Neural sensor-based navigation of wheeled mobile robot in unknown environment

Zenon Hendzel, Marcin Szuster

Rzeszow University of Technology

Abstract: In presented paper a new approach to a collision-free trajectory generating for a wheeled mobile robot with Adaptive Critic Designs and Fuzzy Logic algorithm, is proposed. The presented discrete hierarchical control system consists of a trajectory generating algorithm based on a reactive navigation of the wheeled mobile robot in an unknown 2D environment with static obstacles, and a tracking control system. A strategy of reactive navigation is developed including two main behaviours: an obstacle avoiding behaviour and a goal-seeking behaviour, realised in a form of Adaptive Critic Design algorithms. These simple, individual behaviours are combined by the fuzzy combiner of behaviours that determines influence of the individual behaviours on the trajectory generation process, according to the environment conditions. The tracking control system is composed of two Dual-Heuristic Dynamic Programming algorithms, the supervisory term and the PD controller. Verification of the proposed control algorithm was realised using the mobile robot Pioneer 2-DX, equipped with one laser and eight sonar range finders, that provides object detection.

Keywords: neural dynamic programming, navigation, wheeled mobile robot, neural networks

1. Introduction

The development of mobile robotics in recent years allowed to increase area of its applications. Simultaneously it made realisation of more complex tasks possible and involved necessity of more complicated control systems development. Increase of the wheeled mobile robots (WMRs) constructions complexity, quantity of information received from the environment, and performance of microprocessors, allowed to design control systems capable of generating a WMR motion trajectory in a real time and modifying it according to the environment conditions, e.g. position of obstacles. There are many different approaches to the problem of planning the path of the WMR, e.g. [1, 3, 4, 7, 10–13], but the most popular are global methods in the known environment and local methods that use sensor based systems and can be applied in the unknown environment. Artificial Intelligence (AI) algorithms, as Neural Networks (NNs) or Fuzzy Logic (FL) systems, are widely use to solve this kind of problems. The development of AI methods allowed to apply Bellman's Dynamic Programming (DP) idea in a form of Neural Dynamic Programming (NDP) algorithms, also known as Adaptive Critic Designs (ACDs) [2, 14–16]. ACDs make generating the sub-optimal control law in forward processes possible.

In the presented article a new approach to a collision free trajectory generating for the WMR Pioneer 2-DX, with usage of NDP algorithms, is proposed. Designed hierarchical control system consists of the trajectory generator, based on ACDs in Action Dependant Heuristic Dynamic Programming (ADHDP) configuration, that generate behavioural control signals in the goal-seeking (GS) and the obstacle-avoiding (OA) tasks, and a FL algorithm, that generates signal used to soft switching behavioural control signals. This approach guarantees generation of the trajectory in the complex task of goal-seeking with obstacle-avoiding, and its realisation using the tracking control system with ACDs in Dual Heuristic Programming (DHP) configuration.

The results of researches presented in the article continue authors earlier works related to the path planning [10, 11] and the tracking control [8, 9] of the WMR using NDP methods. The paper is organised in the following way: the first section includes a short introduction into the WMR path planning problems, the second section presents a discrete model of the WMR dynamics. Next section includes the description of the proposed hierarchical control system, with the path planning algorithm. In following sections there are presented results of experiments realised using the WMR Pioneer 2-DX and summary of the research project.

2. Model of the mobile robot Pioneer 2-DX

The WMR Pioneer 2-DX is composed of two driving wheels, a frame and a third, free rolling castor wheel. The WMR weights $m_R = 9$ kg, its basic dimensions are

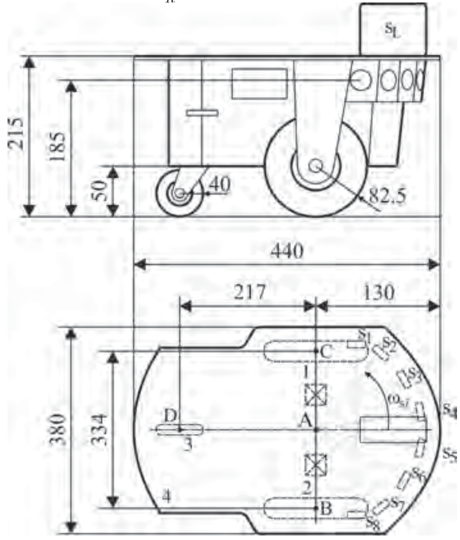

Fig. 1. The wheeled mobile robot Pioneer 2-DX
Rys. 1. Mobilny robot kołowy Pioneer 2-DX

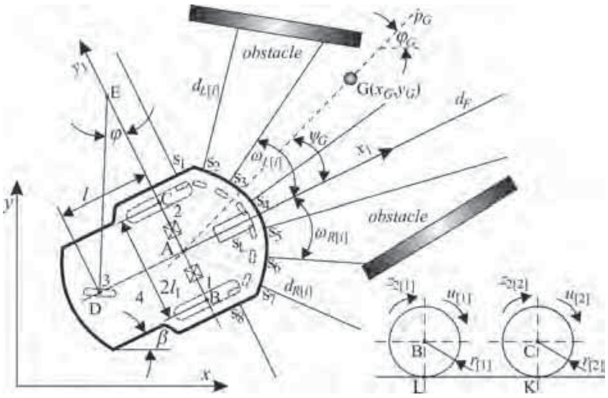

Fig. 2. Scheme of the wheeled mobile robot Pioneer 2-DX in the 2D environment
Rys. 2. Schemat robota mobilnego Pioneer 2-DX w środowisku 2D

shown in fig. 1. Is has eight ultrasonic range finders s_1, ..., s_8 and one laser range finder s_L.

The movement of the non-holonomic WMR Pioneer 2-DX is analyzed in the *xy* plane [5, 6]. The WMR is schematically shown in fig. 2.

