Coherent Adaptive and Rational Action Selection Problem Solution for Mobile Robots

Received 10th March; accepted 26th March.

Piotr Szynkarczyk

Abstract:

The article deals with the issue of controlling the autonomous mobile systems. Considered are problems focusing around the use of distributed architecture controllers. In contrary to hierarchical architecture, all layers of distributed controller are operating simultaneously, and have an access to both: the sensors (receptors), as well as to the executing systems (effectors) of controlled autonomous mobile system.

Simultaneous operation of all layers of the distributed architecture controller entails arising of so called action selection problem. This problem consists in continuous taking of decisions: which one of the layers, or which layers of controller, at given moment should have an influence on the movement of controlled object. Discussed is the proposal of a new solution of this problem. Presented are results of experiments with the use of mobile robot.

Keywords: *robot, data processing, distributed controller*

1. Behavior of Autonomous Mobile Systems

1.1 Preliminary definitions

The behaviors of autonomous mobile system may take different form. In particular, as far as the application of distributed architecture controllers is concerned, among many features they may show, the most essential seems to be:

coherence behavior,
rational behavior,
adaptation abilities.

Quoted hereunder are the selected and adopted in further considerations definitions describing these features.

Coherence behavior: in case of the conflict between the execution of decisions made by some layers of controller, the autonomous mobile system shall not behave in an „undecided manner", i.e. it shall not try to execute different goals, depending on insignificantly varying states of environment.

Rational behavior: in every situation the decision on the behavior of autonomous mobile system shall be undertaken with all knowledge contained in controller taken into consideration, and with all goals or tasks taken into consideration. In another words: in the case when the autonomous mobile system is in the situation when passing over of control of its effectors to one of its layer will result in achieving of one of the autonomous mobile system goals, then such passing over of control shall take place.

Another definition is given in [7]. According to it, the rational behavior consists in the meeting of the following requirements:

Compatibility: the autonomous mobile system, due to irrational control shall not try to execute simultaneously two mutually contradicting actions, e.g. the mobile robot can not try simultaneously to travel forward and backward.

Common currency: in the case of the occurrence of conflict between different layers of controller, i.e. when they are tying to execute conflicting actions (which leads to the loss of compatibility), then in order to provide possibility of rational selection, there must be certain common currency for all layers, defining how the execution of given action is important at the time being.

Consistency: the problem of rationality of given autonomous system can also be expressed in the following way: in the case when being twice in the same situation (understood as a state of environment and the internal state of controller) the autonomous mobile system each time takes different actions, then it can not be regarded as rational.

Transitivity: the action selection is effected on the basis of the currency common for all layers. In the case when selected is the action A, which is regarded as more important that B (that is: $A > B$), and if the action B is recognized as more important than C (that is: $B > C$), then in the case of selection between A and C, selected must be action A (that is: $A > B > C$ must take place).

In the further part of the text [7] one can also find the statement, that coherent behavior is the necessary condition for the occurrence of rational behavior.

In this article, an analysis was made with the use of both rationality definitions, since they are not contradictory and are mutually completing each other.

While studying the behaviors of autonomous mobile systems, the limitations of maintaining of rational behavior should be kept in mind. They are limited due to:

limited possibility of data processing (e.g. limited volume of memory, limited speed of computer processing unit, etc.),
limited possibility of obtaining the data on the state of environment resulting from the finite number of receptors (e.g. distance sensors) with finite accuracies, ranges, etc.,
limited knowledge about the proper rules of action in given situation (anticipation of the consequences of one's actions on the grounds of finite knowledge base),

limited possibilities of the execution of control commands control with zero error is not possible.

In connection with the above it is possible to obtain only so called bounded rationality [9]. The autonomous mobile system can act in a rational way only within the range of certain limitations. From the point of view of external observer (without all or part of these limitations) the loss of actions rationality can take place [8].

The adaptive behavior: in the case of actions in the unknown and dynamic environment, the autonomous mobile system shall reveal adaptation abilities if it is in a position to change its behavior of adapting itself to the variable states of environment.

The additional manifestation of adaptation could be also quicker reactions of controller at the repeated states of environment.

In the case of discussed in this paper class of controllers subjected to the adaptation can be the control layers, or the manner their co-operation is organized, i.e. the action selection method.

2. Distributed Controller with a New Solution to the Action Selection Problem

2.1 Observations of the nature as the premises for construction of adaptive distributed controller

Distributed Architecture

Control separation of individual elementary behaviors into separate, isolated layers is available for observation in the nature in the case of living organisms. It was proven that frog is jumping towards detected insect, using two separate reflexes. One of them allows for making the jump of adequate length, and the second is controlling its direction. The experiments described in [4] show that the frog with damaged part of brain, which is identified as being responsible for the direction of jump, when hunting is jumping straight ahead regardless of the fact, at what angle it may see the insect, but it jumps by proper distance. In [4] described are also tests of the behavior of humans with brain damages. In one of the cases, two patient women had the problems with the distinguishing and taking objects from the shelf in shop. However, their problems were completely different. One of them was unable to indicate any difference in the shape of the observed objects. However, she was in a position to grab correctly each of them, so that they were not falling out of her hand. This would have been not possible, should she be unable to match her grip properly to the shape of objects. The second patient was able to describe in a faultless way each of the objects, however she had a serious problem how to grab them properly in result these objects were falling out of her hands. The complex actions consist than of the set of more simple and independent behaviors. The parts of brain that control these behaviors may even not communicate internally with each other. Both above cases allow to suppose, that the biological control systems, so far the most effective out of those known, have the structure of distributed nature, similarly like the solution suggested in the article.

The Role of Preliminary Data Processing

The data processing in WPD modules (WPD: modules of preliminary data processing, as proposed later in this paper) is effected in order to carry out specialist reduction of information. Each of the layers in order to execute assigned to it elementary task, needs the set of data on the state of environment as well as the internal state of controlled object. Each of the layers requires synthetic description of the situation of autonomous system as seen from different point of view. Here we have an incidence of analogy to biological systems, in which the sensor information is subjected to strong filtration [10]. For example, the eyes of frog are equipped with the movement filters, therefore frogs can see small and mobile objects only [1]. Such simple solution allows for easy differentiation of the dead and motionless insects from living insects, hence mobile. En route of evolution the insects have developed the defensive reflex consisting in lying motionless, defending themselves in this way since they are invisible to frog. The effectiveness of this defensive mechanism is the price which the frogs as a species have to pay for the vast simplification of the image analysis task in their brains. The complication connected with the introduction of non-filtered image analysis would entail the enlargement of brain tissue, hence the increased energy expenditure for its maintenance. This would have forced frogs for more frequent hunting, and therefore should have to pay off in the increase of their effectiveness thanks to a new system of insect detection.

