Artificial Intelligence Methods in Reactive Navigation of Mobile Robots Formation

Zenon Hendzel, Andrzej Burghardt and Marcin Szuster
Dept of Applied Mechanics and Robotics, Rzeszow University of Technology, 8 Powstancow Warszawy St., Rzeszow, Poland

Keywords: Behavioural Control, Adaptive Critic Design, Robots Formation, Reactive Navigation, Neural Networks.

Abstract: The article presents a hierarchical control system build using artificial intelligence methods, that generates a trajectory of the wheeled mobile robots formation, and realises the tracking control task of all agents. The hierarchical control system consists of a navigator, based on a conception of behavioural control signals coordination, and individual tracking control systems for all mobile robots in the formation. The navigator realises a sensor-based approach to the path planning process in the unknown 2-D environment with static obstacles. The navigator presents a new approach to the behavioural control, where one Neural dynamic programming algorithm generates the control signal for the complex behaviour, which is a composition of two individual behaviours: “goal-seeking”and “obstacle avoiding”. Influence of individual behaviours on the navigator control signal depends on the environment conditions and changes fluently. On the basis of control signal generated by the navigator are computed the desired collision-free trajectories for all robots in formation, realised by the tracking control systems. Realisation of generated trajectories guarantees reaching the goal by selected point of the robots formation with obstacles avoiding by all agents. Computer simulations have been conducted to illustrate the process of path planning in the different environment conditions.

1 INTRODUCTION

The problem of large-size objects transport is difficult to solve and expensive in realisation. It requires to use suitably large transport facilities or a group of small cooperating devices. The second conception seems to be more adequate but is harder to apply. The transporters cooperating in a formation in the large-size load transportation task can be also useful after fulfilling the task, but the cooperation of human operators is not always suitable and can lead to dangerous situations. This problem can be solved by using autonomous group of mobile robots, moving in a definite formation with precisely determined position of individual agents in formation.

The tracking control task of the wheeled mobile robot (WMR) is difficult because of its dynamics described by the non-linear equations, and changeable parameters during transportation tasks. The problem of not known or changing parameters of the WMR dynamics model in the tracking control task, is often solved by using adaptive methods in the tracking control system, like modern Artificial inteligence (AI) algorithms, especially neural networks (NNs). The second problem is to coordinate the movement of all agents in the wheeled mobile robots formation (WMRF) to successively complete the task. This type of problem can be solved by using virtual structure algorithms (Egerstedt and Hu, 2001; Ogren and Leonard, 2003). The third problem concerns the conception of sensor-based navigation in generating the trajectory of the WMRF in the unknown environment with static obstacles (Arkin, 1998; Fahimi, 2008; Maaref and Barret, 2002; Millan, 1995). This task is often solved by deriving inspiration from the wold of animals in a form of behavioural methods of WMRF control (Yamaguchi, 1997).

The development of AI methods, like NNs, allowed to apply Bellman’s Dynamic Programming (DP) idea in the form of Neural Dynamic Programming (NDP) algorithms (Sutton and Barto, 1998; Powell, 2007; Prokhorov and Wunsch, 1997; Si et al., 2004), that proved to be very efficient in the control tasks. In the article, the hierarchical control system with NDP algorithms is presented. It consists of three main layers: the highest is the navigator, that generates the desired trajectory of the WMRF, the middle layer is the robots formation control system, that generates desired trajectories for all agents, and the lowest layer consists of the tracking control systems for

Hendzel Z., Szuster M. and Burghardt A..
Artificial Intelligence Methods in Reactive Navigation of Mobile Robots Formation.
DOI: 10.5220/0004113404660473
In *Proceedings of the 4th International Joint Conference on Computational Intelligence* (NCTA-2012), pages 466-473
ISBN: 978-989-8565-33-4

individual agents.

The results of scientific researches presented in the article continuous authors' earlier works related to the tracking control of the WMR problem (Hendzel and Szuster, 2010a; Hendzel and Szuster, 2010b), the WMRF control (A. Burghardt and Giergiel, 2011) and the trajectory generating process in the unknown 2D environment (Burghardt, 2004; Burghardt, 2008; Hendzel, 2004; Hendzel and Szuster, 2011; Hendzel and Szuster, 2012), where were used AI methods. The article is organised in the following way: the first section is an introduction to the WMRs control problem, connected with the tracking control, the WMRF control and path planning. The second section contains description of the WMR dynamics. The section three presents hierarchical control system. Section four contains results of numerical tests. The last section summarises the article.

2 MODEL OF THE WHEELED MOBILE ROBOT

The WMRF is composed of $m = 3$ WMRs Amigobot, that form a virtual structure of a triangle. The WMR Amigobot consists of two driving wheels, a frame and a third castor wheel. It has eight ultrasonic range finders $s_1, \ldots, s_8$ for obstacles detection. Angles between axes of ultrasonic range finders and the axis of the frame of Amigobot are equal $\omega_1 = 144°$, $\omega_2 = 90°$, $\omega_3 = 44°$, $\omega_4 = 12°$, $\omega_5 = -12°$, $\omega_6 = -44°$, $\omega_7 = -90°$, $\omega_8 = -144°$, the range of individual range finder measurements is equal to d_i, $i = 1,\ldots,8$, and the maximal range $d_{mx} = 4$ [m].

