

Decision Making Process of Hexapods in a Model of Complex Terrains

Vladimir Socha^{1,3}, Patrik Kutilek¹, Alexandr Stefek², Lubos Socha³, Jakub Schlenker¹, Karel Hana¹

¹ Czech Technical University in Prague, the Faculty of Biomedical Engineering, Nam. Sitna 3105, 272 01, Kladno, Czech Republic, e-mail: vladimir.socha@fbmi.cvut.cz, kutilek@fbmi.cvut.cz, jakub.schlenker@fbmi.cvut.cz, hanakar1@fbmi.cvut.cz

² University of Defence, Faculty of Military Technology, Kounicova 65, 662 10, Brno, Czech Republic, e-mail: alexandr.stefek@unob.cz

³ Technical University of Košice, Faculty of Aeronautics, Rampová 7, 041 21, Košice, Slovakia, e-mail: lubos.socha@tuke.sk

Abstract: The paper describes behavior of a cognitive control system model, which enables a hexapod to walk in an obstacle-free terrain as well as in a complex terrain including obstacles. This cognitive system model is based on reinforcement learning and assumes the concept of static-stable walking. The decision making process was tested using three different types of terrain models. The results of decision making process trigger actions in the form of changes in the state of six-legged body to maintain stable walking forward. New methods have been developed to describe a group of obstacles of different sizes in a complex terrain. The results suggest a relationship between the predefined number of actions and the maximum total walked distance in terrain. In case of the terrain without obstacles, the optimized actions are the same. Thus, the way of moving the trunk and legs in the terrain is always the same and cyclic. The results also indicate that the maximum total walked distance is reduced due to a growing number of obstacles to overcome. The maximum total walked distance is reduced more significantly in the case of overcoming a greater number of small obstacles compared to the case of smaller number of large obstacles. The way of moving the trunk and legs in the terrain with large obstacles is acyclic. The methods proposed for the study of the cognitive system and the sensory system of a hexapod, for the simulation of six-legged walking, as well as for the characterization of terrain with obstacles may find application in bioengineering, robotics, military system and other fields.

Keywords: complex terrain; obstacles; hexapod; reinforcement learning; static-stable walking

1 Introduction

Hexapod represents a six-legged arthropod with a set of three pairs of legs controlled by a nervous system, [1] while relying on the concept of six-legged undercarriage previously introduced in robotics. The model of the cognitive system used for modelling of a six-legged insect's gait [2] employed a set of sensors (eyes) and actuators (leg muscles), [1, 2]. Generally, the cognitive system sends information to the motoric system and receives information from the sensory system, which is a concept known as the sensor-actuator loop [2]. The central cognitive system is responsible for the strategy of leg coordination and the transition from current-state to future-state [2], see Fig. 1.

The process of learning and decision-making in a cognitive system can be modelled using computational methods [2-4]. A number of methods based on artificial intelligence has been developed to control the trajectory of a walking hexapod robot [5-7], [8] however, these have not been designed for leg coordination [2]. There are also methods for controlling legs on a flat surface without any obstacles [9-13]. The methods do not offer solutions for crossing of small obstacles in the terrain [2] and only a few of the methods have been proposed to allow for this technique of walking [2, 14]. These methods are based i.e. on neural networks [15-16] or genetic algorithms [17]. Only a small number of methods is based on Reinforcement Learning (RL) [2, 3, 9, 10], which is accepted as a method that describes the decision-making process of living organisms [18-21]. The results of decision making process based on RL trigger in the form of changes in state of six-legged body to maintain stable forward walking [20]. An issue in evaluating the effectiveness of methods based on RL, however, is that the total walking distance in terrain covered by the number of actions of legs is affected not only by the algorithm parameter settings but the terrain complexity as well. This shortcoming thus calls for a method which would allow for the terrain complexity description.

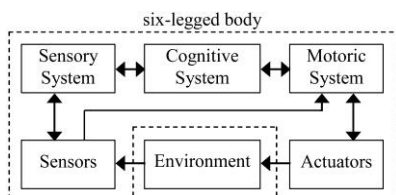


Figure 1

The sensor-actuator loop

A large number of methods focusing on the complexity of terrain has been designed in the field of robotic studies, and some of these are based on the description of rough terrain. Although procedures for planning foot trajectory and body trajectory of walking robots in 3D space have already been designed [22-25],

no method or variable to describe a group of obstacles of different sizes in a complex terrain corresponding to the walked distance or number of actions has been developed yet. In contrast with methods describing obstacles of different heights corresponding to the dimensions of the robot's trunk already introduced [26], a method of assessing lengths (horizontal sizes) of obstacles, however, has never been described comprehensively. A terrain analysis based on the envelope, slope and the curvature of the surface using certain scale has also been designed [27], however failed to provide characteristics of the geometry and size of obstacles in relation to the trunk and/or leg characteristics in terms of size and workspace.