Another nature example of filters superimposed on the signals from receptors can be the case of sleeping squirrel. When the wind is moving branches, it rains, or other noises arise squirrel is sleeping firmly. If, however marten is climbing on the tree trunk, causing the sounds that are more delicate than the rain - the animal reacts immediately [1].

The same receptors are used for different purposes. In the case of frog the same eyes, which are used for hunting insects, are allowing for the differentiation of day from night, seeking partners, avoiding of dangers. Similarly the WPD modules enable the multiple use of the same sensors by different layers of controller. Each layer of controller is receiving an information on the state of environment that is essential only for the correct execution of assigned to it task.

Phenomenon of Adaptation

Under established conditions the individual reflexes (or behaviors) are initially ranked according to certain importance. The priorities are making up a certain prototype of typical behaviors, e.g. more important is the escape from predator than having rest after meal. However, such ranking is changing along the changes within the state of environment this is the adaptation.

The phenomenon of adaptation to the conditions prevailing in environment can be manifested by the behavior consisting in certain kind of "concentration of attention". Such situation can take place for example when in the environment of a given living organism appears large quantity of food (at a time when it is hungry it is an internal state of organism, that can be compared to the case of discharged batteries supplying mobile robot with energy). Then it "pays more attention" to reaching food reflex of seeking the food and reaching it becomes the dominant reflex. The "concentration of attention" on actions connected with taking of food shall be the greater, the more intensive are the external stimuli connected with food presence in the environment and the internal stimuli intensity of hunger. The purpose of such mechanism of the "concentration of attention" is in this case showing of maximum effectiveness in getting of food. However, when predators appears in the environment, the reflexes of hiding and escape from predators are activated and begin to be dominant. In spite of this, it may happen then, that if the exemplary living organism approaches the food sufficiently close to it, so that the intensity of stimuli connected with the presence of food becomes very high, and at the same time sensation of hunger is strong - then activated is behavior connected with seeking of food and reaching to it.

In addition, it was noticed that in the case of living organisms, the intensity of stimulus has great significance for the promptness of reaction to it. Often the more intensive it is, the quicker stimulated by it reflex (or behavior) is activated.

When comparing two living organisms of different history the first one, in the environment in which there was large number of predators, and the second, in the environment in which there was a large quantity of food, but no predators at all then probably in the presence of food the first organism shall be more "careful", that is shall react more quickly for the appearance of predators, on the other hand the second shall be "more concentrated" on food. This is the manifestation of adaptation each of these organisms has adapted itself to the conditions in its environment. Should the conditions change then organisms shall adapt to them anew. This phenomenon was found during the research on the behavior of fish [6]. Documented was the dependence of behaviors on the history of given specimens: it was proven that given fish specimens from the species of bleak, coming from the population exposed to the pressure of pikes, are reacting in a stronger way to the presence of alarm substance than those, that have never encountered it. Fish synthesizes the chemical alarm substance, being released to the water environment only in the case their injuring. For other fish it is a warning signal of danger.

The dependence of behavior on the intensity of stimuli was described in [1] on the example of birds. Observations were concerning the behavior of vultures when eating the carrion. Their group is usually divided into three categories: actually eating, standing directly close to food, but being not admitted to it, and the third category: distant observers. As it turns out, between these groups occurs incessant rotation, the birds are promoted in turn in the direction of eating group. Mechanism of this phenomenon is explained as follows [1]: in result of waiting the sensation of hunger is increasing among the birds, and at the same time the tendency to attacking. In particular the birds waiting close by food are intensely stimulated by the proximity of food. At the same time the eating specimen have already satisfied initial hunger in a great hurry and their tendency to fight is decreasing. From this follows an interesting dependence: in a brawl the winning are not the strongest birds, but those that are actually the most hungry and close to food. Similar example can be found in [6], where described are behaviors of fish. It was found that the delay of reaction in relation to the detection of signal presence depends on the concentration of alarm substance. At the low concentrations this reaction time grows longer.

2.2 Description of the architecture of adaptive distributed controller

The individual layers of controller in proposed solution are by no means internally coupled with each other. The intelligent behavior is not written inside of controller in a form of procedures anticipating all possible situations, but is arising out on the contact point of their actions with the complex external real world. The results of action, that is the change of state of environment and in consequence the change of the internal state of controller, make up the only information channel between the layers of controller.

It is all different with WPD modules that can influence each other (mainly by mutual suppression). The reason for introduction of such mechanism is the existence of specific circumstances, in which actions of some layers (hence the execution of some actions) is completely needless or illogical. For example if the task of robot is collecting of minerals, then when its container is full, then seeking of further minerals becomes completely needless. However, if the robot container is empty, then proceeding to the storeroom is also irrational. The rational behavior of autonomous system in the cases of this kind, can be initially ensured by the introduction of mutual suppression of WPD modules. This is a solution similar to the „subsumption" type architecture, however two fundamental differences appear here:

- the purpose is not obtaining of the complex behaviors by creation of the chain of simpler reflexes, but initial exclusion of the layers, the actions of which under given circumstances are completely needless or irrational,
- the mutual influences take place not between the layers of controller, but between self-contained WPD modules.

From control layers arbiter is receiving an information on required by them control of individual degrees of freedom of autonomous system, and also is receiving signals from decisive neurons. Each of the layers and each of decisive neurons, thanks to supplying them with sorted by WPD modules information on environment and the internal state of autonomous system, are especially active when it is required by arising situation. This is facilitating the rational action selection. However remains to be resolved the problem of the conflict arising when at the same time two or more decisive layer-neuron sets are generating strong signals that is when it is necessary to take decision which of the actions of similar importance at given moment should be executed. The additional difficulty is retaining at the same time of coherence and such structure of controller ensuring that it is adaptive.

These problems are resolved in the proposed method of action selection. The general outline of the architecture of proposed controller is presented in Fig. 1.

WPD modules, decisive neurons and arbiter are jointly making up the decisive mechanism. The decisive mechanism to certain extent is behaving like a neuron network with the number of so called decisive neurons corresponding to the number of layers in the controller. Each such neuron has one input and one output. The signals on the neuron inputs are the signals coming from WPD modules corresponding to the layers. Signals on neuron outputs are deciding which layer at given moment shall control behavior of autonomous mobile system. The output signals are the result of the product of input signals and corresponding to them weights. As a control layer at given moment selected is by arbiter the one of the strongest output signal of attributed to it neuron. Such signals, in the following part of article are called activations.