The discrete notation of the WMR dynamics was derived using the Euler's derivative approximation of the model obtained by applying the Maggie's mathematical formalism (Giergiel J., 2002; Giergiel and Zylski, 2005), and takes the form

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

where $\mathbf{z}_{2\{k\}} = \left[z_{2[1]\{k\}}, z_{2[2]\{k\}}\right]^{\mathrm{T}}$ – the vector of discrete angular velocities, $\mathbf{M}$ – the positive definite inertia matrix, $\mathbf{C}\left(\mathbf{z}_{2\{k\}}\right)\mathbf{z}_{2\{k\}}$ – the vector of Coriolis and centrifugal forces, $\mathbf{F}\left(\mathbf{z}_{2\{k\}}\right)$ – the vector of rolling resistances, τ_d – the vector of bounded disturbances, $\mathbf{u} = \left[u_{[1]}, u_{[2]}\right]^{\mathrm{T}}$ – the vector of tracking control signals, h – the time discretization parameter, k – the index of iteration steps.

The movement of the WMR is analysed in xy plane. The scheme of the WMR in the environment with static obstacles is shown in fig. 1.

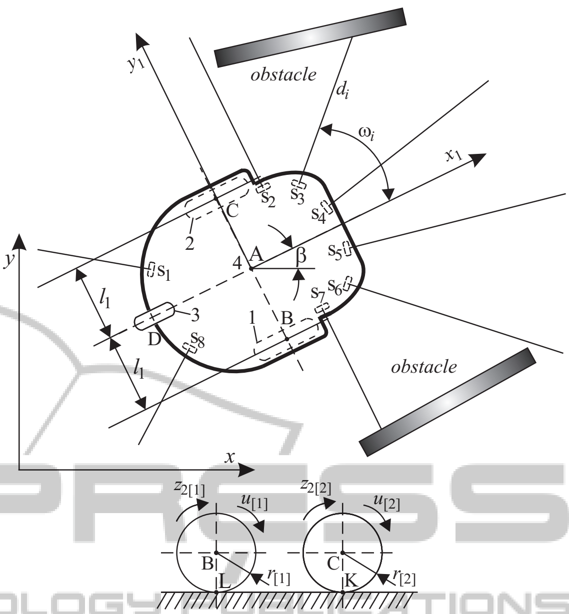

Figure 1: The wheeled mobile robot Amigobot scheme.

3 HIERARCHICAL CONTROL SYSTEM

The proposed hierarchical control system is composed of three main layers. The highest layer is the navigator, that generates the desired trajectory of the point M of the virtual structure. The middle layer is the robots formation control system, that generates desired trajectories for all agents. The lowest layer consists of m tracking control systems for individual agents. Scheme of the proposed hierarchical control system is shown in fig. 2.

3.1 Navigation in the Unknown Environment

The navigator consists of the discrete Action dependant heuristic dynamic programming (ADHDP) structures and the proportional (P) controller. The ADHDP structures are adapted on-line using a reinforcement learning (RL) idea, that bases on the iterative interaction with the environment. The NDP algorithm searches for the optimal action to take. Performing this action minimises the assumed cost function. The presented construction of the navigator is an innovative approach to the trajectory generating process, it uses the P regulator in the navigator to indicate

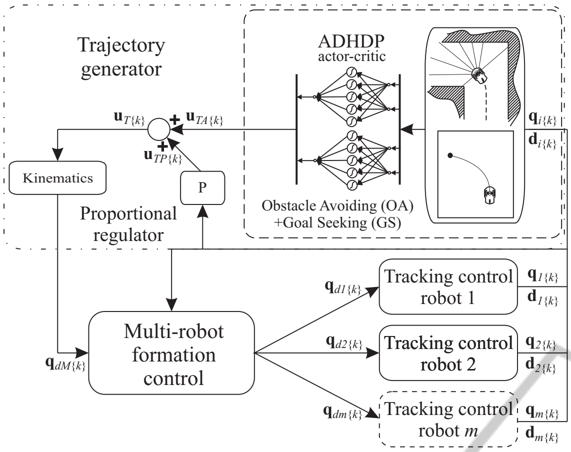

Figure 2: Scheme of the control system for robots formation.

NDP structures adequate control signal at the beginning of the adaptation process to limit exploration and avoid the trial and error learning.

The navigator presents a new approach to the behavioural control, where one Neural dynamic programming algorithm generates the control signal for the complex behaviour, which is a composition of two individual behaviours: "goal-seeking"(GS) and "obstacle avoiding"(OA).

The overall navigator's control signal $\mathbf{u}_{T\{k\}} = \left[u_{Tv\{k\}}, u_{T\dot{\beta}\{k\}}\right]^{\mathrm{T}}$ consists of the control signal $u_{Tv\{k\}}$ that controls the desired velocity of the point M of the virtual structure, and the control signal $u_{T\dot{\beta}\{k\}}$, that corresponds to the angular velocity of the WMRF's self-turn $\dot{\beta}_M$. The overall control signal $\mathbf{u}_{T\{k\}}$ is a sum of control signal generated by the actor-critic ADHDP structures $\mathbf{u}_{TA\{k\}}$ and the P regulator control signals $\mathbf{u}_{BP\{k\}}$, according to equation