The dynamics of the WMR was modelled using Maggie's mathematical formalism [5, 6]. Using Euler's derivative approximation and the state vector $\boldsymbol{z}_{\{k\}} = \left[\boldsymbol{z}_{1\{k\}}^{\mathrm{T}}, \boldsymbol{z}_{2\{k\}}^{\mathrm{T}}\right]^{\mathrm{T}}$, where $\boldsymbol{z}_{2\{k\}}$ corresponds to the vector of continuous angular velocities $\dot{\alpha} = \left[\dot{\alpha}_{[1]}, \dot{\alpha}_{[2]}\right]^{\mathrm{T}}$, we obtained a discrete notation of the WMR dynamics, that can be written in a form

$$\begin{aligned}\mathbf{z}_{1\{k+1\}} &= \mathbf{z}_{1\{k\}} + \mathbf{z}_{2\{k\}}h,\\ \mathbf{z}_{2\{k+1\}} &= -\mathbf{M}^{-1}\left[\mathbf{C}\left(\mathbf{z}_{2\{k\}}\right)\mathbf{z}_{2\{k\}} + \mathbf{F}\left(\mathbf{z}_{2\{k\}}\right) + \tau_{d\{k\}} - \mathbf{u}_{\{k\}}\right]h + \mathbf{z}_{2\{k\}},\end{aligned} \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}(\mathbf{z}_{2\{k\}})$, $\mathbf{F}(\mathbf{z}_{2\{k\}})$ – matrixes and vectors that derive from the WMR dynamics, τ_d – the vector of bounded disturbances, $\mathbf{u}_{\{k\}}$ – the tracking control signal, h – time discretisation parameter, k – index of iteration steps.

The dynamics model of the WMR was described in detail in [5], the closed loop system used in the tracking control system synthesis, was described in detail in [8, 9].

3. Hierarchical control system

The proposed hierarchical control system consists of the tracking control system and the trajectory generator, both build using NDP algorithms. The scheme of the hierarchical control system is shown in fig. 2.

3.1. Tracking control system

The problem of tracking control is defined as searching for the control signal, that minimises tracking errors in the form

$$\begin{aligned}\mathbf{e}_{1\ k} &= \mathbf{z}_{1\ k} - \mathbf{z}_{d1\ k}\,,\\ \mathbf{e}_{2\ k} &= \mathbf{z}_{2\ k} - \mathbf{z}_{d2\ k}\,,\end{aligned} \tag{2}$$

for the desired trajectory $\mathbf{z}_{d\ k} = \mathbf{z}_{d1\ k}^{\mathrm{T}}, \mathbf{z}_{d2\ k}^{\mathrm{T}}\ ^{\mathrm{T}}$, where $\mathbf{z}_{\{k\}} \rightarrow \mathbf{z}_{\mathrm{d}\{k\}}$ when k → ∞, and the control system remains stable. Filtered tracking error $\mathbf{s}_{\{k\}}$ is defined as

$$\mathbf{s}_{\{k\}} = \mathbf{e}_{2\ k} - \Lambda\mathbf{e}_{1\ k}\,, \tag{3}$$

where Λ – a positive defined, fixed diagonal matrix.

In the tracking control system, in detail described in [9], were used ACDs in Dual Heuristic Dynamic Programming (DHP) configuration. The overall tracking control signal

$$\mathbf{u}_{\ k} = \frac{1}{h}\mathbf{M} - \mathbf{u}_{A\ k} + \mathbf{u}_{S\ k}^{*} - \mathbf{u}_{PD\ k} - \mathbf{u}_{E\ k} \tag{4}$$

consists of the ACDs control signal $\mathbf{u}_{A\{k\}}$, the supervisory element control signal $\mathbf{u}_{S\{k\}}^{*}$, the PD control signal $\mathbf{u}_{PD\{k\}}$ and the $\mathbf{u}_{E\{k\}}$ control signal. The supervisory element, derived from the Lyapunov stability theorem, ensures stability, which means that the filtered tracking error $\mathbf{s}_{\{k\}}$ is bounded.

3.2. Trajectory generator

In the presented trajectory generator a strategy of reactive navigation is developed including two main behaviours: OA and GS [1, 3, 10], schematically shown in fig. 4. These simple, individual behaviours are combined by the fuzzy combiner of behaviours (CB), that determines influence of the individual behaviours on the trajectory generation process, according to the environment conditions.

Behavioural control system in the goal-seeking behaviour. Behavioural control signals in the GS behaviour are generated using NDP algorithms in ADHDP configuration, in detail described in [8].

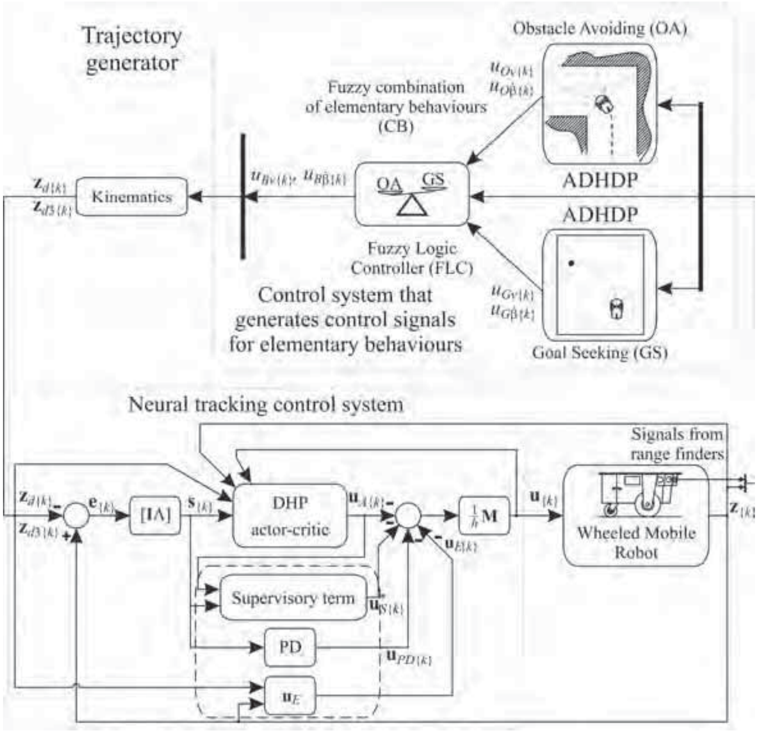

Fig. 3. Scheme of the hierarchical control system
Rys. 3. Schemat hierarchicznego układu sterowania