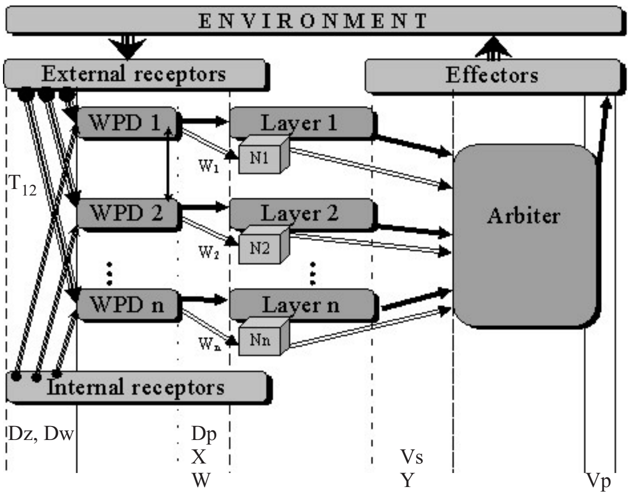

Fig. 1. Proposed architecture of controller.

Explanations to Fig. 1.:

Dz - data collected by WPD modules from external receptors,
Dw - data collected by WPD modules from internal receptors,
Dp - data from receptors, processed by WPD modules for layers,
X - data from receptors, processed by WPD modules for decisive neurons
W - input weights of decisive neurons ($w_1 \ldots w_n$ - weights),
T_{nm} - mutual influence of WPD modules with the numbers n and m,
Vs - control settings with the values generated by individual layers,
Y - output signals of decisive neurons,
Vp - resulting settings of control,
External receptors - receptors collecting data from the autonomous system environment,
Internal receptors - receptors collecting data on internal state of autonomous system,
$WPD_{1 \ldots n}$ - modules of preliminary data processing,
$Layer_{1 \ldots n}$ - layers of controller,
N1...n - decisive neurons,
Arbiter - the module making final action selection.

The values of signals from WPD modules, are resulting mainly from the state of environment, however they are also indirectly dependant on the history, through the influence on them of the internal state of controlled object. The values of weights are changing incessantly. At given moment they are resulting mainly from the history of action of autonomous system, but also to some extent from current situation. These weights consist of two elements:

1. ***Constant part*** - initially attributed to given layer in accordance with the principle of *priority inversion*. The layers of higher and higher levels - executing more and more complex tasks - have generally lower and lower priorities. Suggestion is to call this phenomenon the "*inversion of priorities*". Manifestations of this pheno-menon can be also observed in the actions of human minds. Usually when somebody gets hungry, or when something is worrying him, he is not in a position to think in a creative and abstractive way. This is a manifestation of existing gradation of the importance of reflexes (or behaviors) in living organisms. The most important is ensuring of the basic needs (e.g. safety, food). Only then, when they are ensured - active become remaining reflexes (or behaviors), e.g. satisfying of curiosity.
2. ***Adaptive part*** - result of the incessant process of adaptation of the decisive mechanism.

Taking of decision on the basis of calculated in this way activations is compatible with so called „principle of letting through the highest value for behavior tendency" (discoverer of the principle: Nobel prize winner K. Hartline) [2]. This means that, if in the motivation system of any creature (or autonomous system) two contradicting with each other behavior tendencies are aiming at the activation, then in contrary to completely nonsensical and non adequate to situation behavior (e.g. continuous hesitation, that is the loss of coherence), winning is the tendency with the strongest motivation. This principle is ensuring for the autonomous system the coherence of behaviors.

2.3 The structure and operation of an adaptive distributed controller executing the exemplary task

The Exemplary Task

The exemplary task consists in the collection of the „objects" from „storeroom" and successive transporting them „home". At the same time robot has to avoid collisions with the obstacles and take care for an appropriate state of charging its batteries. Should the batteries be discharged to certain level, then robot should arrive in the „battery charging station". The robot as a complex mechatronic system should monitor the degree of wear of its more important mechanical parts. Should this degree of wear become adequately large, then the robot should proceed to „service station" and when being there turn off itself in order to enable its overhaul and replacement of parts. In the discussed research all service operations are simulated.

The controller executing exemplary task, consists of nine layers (numbered from 0 up to 8):

0. Layer of behaviors connected with avoiding of collisions with obstacles.
1. Layer of the reflex of travel towards "storeroom" (or in other words: towards the „goal").
2. Layer of the reflex of travel towards "home".
3. Layer of the reflex of travel towards "battery charging station".
4. Layer controlling the behavior of robot when turns up in „storeroom" grasping the „object" by manipulator.
5. Layer controlling the behavior of robot when turns up at „home" - putting away the „object" from manipulator.
6. Layer controlling the behavior of robot when turns up in „battery charging station" charging of batteries.
7. Layer controlling behavior of robot when turns up in „service station" turning itself off.
8. Layer of the reflex of travel towards "service station".

WPD Modules - Data Processing for the Needs of Control Layers

The individual layers of exemplary controller are using, through WPD modules the two different kinds of sensors (receptors). These are the state of environment sensors and the autonomous system internal state sensors. The WPD modules make up a certain kind of filters. Their task is a specialist reduction of information, so that the control layers could receive synthetic description of the state of environment, or internal state. For each kind of layer, these sensors can be different, next in some cases different WPD modules can use common sensors.

Similarly like in the case of environment state sensors, the information from sensors of autonomous system internal state is processed by WPD modules. The internal state are for example readings of position and speed of individual degrees of freedom of robot manipulator. Such information can also be filtered, e.g. by giving by WPD module of nonzero output signal, only then when the position, speed or acceleration go beyond certain assumed ranges. All depends on the structure and tasks of given control layer.

WPD Modules - Data Processing for the Needs of Decisive Mechanism

The WPD modules are processing also the information for the needs of decisive mechanism. Also in this case taken into consideration are: state of environment and internal state of autonomous system. The results of such processing are two signals, one of them representing external situation, and the second representing an internal situation.

The signal representing the state of environment, hereinafter called an ***external stimulus***, is the calculated within the states of environment space, distance between the current state of environment, and the state that is intended (or being avoided depending on the kind of task of given layer). This calculation is made for each layer separately, since due to the use of different data from different sensors, each of them is functioning in another space of environment states. For the control layer of mobile robot aiming at some goal this can be the distance to this goal expressed in meters. In this case the space of states is two-dimensional. The space of external states can have any number of dimensions, and this number results from the number of variables, being the data defining the state of environment. For example this can be a number of readings from sensors used for the determination of the distance from the obstacles around the mobile base of robot. In the process of controlling, taken into consideration can be the values other than only those measured in meters. This could be an information expressed in other units, like e.g. seconds, Volts, grams, etc.

The signal representing internal situation, hereinafter called ***internal imperative***, is a calculated in the space of internal states distance between the current internal state, and the state to the arising of which given control layer (using given WPD module) is aiming at, or is trying to avoid it (depending on the kind of the task of given layer). Similarly like in the case of external stimulus, this calculation is made for each layer separately, since due to the use of different data from different sensors, each of them is functioning within another space of internal states. For layer the task of which is to take care of charging the batteries, it shall be the difference between their current state, and the state of full charge. In this case the space of internal states is mono dimensional.