$$\mathbf{u}_{T\{k\}} = \mathbf{u}_{TA\{k\}} + \mathbf{u}_{TP\{k\}}, \tag{2}$$

where $\mathbf{u}_{TP\{k\}} = \mathbf{K}_T\, \mathbf{e}_{T\{k\}}$, $\mathbf{K}_T$ - a fixed matrix of proportional gains,

$$\mathbf{K}_T = \left[\begin{array}{ccc} k_{Tv} & 0 & 0 \\ 0 & k_{TO} & k_{TG} \end{array}\right]. \tag{3}$$

and k_{Tv}, k_{TO}, k_{TG} are fixed positive gains. The errors of the trajectory generating layer are defined in the form

$$\begin{array}{l} e_{v\{k\}} = f\left(d^*_{F\{k\}}\right) f\left(l_{G\{k\}}\right) - v_{A\{k\}}/v^*_A, \\ e_{O\{k\}} = d^*_{R\{k\}} - d^*_{L\{k\}}, \\ e_{G\{k\}} = \varphi_{G\{k\}} - \beta_{M\{k\}}, \end{array} \tag{4}$$

where $f(.)$ – a sigmoidal bipolar function, $d^*_{F\{k\}} = \min\left(d_{i[4]\{k\}}\left(\mathrm{s}_{i4}\right), d_{i[5]\{k\}}\left(\mathrm{s}_{i5}\right)\right)/d_{mx}$ – the normalised distance to the obstacle in the front of the WMRF, $d_{i[j]\{k\}}$ – range of the j - th range finder of the i-th WMR in the formation, $l_{G\{k\}}$ – the distance of WMRF centre (point M) to the point G, $v_{M\{k\}}$ – a realised velocity of the point M of the virtual structure, v^*_M – a maximal defined velocity of the point M, $d_{L\{k\}} = \min\left(d_{1[3]\{k\}}\left(\mathrm{s}_{13}\right), d_{2[3]\{k\}}\left(\mathrm{s}_{23}\right)\right)$, $d_{R\{k\}} = \min\left(d_{1[6]\{k\}}\left(\mathrm{s}_{16}\right), d_{3[6]\{k\}}\left(\mathrm{s}_{36}\right)\right)$, $d^*_{L\{k\}} = 2\left[\left(d_{L\{k\}}/\left(d_{L\{k\}}+d_{R\{k\}}\right)\right) - 0.5\right]$ – the normalised distance to the obstacle on the left side of the WMRF, $d^*_{R\{k\}} = 2\left[\left(d_{R\{k\}}/\left(d_{L\{k\}}+d_{R\{k\}}\right)\right) - 0.5\right]$ – the normalised distance to the obstacle on the right, $e_{G\{k\}}$ – the temporal angle between the x axis and the line p_G, $\beta_{M\{k\}}$ – the temporal angle of the self-turn of the virtual structure.

The main objective of the actor-critic structures in ADHDP configuration is to generate control signals, that minimises the value functions $V_{v\{k\}}$ and $V_{\dot{\beta}\{k\}}$ in the form

$$\begin{array}{l} V_{v\{k\}} = \sum\limits_{k=0}^{n} \gamma^k L_{Cv\{k\}}, \\ V_{\dot{\beta}\{k\}} = \sum\limits_{k=0}^{n} \gamma^k L_{C\dot{\beta}\{k\}}, \end{array} \tag{5}$$

where n – the last step of the finite discrete process, γ – a discount factor $(0 \leq \gamma \leq 1)$, $L_{Cv\{k\}}$ – a local cost in step k for the first actor-critic structure, $L_{C\dot{\beta}\{k\}}$ – a local cost in step k for the second actor-critic structure of the navigator.

ADHDP algorithm used in the proposed navigator does not require a mathematical model of the controlled object or process in NNs adaptation laws.

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

$$\begin{array}{lcl} L_{Cv\{k\}} & = & \frac{1}{2}P_v e^2_{v\{k\}} + \frac{1}{2}R_v u^2_{Tv\{k\}}, \\ L_{C\dot{\beta}\{k\}} & = & \frac{1}{2}P_G e^2_{G\{k\}} + \frac{1}{2}P_O e^2_{O\{k\}} + \frac{1}{2}R_{\dot{\beta}} u^2_{T\dot{\beta}\{k\}}, \end{array} \tag{6}$$

where P_v, P_G, P_O, R_v, $R_{\dot{\beta}}$ – fixed, positive defined scaling rates.

The navigator control signals were generated by two actor-critic structures in ADHDP configuration. ADHDP structure consists of:

- critic RVFL (Random Vector Functional Link) NN, that estimates the suboptimal value functions 5, and generates signals

$$\begin{array}{l} \hat{V}_{v\{k\}} = \mathbf{W}^{\mathrm{T}}_{Cv\{k\}} \mathbf{S}\left(\mathbf{x}_{Cv\{k\}}\right), \\ \hat{V}_{\dot{\beta}\{k\}} = \mathbf{W}^{\mathrm{T}}_{C\dot{\beta}\{k\}} \mathbf{S}\left(\mathbf{x}_{C\dot{\beta}\{k\}}\right), \end{array} \tag{7}$$

where $\mathbf{W}_{Cv\{k\}}$, $\mathbf{W}_{C\dot{\beta}\{k\}}$ – vectors of critic NNs' output-layer weights, $\mathbf{S}(.)$ – a vector of sigmoidal bipolar neurons activation functions, $\mathbf{x}_{Cv\{k\}}$, $\mathbf{x}_{C\dot{\beta}\{k\}}$ – input vectors to the critic NNs. The input vectors to the critic NNs consists of adequate errors and control signals.