The objective of the NDP algorithm is to determine the sub-optimal control law, that minimises the value function $V_{\{k\}}$ $(\mathbf{x}_{\{k\}}, \mathbf{u}_{\{k\}})$ [2, 14–16], which is function of the state $\mathbf{x}_{\{k\}}$ and the control $\mathbf{u}_{\{k\}}$ in general case

$$V_{\{k\}}\left(\boldsymbol{x}_{\{k\}}, \boldsymbol{u}_{\{k\}}\right) = \sum_{k=0}^{n} \gamma^{k} L_{C\{k\}}\left(\boldsymbol{x}_{\{k\}}, \boldsymbol{u}_{\{k\}}\right), \tag{5}$$

where n – last step of the finite discrete process, γ – a discount factor $(0 \leq \gamma \leq 1)$, $L_{C\{k\}}$ $(\mathbf{x}_{\{k\}}, \mathbf{u}_{\{k\}})$ – a local cost in step k.

The generated velocity error $e_{Gv\{k\}}$ and the angle of the WMR's frame turn error $e_{G\dot{\beta}\{k\}}$ for the GS behaviour are defined in the form

$$\begin{aligned} e_{Gv\{k\}} &= f\left(l^{*}_{G\{k\}}\right) - v_{A\{k\}} / v^{*}_{A}, \\ e_{G\dot{\beta}\{k\}} &= \varphi_{G\{k\}} - \beta_{\{k\}}, \end{aligned} \tag{6}$$

where $f(\cdot)$ – a sigmoidal unipolar function, $l^{*}_{G\{k\}} \in <0,1>$ – the normalised distance to the goal G, $l^{*}_{G\{k\}} = l_{G\{k\}} / l_{Gmx}, l_{G\{k\}} = \|A,G\|$, l_{Gmx} – the maximal distance to the goal G, $v_{A\{k\}}$ – a realized velocity of the point A of the WMR, v^{*}_{A} – a maximal defined velocity of the point A, $\varphi_{G\{k\}}$ – an angle between the axis of the WMR's frame and the straight line p_{G}, $\beta_{\{k\}}$ – a temporary angle of the self-turn of the WMR's frame.

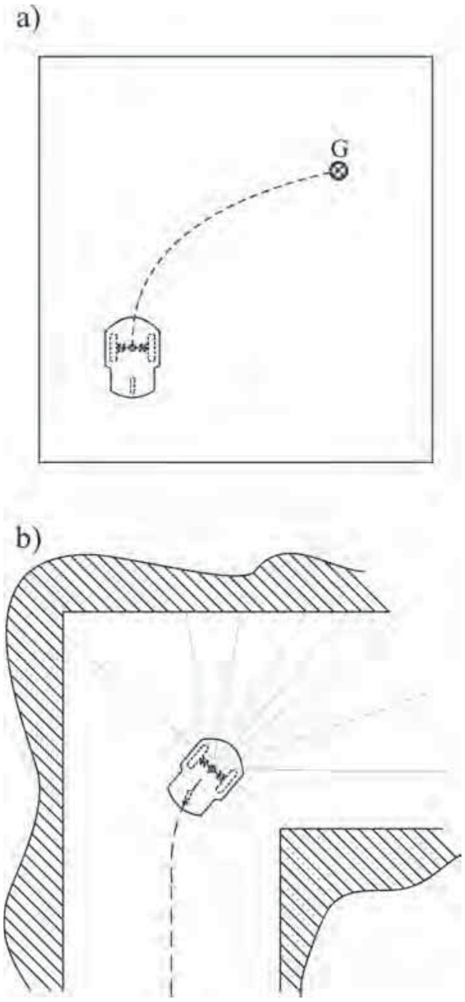

Fig. 4. a) Scheme of the goal-seeking behaviour, b) scheme of the obstacle avoiding behaviour
Rys. 4. a) Schemat realizacji zadania „podążaj do celu", b) schemat realizacji zadania „omijaj przeszkody"

The local costs $L_{CGv\{k\}}$ and $L_{CG\dot{\beta}\{k\}}$ were assumed in the forms

$$\begin{aligned} L_{CGv\{k\}} &= \frac{1}{2} R_{Gv} e^{2}_{Gv\{k\}} + \frac{1}{2} Q_{Gv} e^{2}_{Gv\{k\}}, \\ L_{CG\dot{\beta}\{k\}} &= \frac{1}{2} R_{G\dot{\beta}} e^{2}_{Gv\{k\}} + \frac{1}{2} Q_{G\dot{\beta}} e^{2}_{G\dot{\beta}\{k\}}, \end{aligned} \tag{7}$$

where R_{Gv}, $R_{G\dot{\beta}}$, Q_{Gv}, $Q_{G\dot{\beta}}$ – positive constants, $u_{Gv\{k\}}, u_{G\dot{\beta}\{k\}}$ – the overall behavioural control signals, that consist of control signals generated by actor NNs $\mathbf{u}_{GA\{k\}} = \left[u_{GAv\{k\}}, u_{GA\dot{\beta}\{k\}}\right]^{\mathrm{T}}$, and proportional (P) controller signal

$$\mathbf{u}_{G\{k\}} = \mathbf{u}_{GA\{k\}} + \mathbf{u}_{GP\{k\}} \tag{8}$$

where $\boldsymbol{u}_{GP\{k\}} = \boldsymbol{K}_{GP\{k\}}\left[e_{Gv\{k\}}, e_{G\dot{\beta}\{k\}}\right]^{\mathrm{T}}$, $\mathbf{K}_{GP\ k}$ – a positive defined, fixed diagonal matrix.