Similarly like in the case of external stimulus, the space of internal states can have any number of dimensions, and this number results from the number of variables making up the data defining the internal state.

Finally, when the values of external stimulus (symbol: bz) and internal imperative (symbol: iwew) are known, calculated can be the signal x making up the representation of the data from receptors processed by WPD module for the needs of decisive neurons. This is described by general formula (1):

$$x = \frac{iwew}{bz} \qquad (1)$$

Signal x makes up the input signal for decisive neuron corresponding to the given WPD module.

The example of using the above assumptions in practice can be the action of WPD modules for the part of reflexes of travel towards given goal in the controller executing exemplary task. The external stimuli bz are defined as the functions determined by formula (2).

$$bz_n = \sqrt{(x_r - x_c)^2 + (y_r - y_c)^2} \qquad (2)$$

where:
bz_n - external stimulus for layer No n (n=1,2,3,8)
x_r - coordinate x of robot
y_r - coordinate y of robot
x_c - coordinate x of goal
y_c - coordinate y of goal

The internal imperatives for the part of decisive reflexes of travel towards given goal are determined with the use of function described by formula (3).

$$iwew_n = iwew_{n\max}(1 - e^{-\beta_n \ i}) \qquad (3)$$

where:

$iwew_n$ - internal imperative for layer number n (n=1,2,3,8)

iwewnmax - maximum value of internal imperative for layer number n (n=1,2,3,8)

n - growth rate of signal for layer number n, n > 0, n=(1,2,3,8)

i - variable of the character of time passage from the last presence in given goal

Such functions, with different values of parameters n and iwewnmax for different layers, are defining:

1. For n=1 (layer of the reflex of travel towards "storeroom") iwew1 resulting from the time that has passed from the last presence in „storeroom" (or in other words: „goal").
2. For n=2 (layer of the reflex of travel towards "home") iwew2 resulting from the time that has passed from the last presence at „home".
3. For n=3 (layer of the reflex of travel towards "battery charging station") iwew3 resulting form the simulated degree of batteries discharging.
4. For n=8 (layer of the reflex of travel towards "service station") iwew8 resulting from simulated degree of the wear of mechanical parts of robot.

Using of the function defined by formula (3) for the calculation of the intensity of signals iwew, enables taking into consideration of certain circumstances of long term character. E.g. in the case of layer of travel towards „battery charging station" we are obtaining the effect of iwew growth when the batteries are being „discharged". On the other hand, when they are „charged", the value iwew is suddenly decreasing. Hence the iwew variable is measuring the distance between current state and the state when the batteries are fully „charged".

For n=1, 2 and 3 the parameter iwew of temporal character is also the metric measuring in the space of time the distance between the current state and intended state (which given control layer is aiming at). E.g. for aiming at „home" the intended state is reaching the goal, that is the moment when the time from last presence in this goal is equal to zero.

In all above cases the spaces of internal states are mono-dimensional, and for the measurements used is an internal clock of real time, installed in the on board computer of robot.

The parameters $iwew_{max}$ were selected in a way to reflect the degrees of importance of given layers actions, what in certain sense is the equivalent of priorities of these layers. This is of importance for example in the case when it is necessary to make the action selection between the layers of aiming at „battery charging station" and at „home", and when both values of iwew are maximum. More important is charging of batteries (priority inversion) and so for this layer $iwew_{max}$ is larger.

Decisive Mechanism

The decisive mechanism consists of three kinds of elements: WPD modules in the decisive part, decisive neurons and arbiter. This structure is presented in Fig. 2. being the reminder of the part of Fig. 1.

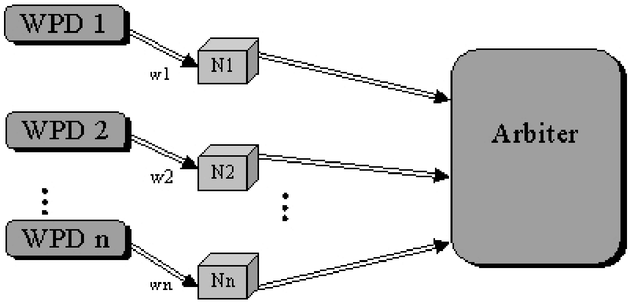

Fig. 2. The structure of decisive mechanism.

Explanations to Fig. 2.:

x_i - signal from WPD modules for decisive neurons (i=1...n),

w_i - input weights of decisive neurons (w1...wn - weights, i=1...n),

y_i - output signals of decisive neurons (i=1...n),

n - number of layers in the controller,

$WPD_{1...n}$ - modules of preliminary data processing,

$N_{1...n}$ - decisive neurons,

Arbiter - module making final action selection.

The output signals from WPD modules, being the input signals x for the decisive neurons are obtained on the basis of calculations according to general formula (1). Every decisive neuron has one input and one output. Each input signal x has corresponding weight w. Their product gives in result corresponding to them output signal y (Fig. 2.), called the activation of decisive neuron (or more generally: layer activation). The function of the transition of decisive neurons is an identity function. At given moment as an acting layer, selected is by arbiter this one which has the largest activation.

The weight of each input signal of decisive neuron consists of two terms (formula (4)):

$$w = w_{st} + w_a \qquad (4)$$

where:

w_{st} - constant part connection weight

w_a - adaptive part of connection weight

Constant parts of weights have to determine preliminary preferences and degrees of action importance of given control layers, e.g. obstacle avoiding layer must have initially the largest weight (for the sake of priorities inversion). The constant parts of weights are the equivalent of „preliminary tuning" of behaviors hierarchy. As shown by laboratory experiments, their values have mainly influence on the rate of reaction of autonomous system to external stimuli.

In the experimental part, tested were five methods of decisive neuron learning (changes of the values of the decisive neurons weights adaptive parts). These methods are described by the below mentioned formulas.