Critics' weights are adapted by the back propagation method of the Temporal Difference error (Powell, 2007; Prokhorov and Wunsch, 1997; Si et al., 2004).

- actor RVFL NNs, that generates the control law $u_{TAv\{k\}}$ or $u_{TA\dot{\beta}\{k\}}$, according to equations

$$
\begin{aligned}
u_{TAv\{k\}} &= \mathbf{W}^{\mathrm{T}}_{Av\{k\}}\mathbf{S}\left(\mathbf{x}_{Av\{k\}}\right), \\
u_{TA\dot{\beta}\{k\}} &= \mathbf{W}^{\mathrm{T}}_{A\dot{\beta}\{k\}}\mathbf{S}\left(\mathbf{x}_{A\dot{\beta}\{k\}}\right),
\end{aligned}
\tag{8}
$$

where $\mathbf{W}_{Av\{k\}}$, $\mathbf{W}_{A\dot{\beta}\{k\}}$ – vectors of actor NNs' output-layer weights, $\mathbf{x}_{Av\{k\}}$, $\mathbf{x}_{A\dot{\beta}\{k\}}$ – input vectors to the actor NNs.

The general scheme of the NDP structure in ADHDP configuration used in the navigator is shown in fig. 3, where **e** is the error vector.

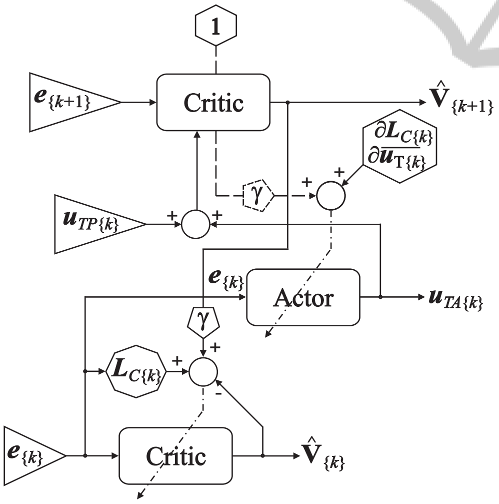

Figure 3: Scheme of the ADHDP structure.

The discrete navigator using ADHDP algorithms was described in detail in (Hendzel and Szuster, 2012).

3.2 Control of the Robots in Formation

The WMRF control system uses the idea of the virtual structure with the centre in point $\mathrm{M}(x_{M\{k\}}, y_{M\{k\}})$, and orientation defined by $\beta_{M\{k\}}$ angle, shown in fig. 4. Position and orientation of the virtual structure change according to control signals of the navigator ($u_{Tv\{k\}}$ and $u_{T\dot{\beta}\{k\}}$), that depend on the environment conditions and assumed localisation of the goal.

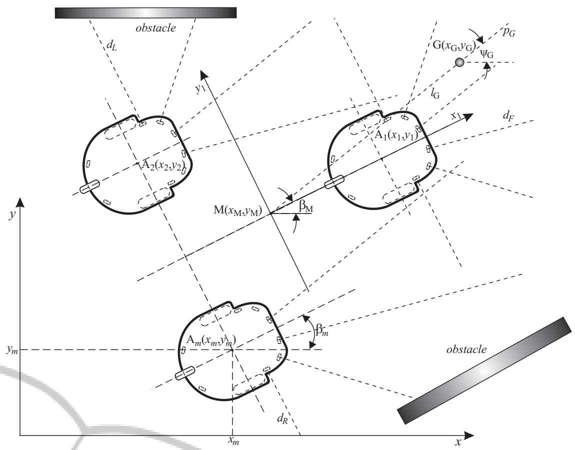

Figure 4: The robots formation scheme.

Positions of characteristic points of the virtual structure are traced by the WMRs points A in the way, that the i-th WMR's point $\mathrm{A}_{i\{k\}}(x_{i\{k\}}, y_{i\{k\}})$ is going to achieve in the next iteration step the desired position $\mathrm{A}_{di\{k\}}(x_{di\{k\}}, y_{di\{k\}})$ computed on the basis of the virtual structure position and orientation. The idea of the WMRF control bases on generating trajectories of point A of the individual WMRs. Determined trajectories guarantee minimisation of errors $\delta_{Li\{k\}}$ and $\psi_{Li\{k\}}$, what results in the trajectories, in which point A_i of the i-th WMR traces point A_{di} of the virtual structure. The idea of formation control is shown in fig. 5.

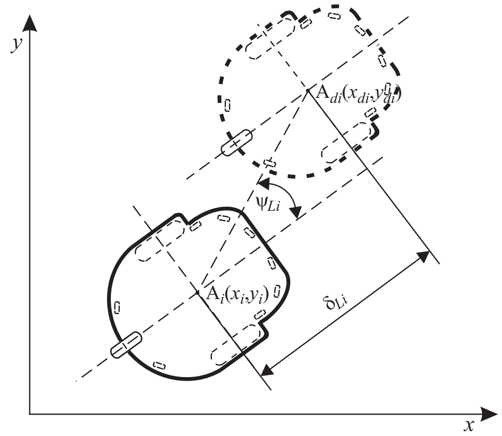

Figure 5: Conception of robots formation control with errors $\delta_{Li\{k\}}$ and $\psi_{Li\{k\}}$.