NDP structures are classified as Reinforcement Learning (RL) methods, where algorithms search for the optimal control law by exploring acceptable control laws and states of the system, and exploiting obtained strategies. Use of the proportional controller in the presented behavioural control system is an innovative approach that limits exploration by prompting the NDP structure proper control signal at the beginning of the NNs' weights adaptation process, what prevents from the trial and error learning.

The behavioural control signals $\mathrm{u}_{GA\{k\}}$ in the GS task are generated by two ADHDP actor-critic structures, composed of;

1) critic, that estimates the suboptimal value function $V_{Gv\{k\}}\left(e_{Gv\{k\}}, u_{Gv\{k\}}\right)$ or $V_{G\dot{\beta}\{k\}}\left(e_{G\dot{\beta}\{k\}}, u_{G\dot{\beta}\{k\}}\right)$, and is realised in the form of Random Vector Functional Link (RVFL) NN with output signal

$$\begin{aligned} \hat{V}_{Gv\{k\}} &= \mathbf{W}^{\mathrm{T}}_{GCv\{k\}} \mathbf{S}\left(\mathbf{x}_{GCv\{k\}}\right), \\ \hat{V}_{G\dot{\beta}\{k\}} &= \mathbf{W}^{\mathrm{T}}_{GC\dot{\beta}\{k\}} \mathbf{S}\left(\mathbf{x}_{GC\dot{\beta}\{k\}}\right), \end{aligned} \tag{9}$$

where $\mathbf{W}_{GCv\{k\}}$, $\mathbf{W}_{GC\dot{\beta}\{k\}}$ – vectors of output-layer weights, $\mathbf{S}(.)$ – the vector of sigmoidal bipolar neurons activation functions, $\mathbf{x}_{GCv\{k\}}$, $\mathbf{x}_{GC\dot{\beta}\{k\}}$ – NNs' input vectors, that contain adequate errors and control signals. Critics' weights are adapted by the back propagation method of the Temporal Difference errors in the form

$$\begin{aligned} e_{GCv\{k\}} &= L_{CGv\{k\}} + \gamma \hat{V}_{Gv\{k+1\}} - \hat{V}_{Gv\{k\}}, \\ e_{GC\dot{\beta}\{k\}} &= L_{CG\dot{\beta}\{k\}} + \gamma \hat{V}_{G\dot{\beta}\{k+1\}} - \hat{V}_{G\dot{\beta}\{k\}}, \end{aligned} \tag{10}$$

2) actor, that generates the suboptimal control law $u_{Gv\{k\}}$ or $u_{G\dot{\beta}\{k\}}$, is realised in the form of RVFL NN with output signal

$$\begin{aligned} u_{GAv\{k\}} &= \mathbf{W}^{\mathrm{T}}_{GAv\{k\}} \mathbf{S}\left(\mathbf{x}_{GAv\{k\}}\right), \\ u_{GA\dot{\beta}\{k\}} &= \mathbf{W}^{\mathrm{T}}_{GA\dot{\beta}\{k\}} \mathbf{S}\left(\mathbf{x}_{GA\dot{\beta}\{k\}}\right), \end{aligned} \tag{11}$$

and its weights are adapted by the back propagation method of errors

$$
\begin{aligned}
e_{GAv\{k\}} &= \frac{\partial L_{CGv\{k\}}}{\partial u_{Gv\{k\}}} + \gamma \frac{\partial \hat{V}_{Gv\{k+1\}}}{\partial u_{Gv\{k\}}}, \\
e_{GA\dot{\beta}\{k\}} &= \frac{\partial L_{CG\dot{\beta}\{k\}}}{\partial u_{G\dot{\beta}\{k\}}} + \gamma \frac{\partial \hat{V}_{G\dot{\beta}\{k+1\}}}{\partial u_{G\dot{\beta}\{k\}}}.
\end{aligned} \tag{12}
$$

In the behavioural control systems were used RVFL NNs with fixed input-layer weights, randomly chosen in the initialization process, set to zero initial output-layer weights and neurons with sigmoidal bipolar activation functions. Each NN had eight neuron activation functions.

Behavioural control system in the obstacle avoiding behaviour. The control system in the OA task is build in the same way that in the GS task. The actor-critic structures minimises the value functions based on errors of generated velocity $e_{Ov\{k\}}$ and the middle of the free space $e_{O\dot{\beta}\{k\}}$. The behavioural control signals $\boldsymbol{u}_{OA\{k\}}$ in the OA task are generated by two ADHDP actor-critic structures and the proportional controller.

Fuzzy combiner of behaviours. We used the Takagi-Sugeno FL model, with triangular or trapezoidal affiliation functions to fuzzy sets. The FL controller contains the rules base that consists of $m = 25$ rules in a form:

$$R_B^m : \mathrm{IF}\ \left(l^*_{G\{k\}}\ \mathrm{is\, lS}\right)\ \mathrm{AND}\ \left(d^*_{O\{k\}}\ \mathrm{is\, dS}\right)\ \ \mathrm{THEN}\ a_{B\{k\}}\ \mathrm{is\ aM} \tag{13}$$

where $d^*_{O\{k\}} \in <0,\ 1>$ – the normalised distance to the obstacle, $d^*_{O\{k\}} = \min\Big(d_{L[1]\{k\}}(s_2),\ d_{L[2]\{k\}}(s_3),\ d_{F[1]\{k\}}(s_L),\ d_{R[1]\{k\}}(s_6),\ d_{R[2]\{k\}}(s_7)\Big)/l_{Omx}$, l_{Omx} – the maximal range of sensors, $a_{B\{k\}}$ – the combination of individual behaviours control signal, “lS”, “dS”, “aM”– linguistic labels of affiliation functions to the fuzzy sets.