Hebb's classical method

$$w_i(k) = wi(k-1) + \eta\, xi(k) - yi(k) \quad (5)$$

where:
- $w_i(k)$ - new weight obtained in the k-step of learning
- k - learning step number k
- i - number of decisive neuron
- $w_i(k-1)$ - previous weight
- η - learning coefficient
- x - input signal to decisive neuron (signal from WPD module)
- y - activation of neuron (y = x * w)

Hebb's method with forgetting

$$w_i(k) = w_i(k-1) + \ x_i(k)\, y_i(k) + \eta\, w_i(k-1) - \gamma\, y_i(k) \quad (6)$$

where:
γ - so called coefficient of forgetting

„Instar" method

$$w_i(k) = wi(k-1) + \eta\, (xi(k) - wi(k-1)) \quad (7)$$

„Outstar" method

$$w_i(k) = w_i(k-1) + \eta\, (y_i(k) - w_i(k-1)) \quad (8)$$

Oja's Method

$$w_i(k) = w_i(k-1) + \eta\, y_i(k)\, [\, x_i(k) - w_i(k-1)\, y_i(k)\,] \quad (9)$$

Let us assume that mobile robot executing the exemplary task, at given moment is moving forward and in effect of this movement is approaching the obstacle. The closer it is, the smaller becomes signal of external stimulus bz corresponding to avoiding of obstacles. At the same time the signal of internal imperative iwew has the value resulting from the history of the movement made so far. Signal x (input signal for decisive neuron, x = iwew / bz where bz is an external stimulus for layer) is growing. Simultaneously, at a time of approaching the obstacle, controller is undergoing the next cycles of learning. Along with the growth of signal x, growing is also weight w, hence larger and larger is activation y of this layer (y = w * x). At the same time, remaining layers are acting being subjected to the external and internal stimuli, provided the WPD layers are transferring to them such signals. However, the layer of avoiding the obstacles has the strongest activation, since due to the proximity of obstacle signal x is large. The layer of avoiding obstacles is therefore taking over the control over robot. When in result of the actions of this layer the obstacle is avoided, the activation of obstacle avoiding layer is suddenly decreasing. Then, depending on the history of robot movement, the actions of controller layers and influencing them external and internal stimuli, the control over robot movements is taken over by another one of control layers. During the whole period of autonomous system operation, the incessant action selection process is taking place. The thesis assuming that in the effect of such mechanism we are obtaining the rational, coherent and adaptive behavior, has been proven on the grounds of experiments with the real, mobile robot Nomad 200 used as the object subjected to control.

3. Adaptive Distributed Controller - Results of Laboratory Experiments

3.1 Experimental stand

All described experiments have been carried out in the laboratory, with the use of mobile robot Nomad 200 and the distributed controller simulation program for the execution of exemplary task. The photograph of typical arrangement of experimental track is shown in Fig. 3.

Fig. 3. Typical arrangement of experimental track.

Photograph (Fig. 3.) shows marked on laboratory floor goals to be reached by robot within the framework of the execution of exemplary task. Seen are also carton boxes making up the obstacles placed on the path of robot. The research has been carried out on three different routes differing from each other in the arrangement of goals and obstacles placed on the way towards them. Goals A, B, C, D in various experiments with exemplary task were becoming „home", „goal" („storeroom"), „battery charging station" (in short „station"), „service". Table 1. shows co-ordinates of goals and their different functions depending on the route.

Table 1. Co-ordinates of goals and their different functions depending on route.

Goal co-ordinates in [m]	Function on route No 1	Function on route No 2	Function on route No 3
(0, 0)	„home"	„home"	„home"
(2.8, -1.54)	„goal"	„service"	„station"
(2.8, 0.74)	„station"	„goal"	„service"
(1.15, 2.5)	„service"	„station"	„goal"

3.2 Comparison of learning methods

During one of series of experiments compared were actions of controller at the use of five various learning methods, according to formulas from (5) up to (9). Operation of controller at the use of learning method „instar" according to formula (7) has been tested and documented in the previous research sub-programs.

The learning methods other than „instar" were requiring that the initial values of adaptation weights are larger than zero. Otherwise their values were not changing and remained on zero level. To avoid this at the beginning of described hereunder experiments all adaptation parts of weights were initially set to value of 0.000001.

The first one out of tested under this sub-program learning methods, was the classical Hebb's method as described by formula (5). In the course of experiments, for =0.1, the activations of layers were growing rapidly up to large values. In practice, from certain moment, the robot started to avoid non-existent obstacles, since the values of responsible for it layer activation was growing most rapidly. Activations of all layers were quickly reaching the numbers considered by on board computer of robot as the infinity (e.g. Fig. 4.). Decreasing of the value was reducing only the rate of growth of controller layers activations.

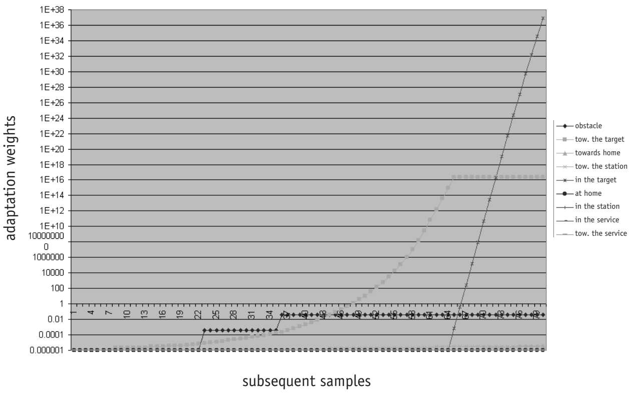

Fig. 4. Adaptation weights in the experiment with the use of classical, Hebb's learning method (2 samples ≈ 1 second).

The second of the tested learning methods of adaptation mechanism was so called Hebb's method with forgetting, as described by formula (6). This method is introducing weights decreasing mechanism at the absence of activation to the degree regulated by so called forgetting coefficient. Unfortunately at the various values of learning and forgetting coefficients, obtained were results very close to the results from the previous series of experiments. All experiments were ended when the robot began to avoid non-existing obstacles. Robot was also crossing consecutive goals not executing the tasks it had to execute after reaching them. As one can see, also this learning method has failed allow proper control of the robot with use of the decisive mechanism, as proposed in the paper.

The next tested method of learning of the decisive mechanism was „outstar" method described by formula (8). During the experiments for =0.1 there were irregularities in the behavior of robot:

1. Obstacles were often rammed.
2. After certain period of experiment duration, robot stopped to came to stop at goals, and execute attributed to them simulated tasks (e.g. charging of batteries and because of that, and soon upon leaving the „station" was returning there repeatedly).
3. At the final stage robot was traveling between „home" and „goal" only.
4. Robot has never reached the „service".

The exemplary route covered by robot is shown in Fig. 5.

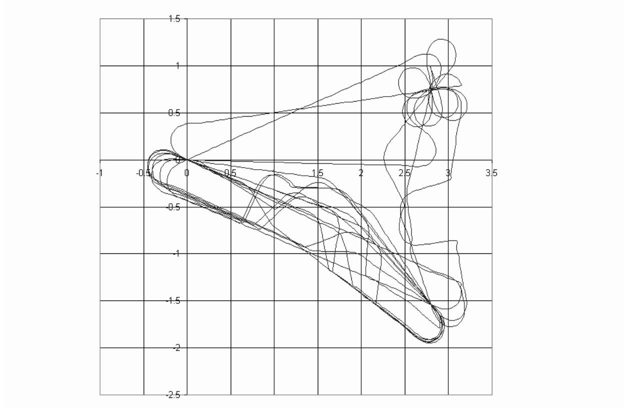

Fig. 5. The route covered by robot during the experiments with the „outstar" type learning method (scale in meters).