The location of the i-th WMR is defined by the coordinates $\mathbf{q}_{i\{k\}} = \left[x_{i\{k\}}, y_{i\{k\}}, \beta_{i\{k\}}\right]^T$, where $\beta_{i\{k\}}$ is an angle of the self turn of the i-th WMR's frame.

We assumed the WMRF control signals in the form

$$u_{Fvi\{k\}} = k_{F1}\delta_{Li\{k\}}\cos\left(\psi_{Li\{k\}}\right),$$
$$u_{F\dot{\beta}i\{k\}} = k_{F1}\sin\left(\psi_{Li\{k\}}\right)\cos\left(\psi_{Li\{k\}}\right) + k_{F2}\psi_{Li\{k\}}, \tag{9}$$

where k_{F1}, k_{F2} – positive constants. The presented formation control system was presented in detail in (A. Burghardt and Giergiel, 2011).

On the basis of the WMRF control signals $u_{Fvi\{k\}}$ and $u_{F\dot{\beta}i\{k\}}$ the angular velocities of i-th WMR proper wheels are calculated according to equation

$$\begin{bmatrix} z_{di2[1]\{k\}} \\ z_{di2[2]\{k\}} \end{bmatrix} = \frac{1}{r}\begin{bmatrix} v_M^* & \dot{\beta}^* l_1 \\ v_M^* & -\dot{\beta}^* l_1 \end{bmatrix}\begin{bmatrix} u_{Fvi\{k\}} \\ u_{F\dot{\beta}i\{k\}} \end{bmatrix}, \tag{10}$$

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.

3.3 Tracking Control System

The discrete tracking control system realises trajectory generated for an individual agent. It generates control signals for the WMR driving systems. Realisation of the tracking control signals allows the point A of the WMR to keeps its position in the moving virtual structure of the WMRF. The tracking control system uses NDP algorithms in Dual heuristic dynamic programming (DHP) configuration and was described in detail in (Hendzel and Szuster, 2010b). The overall tracking control signal consist of control signal generated by two NDP structures $\mathbf{u}_{A\{k\}} = \left[u_{A[1]\{k\}}, u_{A[2]\{k\}}\right]^{\mathrm{T}}$, the PD control signal $\mathbf{u}_{PD\{k\}}$, the supervisory term control signal $\mathbf{u}^*_{S\{k\}}$ derived from the Lyapunov stability theorem, and control signal $\mathbf{u}_{E\{k\}}$, that derives from discretisation of the WMR's model in the closed system loop. The overall tracking control signal is assumed in the form

$$\mathbf{u}_{\{k\}} = \frac{1}{h}\mathbf{M}\left\{-\mathbf{u}_{A\{k\}} + \mathbf{u}^*_{S\{k\}} - \mathbf{u}_{PD\{k\}} - \mathbf{u}_{E\{k\}}\right\}. \tag{11}$$

Scheme of the tracking control system for the i-th WMR Amigobot in the formation is shown in fig. 6.

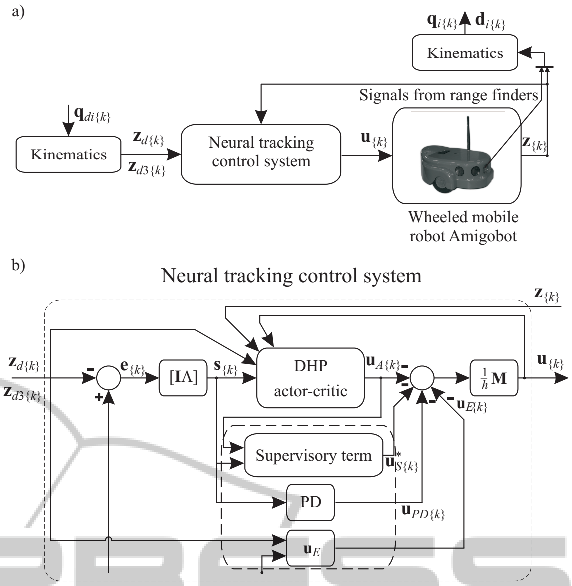

Figure 6: a) The tracking control system of the i-th mobile robot, b) the neural tracking control system.

On the basis of the simulated range finders measurements the proposed hierarchical control system generated the collision-free trajectory of the WMRF point M, and the tracking control signals for all agents, that allowed to realise the planed trajectory. The generated paths start in the point S, marked by the triangles on the figure, and ends in the goal G, marked by the **X** mark. In fig. 7 is shown the environment map with the path of the points A of all WMRs of the formation for the goal G(7.0, 9.0).

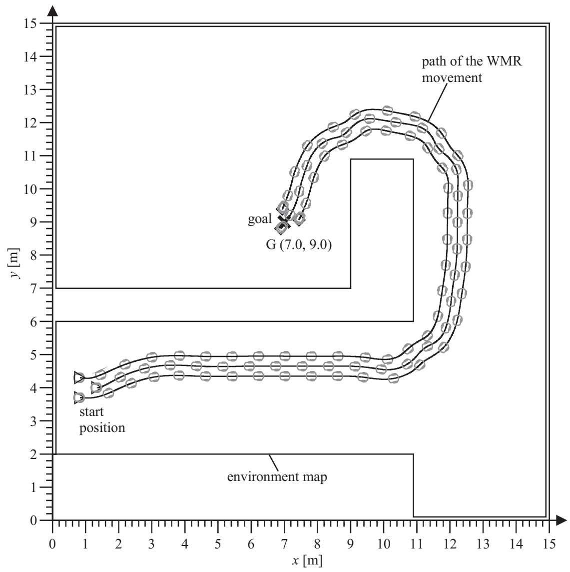

Figure 7: The environment map with the path of the wheeled mobile robots formation to the goal G(7.0, 9.0).