Scheme of the rules base is shown in fig. 5, where linguistic labels of particular affiliation functions to the fuzzy sets are: “WS0”– very small, near zero, “WS”– very small, “S”– small, “M”– medium, “B”– big, “WB”– very big, “WB1”– very big, near one.

The trajectory generation system generates the control signal $\mathbf{u}_{B\{k\}} = \ u_{Bv\ k}, u_{B\dot{\ }\ k}\ {}^{\mathrm{T}}$, on the basis of control signals generated for the individual behaviours; the goal-seeking $\mathbf{u}_{G\{k\}} = \left[u_{Gv\{k\}}, u_{G\dot{\beta}\{k\}}\right]^{\mathrm{T}}$ and the obstacle avoiding $\mathbf{u}_{O\{k\}} = \left[u_{Ov\{k\}}, u_{O\dot{\beta}\{k\}}\right]^{\mathrm{T}}$, according to equation

$$\mathbf{u}_{B\{k\}} = a_{B\{k\}}\mathbf{u}_{G\{k\}} + \left(1 - a_{B\{k\}}\right)\mathbf{u}_{O\{k\}}\,. \tag{14}$$

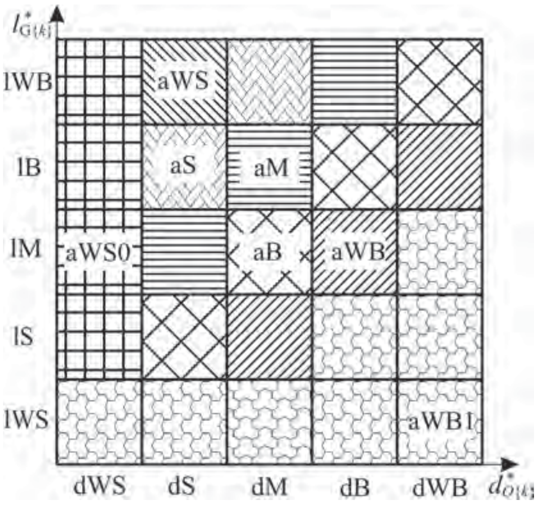

Fig. 5. Scheme of the rules base with fuzzy logic
Rys. 5. Schemat bazy reguł układu z logiką rozmytą

In the global co-ordinate system xy position of the WMR is described by $\left[x_{A\{k\}}, y_{A\{k\}}, \beta_{\{k\}}\right]^{\mathrm{T}}$, where $\left(x_{A\{k\}}, y_{A\{k\}}\right)$ are co-ordinates of the point A. The angular velocities for proper wheels are calculated according to equation

$$\begin{bmatrix} z_{d2[1]\{k\}} \\ z_{d2[2]\{k\}} \end{bmatrix} = \frac{1}{r} \begin{bmatrix} v_A^* & \dot{\beta}^* l_1 \\ v_A^* & -\dot{\beta}^* l_1 \end{bmatrix} \begin{bmatrix} u_{Bv\{k\}} \\ u_{B\dot{\beta}\{k\}} \end{bmatrix}, \tag{15}$$

where $\dot{\beta}^*$ – a maximal defined angular velocity of the self turn of the WMR frame, l_1, $r = r_{[1]} = r_{[2]}$ – the lengths that derive from the WMR geometry.

4. Experiment results

Verification of the proposed control algorithm was realised by a series of experiments using the WMR Pioneer 2-DX in the laboratory environment. The experimental system consists of the WMR Pionner 2-DX and PC with the dSpace DS1102 digital signal processing board, MATLAB and dSpace Control Desk software. In this section, for the sake of simplicity, all variables are presented in a continuous domain of the time and there is not used k index, $h = 0.01$ s.

On the basis of range finder signals the proposed control system generated the collision free path of the point A of the WMR’s frame from the starting point to the goal. The environment maps with trajectories of the point A, positions of obstacles localised by range finders and the destination in points G(0.8, 4.2) and G(3.4, 5.0), are shown in fig. 6a) and b). In the figure the start position of the WMRis marked by the triangle, the goal is marked by the ”X”.

The map of the environment was projected in the way, that none of the behavioural control systems in the OA or the GS task are able to generate the successive path, it is possible on the basis of the control signal generated by the presented algorithm with the fuzzy coordinator of the behaviours. Obstacles detected by the sensors system are pointed by black dots in fig. 6. The quality of measurements depends on the type of used range finders, and has a significant influence on the trajectory generating process. The best mapping of the obstacles localization was realized using the laser range finder, but in measurements taken by sonars errors occurred. The localisations of obstacles were computed on the basis of sensors readings, known geometry of the sensors system, localisation of the point A and orientation of the WMR’s frame, measured using incremental encoders. Errors in sensors readings and measurements of the realised angles of the self-turn of wheels influence on computed localisations of detected obstacles in coordinates of the map, what cause a difference between actual and computed localisations of obstacles in fig. 6.a) and b).

On the basis of the WMR's sensor system signals was generated the FL combination of behaviours control signal a_{B} shown in fig. 7.a), for the goal G(0.8, 4.2), and the overall trajectory generator control signals u_{Bv} and $u_{B\dot{\beta}}$, shown in fig. 7.b). The control signals u_{Bv} and $u_{B\dot{\beta}}$ are a fuzzy combination of behavioural control signals u_{Ov} and $u_{O\dot{\beta}}$ for the OA behaviour, presented in fig. 7c), and the control signals u_{Gv} and $u_{G\dot{\beta}}$ for the GS behaviour, shown in fig. 7d).

The control signals in the GS behaviour are smooth, because are computed according to eq. (6) on the basis of localisation of points A and G, and orientation of the WMR's frame. The control signals in the OA behaviour and the α_{B} signal depend on disturbed sensors readings, therefore are not smooth.