Fig. 6. is the graph in logarithmic scale of the run in time of the adaptation weights during the experiment carried out within this series. On the graph one can see that the values of adaptation weights of certain layers (e.g. „at_home" and „in_station") have fallen and grew no more, and the values of others were growing continuously. The fluctuation of their growth enabled the selection of different layers in given moments, however the growth trend remained all the time.

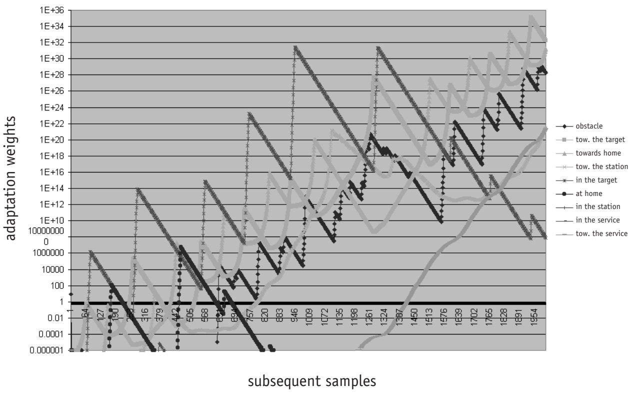

Fig. 6. Adaptation weights in experiments with „outstar" learning method (2 sample ≈ 1 second).

Decrease of the learning coefficient e.g. to the value η=0.01 brought positive effect in a form of correct behavior of robot, similar to one observed in the case of „instar" method. In spite of this, the learning method of „outstar" type decisive mechanism is not making up a good solution for the discussed distributed controller. The reason is the incessant growth trend of adaptation weights. This rate of growth can be regulated by learning coefficient, however this growth is taking place all the time. With adequately long times of controller operation, the weights have to achieve the values regarded by the computer as infinity.

The last tested method of the decisive mechanism learning was Oja's method described by formula (9). In the course of carrying out this series of experiments, occurring were the following irregularities in robot behavior:

1. Robot was avoiding non-existing obstacles.
2. During the travel to given goal, often there was a loss of coherence and so the hesitations between different layers of the aiming at goal reflexes (Fig. 7., vicinity of samples No 55, No 115, No 125).
3. Robot not a single time has stopped upon reaching of given goal, hence has not performed any tasks.

In [3] one can read, that with the Oja's method the fluctuations of weights may happen, which actually took place. In particular, when robot was close to the goal, the adaptation parts of weights were oscillating wildly. In result one may observe in Fig. 8. strong fluctuations of layers activation.

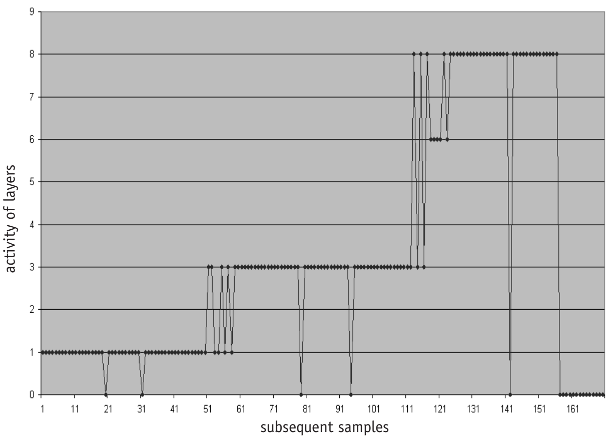

Fig. 7. Activity of layers in the experiment with Oja's learning method (2 samples ≈ 1 second).

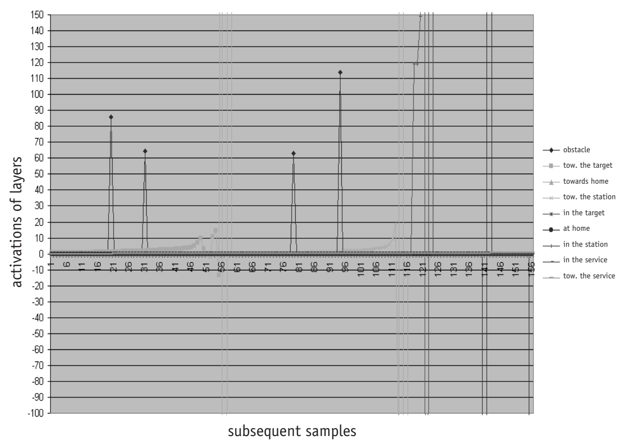

Fig. 8. Layers activation in the experiments with Oja's learning method (2 samples ≈ 1 second).

3.3 Adaptive features of controller

The experiments were carried out on route No 1, with the use of „instar" learning method (formula (7)), The experiments in this research sub-program were consisting in:

1) interruption of experiments at certain moments, manual guiding of the robot into various points of test track and renewed giving back the control of robot movement to its distributed controller,
2) interruption of experiments at certain moments, holding of robot in place and releasing it after the passage of certain time,
3) forcing the robot to remain within certain area of test track with the appropriate use of mobile obstacles,
4) rearranging of obstacles on the way to goals,
5) putting the robot in motion on three different routes (routes No 1, No 2, No 3) without any modification of parameters of distributed controller,
6) various combinations of the a/m factors.

The adaptation phenomenon has manifested itself in two ways within this series of experiments:

1. Controller was immediately adapting itself to the suddenly changed state of environment (e.g. forced by operator new position of mobile base) and was changing its current priorities. Robot depending on the new, external and internal stimuli, as well as its history, was modifying accordingly its current goal. For example, if the state of discharging the batteries was adequately big, and the distance to the „station " in a manually forced new position of robot was adequately small, it was reaching the „station" in spite of fact that before operator's intervention it was traveling to the „storeroom". If however at small distance from „station" the state of batteries discharging is low, robot was continuing its travel towards „storeroom" from a new forced manually position.
2. Controller was modifying its general behavior in the consequence of operator's interference, or even the change of route (switching routes between No 1, No 2, No 3). For example, if during the movement along given route after given time it was reaching the „station", then in a new situation after operator's interruption it was not doing this at the same moment. Controller was executing new plan. Robot was behaving in rational and coherent way, but under completely new conditions caused by the interference of operator. Robot was adapting itself on up-to-date basis with new situation.

The trajectories along which robot was moving were not revealing any irrational behaviors, e.g. repeated return to one of the goals, ramming of obstacles or making circles around without reaching none of the goals. There was no single case of the loss of coherence.