4 NUMERICAL EXPERIMENTS RESULTS

In this section, for the sake of simplicity, all variables are presented in a continuous domain of the time with the time axis t on diagrams, not as the discrete variables, and there is not used index k, $h = 0.01$ [s].

Taking into account the behavioural conception of the trajectory generating problem, the map of the environment was projected in the way, that the successive path can not be generated using only one of the behavioural control signals, for GS or OA task. The proposed navigator generates control signals, that make planning of the path in the complex task of obstacle avoiding and goal seeking possible. The localisations of obstacles were computed on the basis of simulated range finders readings, localisation and orientation of the WMRs in the modelled environment.

The overall trajectory generator control signals u_{Tv} and $u_{T\dot{\beta}}$, shown in fig. 8.a) and b), consists of control signals generated by the NDP structures in ADHDP configuration (u_{TAv}, $u_{TA\dot{\beta}}$) and control signals of the P controller (u_{TPv}, $u_{TP\dot{\beta}}$). The values of control signals generated by the P controller are small in a comparison with the ADHDP structures actors' control signals u_{TAv} and $u_{TA\dot{\beta}}$.

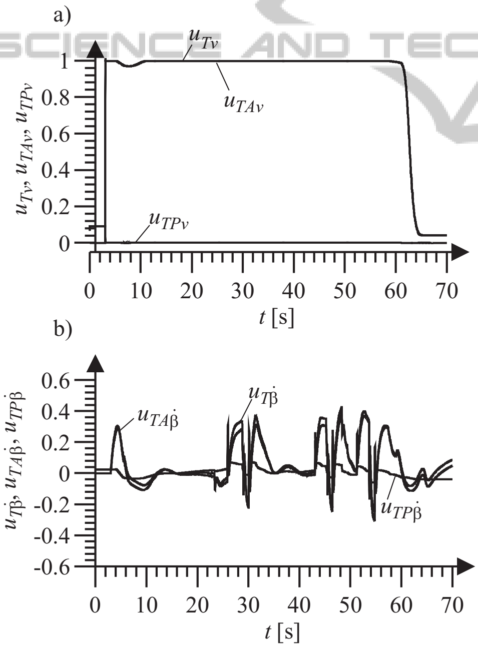

Figure 8: a) The navigator control signals u_{Tv}, u_{TAv} and u_{TPv}, b) the navigator control signals $u_{T\dot{\beta}}$, $u_{TA\dot{\beta}}$ and $u_{TP\dot{\beta}}$.

The distance to the goal G of the virtual structure M point is shown in fig. 9.a). It is consequently reduced during the numerical test, to the value near zero. The angle ψ_G is shown in fig. 9.b).

In fig. 10.a) and b) are shown values of the actor's ($\mathbf{W}_{TAv}$) and the critic's ($\mathbf{W}_{TCv}$) NN weights of the ADHDP structure, that generates the navigator's control signal u_{TAv}, that influences on the linear velocity of the WMRF virtual structure. Values of the actor's ($\mathbf{W}_{TA\dot{\beta}}$) and the critic's ($\mathbf{W}_{TC\dot{\beta}}$) NN weights

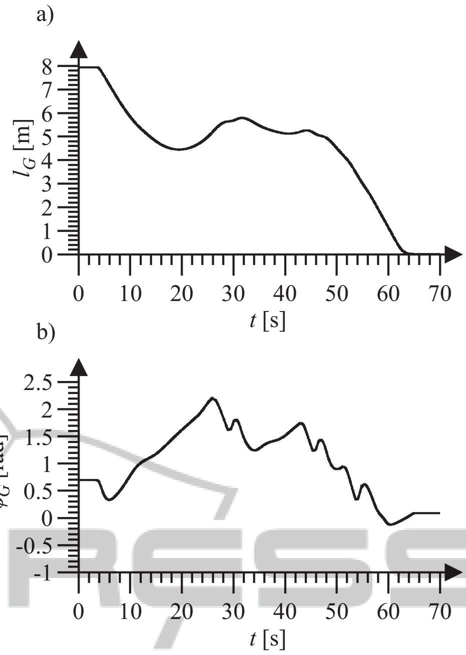

Figure 9: a) The distance to the goal G, l_G, b) the angle ψ_G.

of the ADHDP structure, that generates the navigator's control signal $u_{TA\dot{\beta}}$, are shown in fig. 10.c) and d). This signal controls the angle of the WMRF virtual structure turn. Weights of NNs are bounded and converge to the fixed values.

On the basis of the overall control signals generated by the navigator, u_{Tv} and $u_{T\dot{\beta}}$, taking into account position of the point A of i-th agent in the WMRF, desired angular velocities of the third WMR $z_{d32[1]}$ and $z_{d32[2]}$ (fig. 11.a)) were computed. The desired trajectory was realised, with the tracking errors shown in fig. 11.b), using the tracking control system with the overall tracking control signals $u_{3[1]}$, $u_{3[2]}$, shown in fig. 11.c).