Values of the actor's ($\mathbf{W}_{GA1}$) and the critic's ($\mathbf{W}_{GC1}$) NN weights of the ADHDP structure, that generates the behavioural control signal u_{GAv} in the GS behaviour, are shown in fig. 7a) and b). Weights of NNs are bounded and converge to the fixed values.

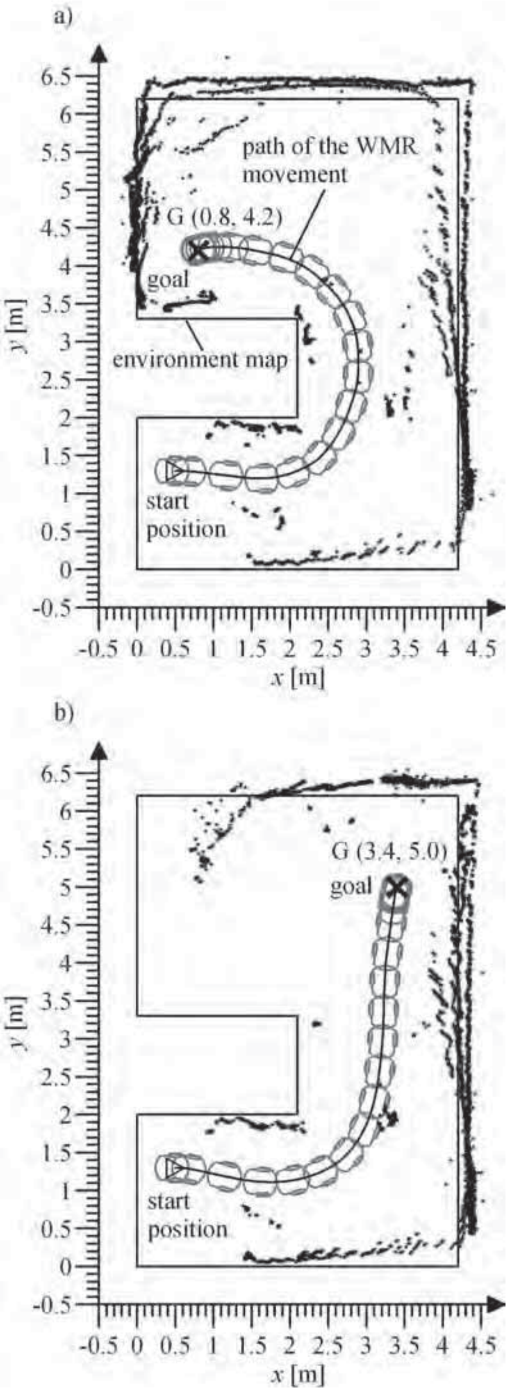

Fig. 6. The environment map with the path of the point A of the Pioneer 2-DX to the goal: a) G(0.8, 4.2), b) G(3.4, 5.0)

Rys. 6. Mapa otoczenia z torem ruchu punktu A mobilnego robota Pioneer 2-DX do celu: a) G(0,8, 4,2), b) G(3,4, 5,0)

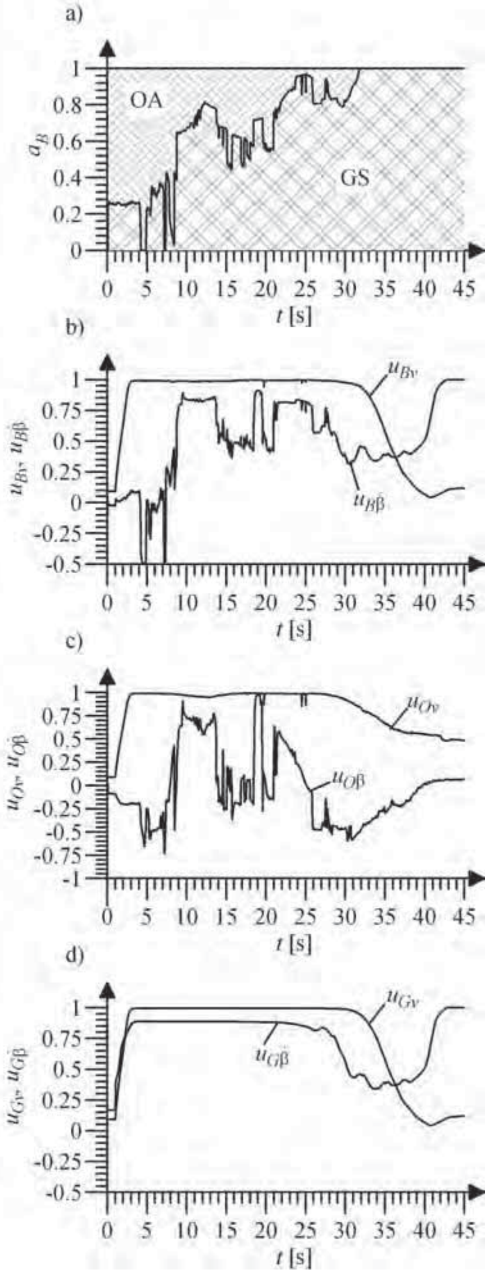

Fig. 7. a) Signal a_B, b) overall control signals u_{Bv} and $u_{B\dot{\beta}}$, c) control signals u_{Ov} and $u_{O\dot{\beta}}$ for the obstacle avoiding behaviour, d) control signals u_{Gv} and $u_{G\dot{\beta}}$ for the goal-seeking behaviour

Rys. 7. a) Sygnał a_B, b) całkowite sygnały sterowania u_{Bv} i $u_{B\dot{\beta}}$, c) sygnały sterowania u_{Ov} i $u_{O\dot{\beta}}$ w zadaniu „omijaj przeszkody”, d) sygnały sterowania u_{Gv} i $u_{G\dot{\beta}}$ w zadaniu „podążaj do celu”