Thus, the aim of this article is to introduce a method of describing a complex terrain and analyse the applicability of the method based on RL to enable stepping over obstacles in a terrain. The methods for the study of the six-legged walking, as well as for the characterization of terrain with obstacles could find application in robotics, military systems, rehabilitation and other fields.

2 Methods

2.1 Model of the Decision-Making Process of Hexapod

Walking through a complex terrain with obstacles is made possible through acyclic gait. The assumed model of cognitive system uses decision-making process based on RL for coordination of the legs of hexapod [2, 20, 21]. To ensure static stable locomotion [20, 21], the cognitive system uses the known states of the position of the legs to maintain static stability [13]. For the decision-making, knowledge of the condition for static-stable posture is assumed [20, 21]. The static-stable walking is represented by vector $t=(l_{R1}, l_{L1}, l_{R2}, l_{L2}, l_{R3}, l_{L3})$, $t \in T$, which describes the 15 states of six-legged body to maintain static stability [2], where binary (i.e. true/false) variables (l_{R1}, \dots, l_{L3}) represent the states of left (L) and right (R) legs: 1 - leg is in the swing phase; 0 - stance phase of the leg, [2, 20, 21].

To transport the trunk, the legs change their position in relation to the trunk. If the movement of each leg is autonomous, we can describe the position of each leg in leg workspace (LW) by the value n_i , [2]. This value represents the requirement for leg movement in LW, [23]. Maximum front position in LW is represented by the value of 0. Value 1 represents maximum requirement for leg movement of the leg in back position, [2]. Vector $r = (n_{R1}, n_{L1}, n_{R2}, n_{L2}, n_{R3}, n_{L3})$, $r \in R$, represents requirements for all legs movements.

Normally, the feet must not touch the obstacles, [22]. If we know where the obstacles are located in LW, it is possible to identify possible step lengths within the LW. Vector $p = (k_{R1}, k_{L1}, k_{R2}, k_{L2}, k_{R3}, k_{L3})$, $p \in P$, represents information about maximum possible step lengths k_i of all legs, [2].

The designed cognitive system uses Q-learning (QL) as a RL technique, [2, 20, 21], see Fig. 2. In the QL, the state-action pairs are represented by a Q-table, [2]. The Q-table stores the information about the relations between the states p , r and proper actions t . The actions represent changes in the state of the six-legged body to maintain static stable walking forward. States for each leg are represented by variables $n_i \in N$ and $k_i \in K$, and the actions of each leg are represented by $l_i \in L$ variables, [2].

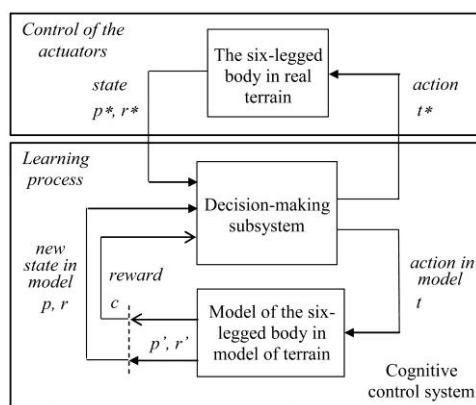


Figure 2

Cognitive control system based on reinforcement learning, six-legged body and terrain model [2]

Based on the information stored in the Q-table, the current situation is evaluated in order to select the best (most reward-promising) action to accomplish the task, [2, 21]. The new state, entered as a consequence of the execution of action, is evaluated by the reinforcement function. Its qualitative criterion (reinforcement) is used by the algorithm that adjusts the Q values, [2, 18, 19]. The RL algorithm applied is described in Table 1, [2]. The learning algorithm searches for possible options of walking and evaluates each action $t \in T$. The p^* and r^* are the initial states and the initial positions of the six-legged body in a terrain model. Let $MaxQ(p', r', t)$ be the maximum Q value of the next state (p', r') related to all the possible actions $t \in T$, [2], Table 1. For each new state of the six-legged body, different possible actions can be tested n times. The range of distance $l_{required}$ for the prediction of t actions in the terrain model is determined by the capabilities of the sensors. The variable for determining the (reached) total walking distance l_{walked} , i.e. maximum total walking distance after the learning process, is a predefined m number of t actions to cover this total distance. The value of c is the immediate (or expected) reward for the state change from the old state (p, r) to the

new state (p', r') . We assume that the success of the action t is defined by the covered distance represented by l , thus the reward is $c \approx l$, $0 \leq l \leq l_{max}$ [2]. The learning rate $\alpha(p, r, t)$ (see Table 1) is used to ensure the convergence of the iterative procedure, [2, 19]:

$$\alpha = \frac{\alpha_0 n_0}{n_0 + n_\alpha(p, r, t)} \quad (1)$$

$n_\alpha(p, r, t)$ is the number of times the $Q(p, r, t)$ value, i.e. state-action pair (p, r, t) , has been called during learning, [20]. The α_0 and n_0 are parameters used to control the convergence of the iterative procedure. Based on previous experiments, we set $\alpha_0=1$ and $n_0=1$. The γ is the discount factor in the range $0 \leq \gamma < 1$, we set $\gamma=0.9$ [20].