In fig. 12.a) and b) are shown values of the actor's ($\mathbf{W}_{A31}$) and the critic's ($\mathbf{W}_{C31}$) NN weights of the DHP structure, that generates the tracking control signal $u_{[1]}$ for the third agent in the WMRF. Weights of NNs are bounded and converge to the fixed values.

5 SUMMARY

The proposed hierarchical control system generates the trajectory for the WMRF and realises tracking control of all agents. The generated trajectory is collision free and allows to reach the goal by the selected point of the WMRF. The sensor-based navigator was builded using NDP algorithms in ADHDP configuration. It is based on the behavioural control con-

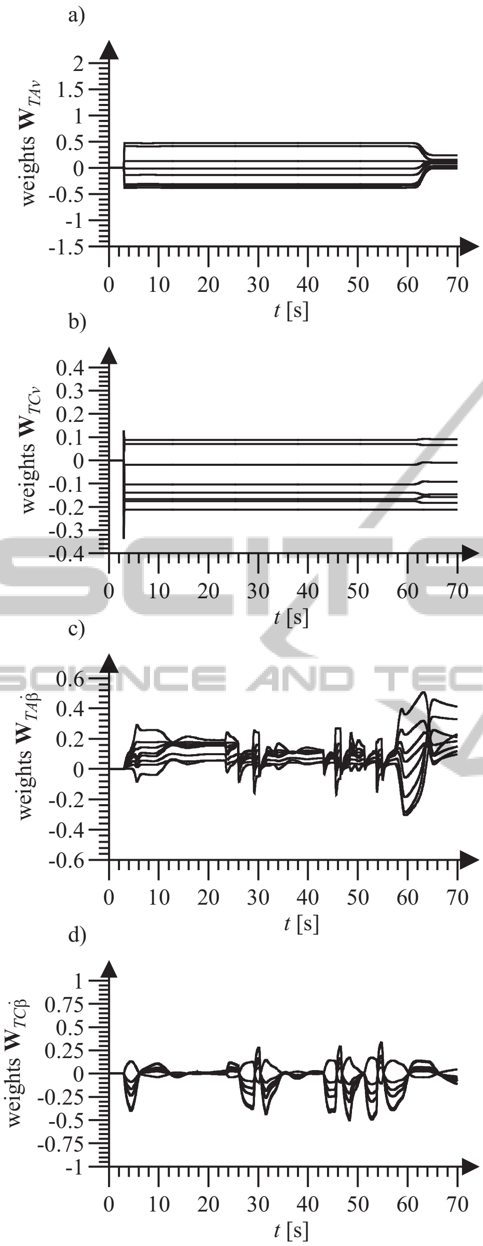

Figure 10: a) Weights of the ADHDP actor 1 NN $\mathbf{W}_{TAv}$, b) weights of the ADHDP critic 1 NN $\mathbf{W}_{TCv}$, c) weights of the ADHDP actor 2 NN $\mathbf{W}_{TA\dot{\beta}}$, d) weights of the ADHDP critic 2 NN $\mathbf{W}_{TC\dot{\beta}}$.

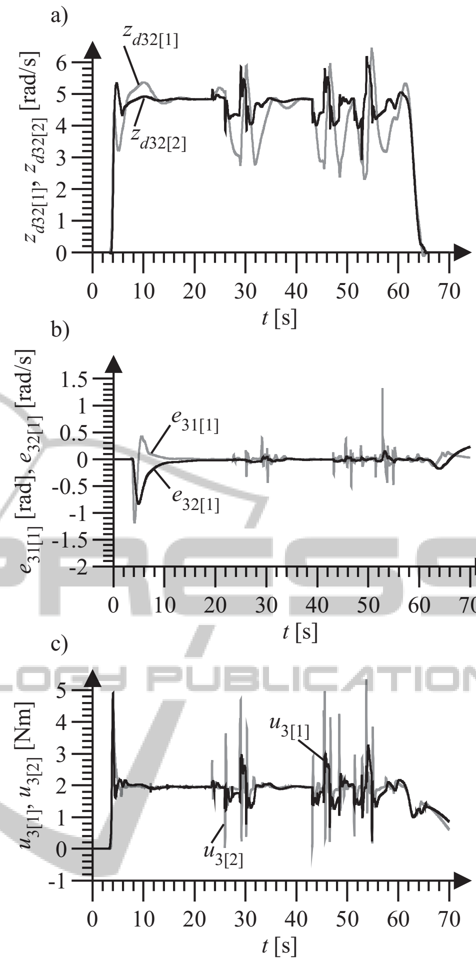

Figure 11: a) Desired $\left(z_{d32[1]}, z_{d32[2]}\right)$ angular velocities of the i-th agent, b) the tracking errors for the first $\left(e_{31[1]}, e_{32[1]}\right)$, c) the overall tracking control signals $u_{3[1]}$ and $u_{3[2]}$.

ception with use of only two individual behaviours: "goal seeking"and "obstacle avoiding", combined to generate collision free trajectory that allows to reach the goal. The new approach to the behaviour control used in the presented navigator allows to unite two individual behavioural control systems for individual behaviours into one structure. Appropriate position of the agent in the formation was ensured by the virtual structure control algorithm, with control signals derived using the Lyapunov stability theory. The trajectory generated for the individual agent was realised using the tracking control system with DHP structures. The proposed hierarchical control system using AI methods, works on-line and does not require the preliminary learning of NNs. Computer simulations conducted to illustrate the path planning process in the different environment conditions confirmed the correctness of the assumed conception of the WMRF hierarchical control. The selected point of the virtual structure reaches the goal, while no one of agents collides obstacles. The next step of the presented researches will be verification of the proposed control system using three Amigobot WMRs in the laboratory environment.