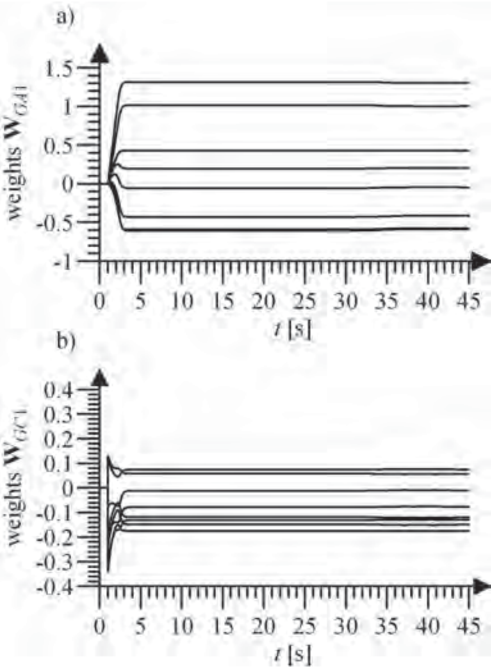

Fig. 8. a) Weights of the ADHDP actor 1 NN, b) weights of the ADHDP critic 1 NN

Rys. 8. a) Wagi sieci neuronowej aktora $\mathbf{W}_{GA1}$ struktury ADHDP, b) wagi sieci neuronowej krytyka $\mathbf{W}_{GC1}$ struktury ADHDP

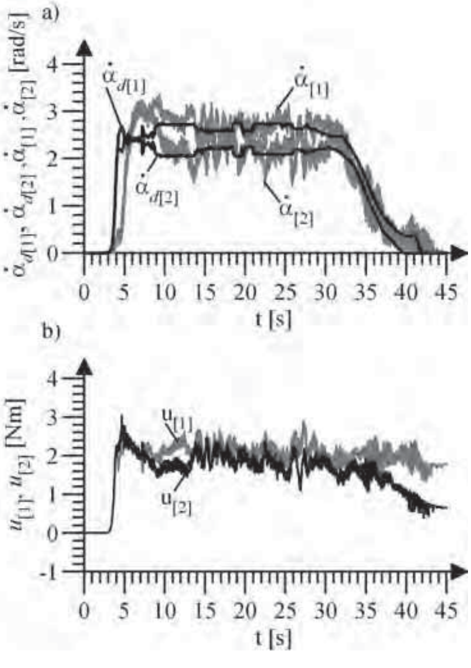

Fig. 9. a) Desired ($\dot\alpha_{d[1]}$,$\dot\alpha_{d[2]}$) and realized ($\dot\alpha_{[1]}$,$\dot\alpha_{[2]}$) angular velocities, b) the overall tracking control signals u_1 and u_2

Rys. 9. a) Zadane ($\dot\alpha_{d[1]}$ $\dot\alpha_{d[2]}$) i zrealizowane ($\dot\alpha_{[1]}$ $\dot\alpha_{[2]}$) prędkości kątowe obrotu kół robota mobilnego, b) całkowite sygnały sterowania ruchem nadążnym u_1 u_2

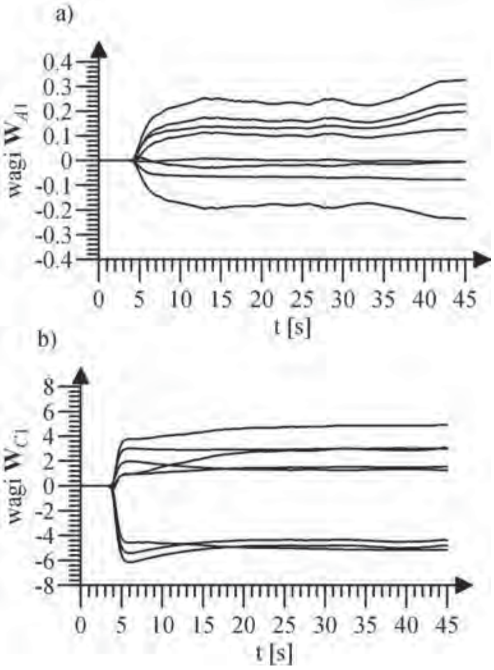

Fig. 10. a) Weights of the DHP actor 1 NN, b) weights of the DHP critic 1 NN

Rys. 10. a) Wagi sieci neuronowej aktora $\mathbf{W}_{A1}$ struktury DHP, b) wagi sieci neuronowej krytyka $\mathbf{W}_{C1}$ struktury DHP

On the basis of the overall trajectory generator control signals u_{Bv} and $u_{B\dot\beta}$, according to the eq. (15), were computed desired angular velocities of the WMR ($\dot\alpha_{d[1]}$, $\dot\alpha_{d[2]}$ that denote to $z_{d2[1]}$ and $z_{d2[2]}$), realised using the tracking control system with the overall tracking control signals $u_{[1]}$, $u_{[2]}$, shown in fig. 7b). The desired and realised ($\dot\alpha_{[1]}$, $\dot\alpha_{[2]}$) angular velocities of the WMR are shown in fig. 7a).

Values of the actor's ($\mathbf{W}_{A1}$) and the critic's ($\mathbf{W}_{C1}$) NN weights of the NDP structure in DHP configuration, that generates the tracking control signal $u_{[1]}$, are shown in fig. 10 a) and b). Weights of NNs are bounded and converge to the fixed values.

5. Summary

The proposed hierarchical control system, with NDP structures in ADHDP configuration in the trajectory generator and DHP algorithms in the tracking control system, generates and realises the collision free trajectory of the WMR Pioneer 2-DX in the unknown 2D environment with static obstacles. The trajectory generator consists of the FL controller and two behavioural control systems for the OA and the GS behaviour. The FL system generates control signal used to soft switching of the behavioural control signals. Each of the behavioural control algorithms consist of ACDs and the proportional regulator, what is an innovative approach that prevents from the time consuming trial and error learning. The generated trajectory provides, that the point A of the WMR Pioneer 2-DX reaches the goal. Significant influence on the trajectory generating process have a quality of measurements and a type of used range finders. The projected hierarchical control system with sensor-based navigator works on-line and does not require the preliminary learning of NNs.