Table 1
Algorithm for RL of the hexapod cognitive control system [2, 21]

-
1. Set the initial parameters of the learning process
 2. Identify the state $p = p^* \in P$ and state $r = r^* \in R$ of the six-legged body in real terrain
 3. Select and execute an $t \in T$ action of the model of six-legged body in the model of terrain
 4. Identify the new state $p' \in P$ and $r' \in R$ of the model of six-legged body, and obtain the immediate reward c and update the learning rate α
 5. Update the Q-table, i.e. Q value:
 $Q(p, r, t) = Q(p, r, t) + \alpha (c + \gamma \text{Max}_{t'} Q(p', r', t) - Q(p, r, t))$
 6. Assign $p = p'$ and $r = r'$
 7. While $n \neq n_{repetitions}$ return to 3
 8. While $l_{walked} \neq l_{required}$ return to 2
 9. Execute the most appropriate action $t^* = t \in T$ of the six-legged body in real terrain
-

The strategy used to select the action during learning is based on Boltzmann's exploration [2, 20, 28]:

$$P(t) = \frac{e^{\frac{Q(p, r, t)}{E}}}{\sum_{t \in T} e^{\frac{Q(p, r, t)}{E}}} \quad (2)$$

where E is a parameter known as the computational temperature. High temperatures cause all actions to be nearly equiprobable, whereas low temperatures cause greedy action selections. The parameter value E decreases gradually by:

$$E_{n+1} = E_{min} + \beta \cdot (E_n - E_{min}) \quad (3)$$

where n is the number of the cycle repetition, i.e. the number of iteration cycles. Based on previous experiments and recommendations [2, 20, 28], the values of the parameter are $E_0 = E_{max} = 0.9$, $E_{min} = 0$ a $\beta = 0.9$. RL parameters and their descriptions can be found in [2, 20].

Table 1 presents the algorithm used to explore the possible states and actions of a six-legged body in a modelled terrain with obstacles. The cognitive system tries to cover a required maximum distance $l_{required}$, and thus the number of iterations is determined by the size of the terrain model. The number of $n_{repetitions} > 0$ if (repetitions) selecting an action t is given by the quality of the cognitive abilities of the system (we use $n=1$). The learning process is stopped when the required maximum distance $l_{required}$ or the number of cycles (i.e. locomotion) to walk through the same terrain model by an m number of t actions is achieved. Of course, the distance $l_{required}$ can be reached by a certain number of cycles to walk through the entire length of the terrain during learning. After learning, the final, optimized (t) actions at the (p, r) states are performed [2]. It is assumed that the total walked distance l_{walked} as well as the m number of t actions will decrease in a more complex terrain and increase in a less complex one. In the following sections the focus will be put on the study of the relationship between the number of cycles required to walk through the entire length of the terrain to achieve the maximum distance, number m of t actions to cover the maximum distance and the terrain complexity.

2.2 Method for Terrain Description and Gait Evaluation

The second aim of article is to introduce a method for describing a complex terrain and analyse the applicability of the method based on RL to enable stepping over obstacles in a terrain. However, a method for assessing the obstacles of different lengths has never been described comprehensively before. We assume that the cognitive system should select the best walking strategy to overcome the longest (i.e. maximum predefined) distance. Also, we assume that obstacles in a terrain are already identified, e.g. using methods based on recursive density estimation and evolving Takagi–Sugeno fuzzy systems [29] or 2D laser range finder and obstacle/gap detection based on edge detection [30]. An example of a terrain model with complicated distribution of obstacles is shown in Fig. 3 exported from the simulation software (MatLab, MathWorks Inc.) including descriptions.

To verify the designed RL based control system, it is necessary to describe lengths (horizontal sizes, [30]) of the obstacles in the direction of movement to consider only those obstacles which are on the path of the hexapod and corresponding with the section of the complex terrain, i.e. the covered total walking distance l_{walked} . Two types of obstacles need to be distinguished: small obstacles which can be stepped over by one leg without translating the trunk and large obstacles which have to be overcome by translating the trunk and can't be stepped over by a single leg movement.