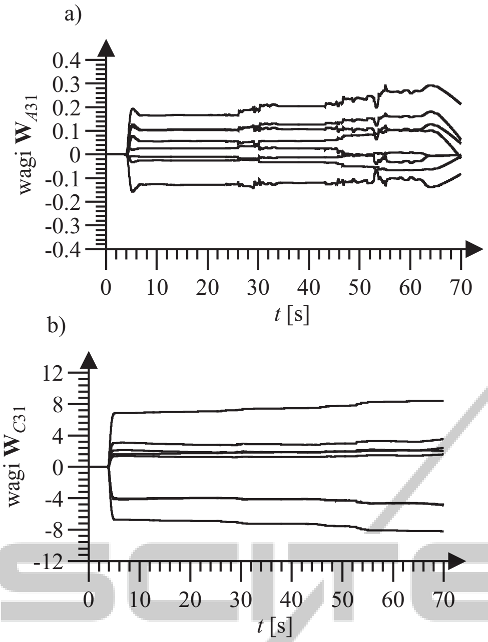

Figure 12: a) Weights of the DHP actor 1 NN $\mathbf{W}_{A31}$, b) weights of the DHP critic 1 NN $\mathbf{W}_{C31}$.

ACKNOWLEDGEMENTS

This paper is supported by Polish Government under Contract N N501 068838.

REFERENCES

A. Burghardt, T. B. and Giergiel, J. (2011). Control of robots' formation in unknown surroundings environment. In *Dynamical Systems. Nonlinear dynamic and control, Proc. of Conference on Dynamical Systems, Theory and Applications*. WPL.

Arkin, R. (1998). *Behavior-Based Robotics*. Intelligent Robots and Autonomous Agents. MIT Press.

Burghardt, A. (2004). Behavioural control of wheeled minirobot, (in polish). *PAK*, 11:26–29.

Burghardt, A. (2008). Proposal for a rapid prototyping environment for algorithms intended for autonomous mobile robot control. *Mechanics and Mechanical Engineering*, 12:5–16.

Egerstedt, M. and Hu, X. (2001). Formation constrained multi-agent control. *IEEE Transactions on Robotics and Automation*, 17(6):947–951.

Fahimi, F. (2008). *Autonomous Robots: Modeling, Path Planning, and Control*. Springer.

Giergiel, J. and Zylski, W. (2005). Description of motion of a mobile robot by maggie's equations. *J. Theor. App. Mech.*, 43:511–521.

Giergiel J., Hendzel Z., Z. W. (2002). *Modeling and control of wheeled mobile robots (in Polish)*. PWN.

Hendzel, Z. (2004). Fuzzy reactive control of wheeled mobile robot. *J. Theor. App. Mech.*, 42:503–517.

Hendzel, Z. and Szuster, M. (2010a). Discrete action dependant heuristic dynamic programming in wheeled mobile robot control. *Solid State Phenomena*, 164:419–424.

Hendzel, Z. and Szuster, M. (2010b). Discrete model-based adaptive critic designs in wheeled mobile robot control. In *Proceedings of the 10th international conference on Artifical intelligence and soft computing: Part II*, ICAISC'10, pages 264–271, Berlin, Heidelberg. Springer-Verlag.

Hendzel, Z. and Szuster, M. (2011). Neural dynamic programming in behavioural control of wheeled mobile robot (in polish). *Acta Mechanica et Automatica*, 5(1):28–36.

Hendzel, Z. and Szuster, M. (2012). Neural dynamic programming in reactive navigation of wheeled mobile robot. *LNCS*, 7268:450–457.

Maaref, H. and Barret, C. (2002). Sensor-based navigation of a mobile robot in an indoor environment. *Robotics and Autonomous Systems*, 38(1):1–18.

Millan, J. D. (1995). Reinforcement learning of goal-directed obstacle-avoiding reaction strategies in an autonomous mobile robot. *Robotics and Autonomous Systems*, 15:237–246.

Ogren, P. and Leonard, N. (2003). Obstacle avoidance in formation. Proceedings of 2003 ICRA.

Powell, W. B. (2007). *Approximate Dynamic Programming: Solving the Curses of Dimensionality (Wiley Series in Probability and Statistics)*. Wiley-Interscience.

Prokhorov, D. V. and Wunsch, D. C. (1997). Adaptive critic designs. *IEEE Transactions on Neural Networks*, 8(5):997–1007.

Si, J., Barto, A. G., Powell, W. B., and Wunsch, D. (2004). *Handbook of Learning and Approximate Dynamic Programming (IEEE Press Series on Computational Intelligence)*. Wiley-IEEE Press.

Sutton, R. and Barto, A. (1998). *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning. MIT Press.

Yamaguchi, H. (1997). Adaptive formation control for distributed autonomous mobile robot groups. In *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, volume 3, pages 2300–2305.