One of the experiments of this series was consisting in the fact, that when robot was first time on its way to the „storeroom", it was directed by hand to the point located closely by the „service" (point A on Fig. 9). However due to the fact that internal imperative for the layer of aiming at „service" was very small at the moment (that is the simulated degree of robot mechanical parts wear was small), in spite of the proximity of this goal robot has renewed its aiming at the „storeroom". After unaided reaching of „storeroom" and „home", robot was again directed in the vicinity of „service” (point B in Fig. 9.). Also this time it has not traveled to this goal, but in return of originally planned „storeroom”, robot was heading towards the „station”. In addition during one of the subsequent travels „home” it was forced (by proper shaping of the obstacle) to travel along longer way avoiding of the obstacle on its left side (point C in Fig. 9.).

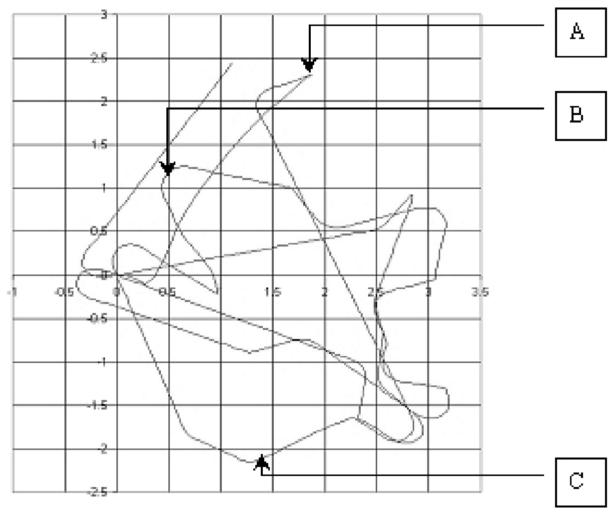

Fig. 9. The route covered by robot (scale in meters).

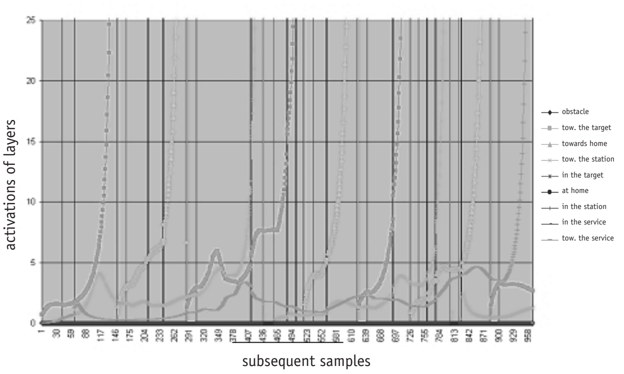

Fig. 10. Activations of controller layers (2 samples ≈ 1 second).

3.4 Problem of action selection

The coherence

The coherence of proposed controller equipped with the decisive mechanism has been proven on the grounds of series of the experiments. The coherence can be tested directly with the use of layers activity graphs. Fig. 11. presents the course of experiment in which the loss of coherence has not taken place, and Fig. 12. the case in which in result of intentional experimental actions the loss of coherence took place (vicinity of points A, B and C).

All discussed experiments, not counting those in which the disturbances of decisive mechanism functioning were intentionally introduced by changing of some of its parameters or learning method, show that the established method of action selection by means of decisive mechanism allows for the achieving of coherent behavior. The coherent behavior was achieved regardless of used experimental track (routes No 1,2,3), of some parameters of controller and of rapid changes of the environment state.

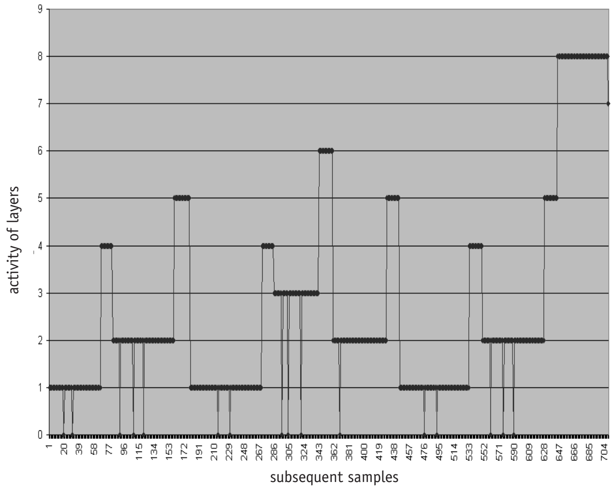

Fig. 11. The coherent activity of layers (2 samples 1 second).

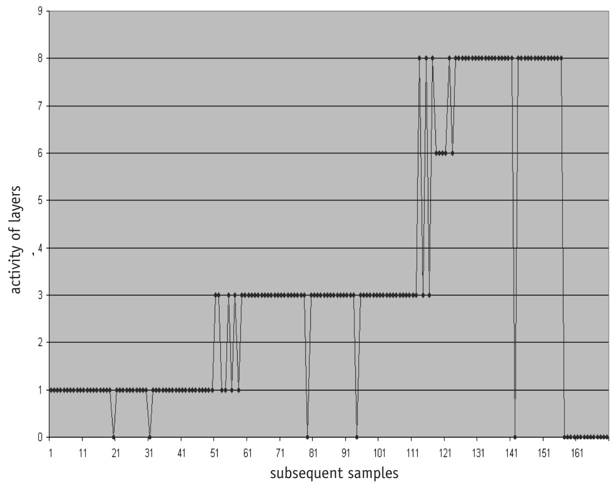

Fig. 12. The loss of coherence: points A, B and C (2 samples ≈ 1 second).

The rationality

The rational behavior of controller is already ensured provisionally by the very method of its operation - by the appropriate designing of WPD modules. By their appropriate design, it was provisionally established in advance which layer of controller be activated at the occurrence of given state of environment and given internal state. There exists certain kind of information filtration, described in paragraphs pertaining to the premises based on nature.

About the selection of individual layers as the controlling layers, deciding are not only WPD modules since these can activate to operation many layers at the same time. The action selection results to main extent from the operation of decisive mechanism. At the decision taking about which layer at given moment shall be controlling the behavior of autonomous mobile system, taken into consideration are both the external circumstances (the state of environment), the internal ones (internal state) as well as present history of actions (adaptation parts of weights). Built within the framework of this research control system belongs therefore to the class of so called situated embodied agents. The behavior of such systems is not basing on any internally complex representation of the state of environment, models of such environment nor the actions of symbolic nature (as e.g. in expert systems). The main role in the operation of such systems are playing the conditions and connections of sensory-motoric type, however these are not the reactive systems only (acting in result of co-operation of simple reflexes). The situated systems are adaptive and have certain history. In result, under identical situations (understood only as the states of environment) they do not have to behave in the same manner. Their behavior depends also on their past and internal states.