Acknowledgements

This research was realized within a framework of research project No. U-8314/DS/M.

Apparatus/equipment purchased in the project No. POPW.01.03.00-18-012/09 from the structural funds, the Development of Eastern Poland Operational Programme co-financed by the European Union, the European Regional Development Fund.

References

1. Arkin R., *Behavior-Based Robotics*, MIT Press, Cambridge 1998.
2. Barto A., Sutton R., *Reinforcement learning: an introduction*, MIT Press, Cambridge 1998.
3. Burghardt A., *Proposal for a rapid prototyping environment for algorithms intended for autonomous mobile robot control*, "Mechanics and Mechanical Engineering", 2008, Vol. 12, 5–16.
4. Fahimi F., *Autonomous Robots. Modeling, Path Planning, and Control*, Springer, New York 2009.
5. Giergiel J., Hendzel Z., Zylski W., *Modeling and Control of Wheeled Mobile Robots* (in Polish), WNT, Warsaw 2002.
6. Giergiel J., Zylski W., *Description of Motion of a Mobile Robot by Maggie's Equations*, "J. Theor. App. Mech.", 2005, Vol. 43, 511–521.
7. Hendzel Z., *Fuzzy reactive control of wheeled mobile robot*, "J. Theor. App. Mech.", 2004, Vol. 42, 503–517.
8. Hendzel Z., Szuster M., *Discrete Action Dependant Heuristic Dynamic Programming in Wheeled Mobile Robot Control*, "Solid State Phenomena", 2010, Vol. 164, 419–424.
9. Hendzel Z., Szuster M., *Discrete Model-Based Adaptive Critic Designs in Wheeled Mobile Robot Control*, LNAI, 2010, Vol. 6114, 264–271.
10. Hendzel Z., Szuster M., *Neural Dynamic Programming in Behavioural Control of Wheeled Mobile Robot*, (in Polish), "Acta Mechanica et Automatica", 2011, Vol. 5, No. 1, 28–36.
11. Hendzel Z., Szuster M., *Neural Dynamic Programming in Reactive Navigation of Wheeled Mobile Robot*, Rutkowski L. et al. (eds.): ICAISC 2012, Part II, LNCS, Vol. 7268, 450–457.
12. Maaref H., Barret C., *Sensor-based Navigation of a Mobile Robot in an Indoor Environment*, "Robotics and Autonomous Systems", 2002, Vol. 38, 1–18.
13. Millan J., *Reinforcement Learning of Goal-Directed Obstacle-Avoiding Reaction Strategies in an Autonomous Mobile Robot*, "Robotics and Autonomous Systems", 1995, Vol. 15, 275–299.
14. Powell W., *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, Willey-Interscience, Princeton 2007.
15. Si J., Barto A., Powell W., Wunsch D., *Handbook of Learning and Approximate Dynamic Programming*, IEEE Press, Wiley-Interscience 2004.
16. Prokhorov D., Wunch D., *Adaptive Critic Designs*. "IEEE Transactions on Neural Networks", 1997, Vol. 8, 997–1007.

■

Zastosowanie aproksymacyjnego programowania dynamicznego w nawigacji mobilnego robota kołowego

Streszczenie: W prezentowanym artykule zaproponowano nowe podejście do generowania bezkolizyjnych trajektorii ruchu mobilnych robotów kołowych z zastosowaniem algorytmów adaptacyjnego krytyka oraz układów z logiką rozmytą. Zaprezentowany hierarchiczny układ sterowania składa się z warstwy generowania trajektorii ruchu bazującej na idei odruchowej nawigacji mobilnego robota kołowego w nieznanym środowisku 2D ze statycznymi przeszkodami oraz warstwy sterowania ruchem nadążnym. Sterowanie odruchowe obejmuje dwa podstawowe zadania: omijanie przeszkód oraz podążanie do celu, zrealizowane z zastosowaniem algorytmów adaptacyjnego krytyka. Te proste zachowania są łączone przez układ z logiką rozmytą, który określa wpływ poszczególnych zachowań na proces generowania trajektorii w zależności od warunków otoczenia sterowanego obiektu. Weryfikacja zaproponowanego algorytmu sterowania została zrealizowana z zastosowaniem mobilnego robota kołowego Pioneer 2-DX, wyposażonego w dalmierz laserowy i osiem sonarów ultradźwiękowych, służących do wykrywania przeszkód.

Słowa kluczowe: aproksymacyjne programowanie dynamiczne, sterowanie behawioralne, mobilny robot, nawigacja

---

Prof. Zenon Hendzel, DSc, PhD, Eng.

Graduated from the AGH University of Science and Technology in Krakow (1973), the PhD degree obtained in 1980 at the Faculty of Electronics, University of Technology in Warsaw. Doctoral degree, as well as the title of Professor received at the Faculty of Mechanical Engineering and Robotics, AGH University of Science and Technology in Krakow (1997), (2007). Since 1973 is active an employee of the University of Technology in Rzeszow. The main areas of his research are: nonlinear control systems, robust control, adaptive control, methods of artificial intelligence, and control issues of the mobile wheeled robots.

e-mail: zenhen@prz.edu.pl

Marcin Szuster, PhD

He received the MSc degree in mechatronics and the PhD degree in mechanics from the Rzeszow University of Technology in 2007 and 2012, respectively. He works at the Department of Applied Mechanic and Robotics of the Rzeszow University of Technology. His research interests are in adaptive dynamic programming, behavioural control and path planning, control of robots and various methods of artificial intelligence, especially neural networks.

e-mail: mszuster@prze.edu.pl