. When intending to perform the simulation process, first, an algorithm was created and later implemented in a programming language. Because of great complexity of the algorithm, a program simulating the behavior of the system working on a computational cluster was created. This allowed to decrease the time of calculations considerably and to enlarge the precision of the obtained results as well. Unfortunately, the program used the power of parallel computers only at a given moment. The possibility of parallelization of calculations exists on every time level which will result in considerable acceleration of calculations.

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 29/06/2026 14:38:26

References

[1] Bellman R., *Adaptive control processes*, Princeton, (1961).
[2] Feldbaum A.A., *Optimal Control Systems*, Academic Press, (1965).
[3] Kulikowski R. *Procesy optymalne i adaptacyjne w układach regulacji automatycznej*, Warszawa, PWN, (1965).
[4] Rishel R., *A nonlinear discrete time stochastic adaptive control problem. Theory and applications of nonlinear control systems*. Sel. Pap. 7th Int. Symp. Math. Theory Networks Systems, (1985) 585.
[5] Harris L., Rishel R., *An algorithm for a solution of a stochastic adaptive linear quadratic optimal control problem*. IEEE Transactions on Automatic Control, 31 (1986) 1165.
[6] Beneš V.E., Karatzas I., Rishel R., *The separation principle for a Bayesian adaptive control problem with no strict-sense oprimal law*. Stochastic Monographs 5, (1991) 121.
[7] Banek T., Kozłowski E., *Adaptive control of system entropy*. Control and Cybernetics, 35(2) (2006) 279.
[8] Banek T., Kozłowski E., *Active and passive learning in control processes application of the entropy concept*. Systems Science, 31(2) (2005) 29.
[9] Banek T., Kozłowski E., *Self- learning jako zadanie sterowania dualnego*. *Systemowo-komputerowe wspomaganie zarządzania wiedzą*, ed. Kulikowski R., Bubnicki Z., Kacprzyk J., Akademicka oficyna Wydawnicza EXIT, (2006) 157.
[10] Saridis G., *Entropy formulation of optimal and adaptive control*. IEEE Transactions on Automatic Control, 33 (1988) 713.