Carrying out the analysis of the rationality of autonomous mobile systems it is necessary to keep in mind the phenomenon of so called bounded rationality [9]. Example of the occurrence of this phenomenon can be the situation, in which robot is avoiding the obstacle making excessive number of maneuvers of obstacle avoiding. Oneself can imagine the case in which e.g. the last maneuver was not already necessary should the robot travel straight on to the goal, it would not have collided with the obstacle. From the point of view of man-observer, behavior of the robot is irrational. However, there is no irrational behavior from the point of view of controller executing correctly its program and utilizing in full the data from sensors.

In the experimental research the irrational behavior of robot was encountered. The irrationality of behavior was lying in the fact of e.g.: omitting the execution of tasks to be performed in given goal, lack of coherence of behaviors (it is impossible to talk about rational behavior without coherence), avoiding of non-existing obstacles, ramming of obstacles, making circles around one only goal, execution of selected part of tasks only with full omission of the remaining, the behavior resulting from internal states only without taking into consideration the state of environment. None of these phenomena does occur in the case of controller with the mechanism taught by „instar" method, with properly selected parameters.

In the light of definition given in section 1.1 of this paper, proposed solution of the problem of action selection ensures the rational behavior of autonomous system. It is a direct consequence of described in the article structure of discussed controller. The structure and operation of such controller are the direct realization of the idea as described by the above definition. This has been proven by the experiments.

The rationality of described herein adaptive distributed controller can be also considered in the light of definition according to [7]:

1. Compatibility:
   In connection with the fact, that the behavior of autonomous mobile system is the result of selection of one of control layers as currently controlling its movements, there is no possibility of arising the conflict in a form of attempt to execute simultaneously two contradictory operations, like e.g. simultaneous travel forward and reverse. Such conflicts also do not arise within the confines of individual layers. Consequently the compatibility is always retained.
2. Common currency:
   The common currency are the levels of activations of controller layers, that are then compared by an arbiter. These signals are defined uniquely, and consequently this postulate is met.
3. Consequence:
   The consequence of controller actions is always retained. This follows directly from the fact that, all behaviors of autonomous mobile system are the result of control resulting from internal states only as well as the states of this system environment. No other factors have the influence on the course of control. Consequently, always in the case of the same states of environment and internal states of controller, the behavior of autonomous mobile system is the same.
4. Transitivity:
   The activations of controller layers are expressed numerically. Each such value can be compared with another, in accordance with the basic rules of mathematics. In this way the postulate of transitivity is met by definition.

Since coherence also has been proven, the autonomous mobile system with proposed distributed adaptive controller architecture is rational.

Adaptation

In connection with the existence of decisive mechanism, operation of which is based among other things on the dynamic assigning of decisive parts of weights to the layers, the described controller shows the adaptation features. The results of research carried out can serve as a proof of the above. These results show that the controller is adapting itself to the variable states of environment and is in a position without any adjustments by man to control the movements of robot on various routes (routes No 1, No 2, No 3).

4. Conclusion

In the article presented are problems connected with the issue of action selection in the distributed controllers of autonomous mobile systems. Discussed in the article adaptive distributed controller, equipped with action selection mechanism makes up the solution that is ensuring to autonomous mobile system the coherent and rational behavior with simultaneous showing of adaptation properties. This is proven by the results of experimental research carried out during the realization of exemplary task, as described in the experimental section.

The exemplary task is one out of many possible cases of using proposed methods of action selection. It has been formulated only for the needs of carrying out of the experimental research on real object mobile robot Nomad 200. The usefulness of proposed solution is growing along with the complexity of distributed controller, since it is resolving in an unambiguous way the problem of action selection. If it is possible to meet the required assumptions, then the tasks assigned for distributed adaptive controller can be quite complex.

One of the more interesting conclusions of this work is the observation, that when building more and more complex control systems, the boundary (seemingly distinct) between the hierarchic architecture and distributed

(parallel) architecture is fading away. The problem of action selection in adopted solution is partly resolved by WPD modules (preliminary data processing). The results of decisions made by control layers are transferred to effectors, which in turn have controllers of low level ensuring e.g. maintaining of required rotational speed of motors. Therefore there is point of view, from which each of controller layers (consisting among other things of sensors, WPD module, control layer, low level controllers of effectors) is a hierarchical structure. In the final effect one can look at the system as at hybrid structure distributed system consisting of sub-systems of hierarchical character.
The very important factor influencing the behavior of autonomous mobile system, turned out to be the phenomenon of so called bounded rationality. The influence of this phenomenon can be seen in practice in all discussed cases of the autonomous mobile systems controllers.

AUTHOR
Piotr Szynkarczyk – Head of Intelligent Mobile Systems Division, Industrial Research Institute for Automation and Measurements, Warsaw, Poland, e-mail: pszynkarczyk@piap.pl.

Bibliography and references

[1] Dröscher V. B., *Reguła przetrwania (Rule of Survival)*, Państwowy Instytut Wydawniczy, second edition, Warszawa 1996.
[2] Dröscher V. B., *Białe lwy muszą umrzeć (White Lions Have to Die)*, Państwowy Instytut Wydawniczy, first edition, Warszawa 1997.
[3] Hertz J., Krogh A., Palmer R., *Wstęp do teorii obliczeń neuronowych (Introduction to the Theory of Neuron Calculations)*, Warszawa: Wydawnictwa Naukowo-Techniczne, first edition, 1993.
[4] Ingram J., *Płonący dom (House on Fire)*, Warszawa: Prószyński i S-ka, 1996.
[5] Jacak W., “Roboty inteligentne - Metody planowania działań i ruchów” (“Intelligent Robots – Planning Methods of Actions and Movements”), Wrocław: Publishing of Wroclaw Polytechnic Scientific Works of Technical Cybernetics Institute of WP, series Monographs, 1991.
[6] Jachner A., “Komunikacja chemiczna u ryb słodkowodnych” (“Chemical Communication of Fresh Water Fish”), in *Wiadomości Ekologiczne (Ecology News)*, vol. XLI, no. 3, Dziekanów Leśny: Publishing House of the Polish Academy of Science, 1995, pp. 153-175.
[7] McFarland D., Bösser Th., *Intelligent Behavior in Animals and Robots*, A Bradford Book, The MIT Press, Cambridge, Massachusetts – London, England, 1993.
[8] Russell S.J., „An Architecture for Bounded Rationality”, in: SIGART Bulletin, 1991, no. 2(4), pp. 146-150.
[9] Simon, H.A., „Bounded rationality”, in: J. Eatwell, M. Millgate & P. Newman (ed.): *The New Palgrave: A Dictionary of Economics,* London and Basingstoke: Macmillan, 1988.
[10] Tadeusiewicz R., *W stronę uśmiechniętych maszyn (Towards Smiling Machines)*, Warszawa: Wydawnictwa ALFA, first edition, 1989.