The maximum possible step length d_{kSmax} in a terrain without obstacles is defined by the LW geometry [2, 17, 30]. The lengths of small obstacles are smaller or equal to the length of the maximum possible step d_{kSmax} , Fig. 3, [2]. The length of large obstacles is greater than the maximum possible step length d_{kSmax} , but still

has to be smaller than the maximum walking distance by one action t to ensure movement of the trunk in a terrain, i.e. l_{max} , see Fig. 3. Theoretically, l_{max} can reach up to $2 \cdot d_{kSmax}$, assuming the static-stable locomotion and geometry of the trunk and LW as described in [2, 21]. It also has to be ensured that two obstacles larger than maximum possible step length d_{kSmax} are not situated in both workspaces of the left and right leg at the same time, i.e. in the workspaces of the two opposite legs. If the condition above is met, then the path corresponding to the section of the complex terrain with obstacles, i.e. reached total walking distance l_{walked} , can be described by a numeral denoting the number of small obstacles and a numeral informing about the number of large obstacles, see Fig. 3. A more objective description would be the following: The path corresponding to the section of the complex terrain with obstacles, i.e. covered total walking distance l_{walked} , is described by numerals that represent the number of small obstacles to be overcome by the left side (LS) of the body (left legs), the number of the large obstacles to be overcome by the LS of the body (left legs), as well as the number of the small obstacles to be overcome by the right side (RS) of the body (right legs), and finally the number of large obstacles to be overcome by the RS of the body (right legs).

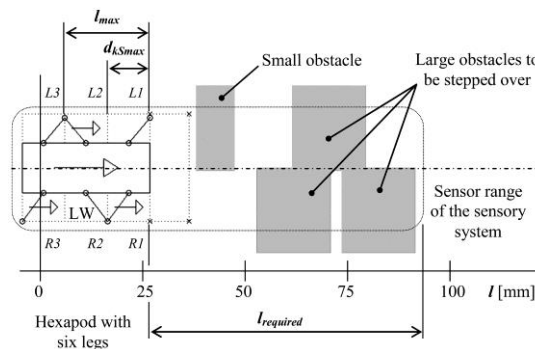


Figure 3

Simplified model of a six-legged body (with legs in swing/stance phase) and sensory information in a terrain model with obstacles [2]

The remaining question is what kind of relationship, if any, is there between the covered total walked distance l_{walked} , the predefined m number of t actions and the numbers of cycles (to reach the maximum total walked distance during learning) to walk over the terrain with specific types and numbers of obstacles. The information about possible relationship could be used to develop algorithms for effective RL, thus reducing the time and complexity of computing the learning process.

2.3 Verification Method

The verification is focused on testing the adopted approach to evaluate the hexapod's decisions in complex terrain. For this purpose, the second aim of article is to introduce a method of describing a complex terrain and analyse the applicability of the method based on RL to enable stepping over obstacles in a terrain. However, a method for assessing obstacles of different lengths has never been described comprehensively before. We assume that the cognitive system of hexapod should select the best walking strategy (i.e. the most appropriate actions t^*) to overcome the longest (i.e. maximum predefined) distance.

Table 2

Relationship between the predefined number of actions and the used types of complex terrains with small and large obstacles (LS – left side, RS – right side of the hexapod's body)

Terrain	Type of obstacles	Predefined number m of actions				
		3	6	9	12	15
		Number of obstacles to be overcome				
1.	small LS	0	0	0	0	0
	large LS	0	0	0	0	0
	small RS	0	0	0	0	0
	large RS	0	0	0	0	0
2.	small LS	2	3	4	5	6
	large LS	0	0	0	0	0
	small RS	0	2	3	4	4
	large RS	0	0	0	0	0
3.	small LS	2	3	3	3	4
	large LS	0	0	1	1	1
	small RS	0	1	1	1	2
	large RS	0	0	1	2	2

The method for the decision making based on RL described above was tested on models of terrain based on method for describing complex terrains in a MatLab software environment. The terrain model with obstacles is (on purpose) designed to be complex and ill-structured. It is assumed that the terrain model is to be explored and identified by the sensory system. The initial parameters used for the learning algorithm define the initial position of the six-legged body in the terrain model and the characteristics of terrain obstacles (dimensions and positions), see Fig. 3. Then the maximum distance $l_{required}=140$ mm from the six-legged body (i.e. trunk) to the farthest detected obstacle or target in the terrain is chosen. During the prediction of walking, the maximum number m of predicted t^* actions of the six-legged body in the model of terrain represents another limiting factor, see Table 2.

The issue with the evaluation of the effectiveness of setting algorithm parameters, however, is that the covered total walked distance l_{walked} achieved by the predefined m number of t actions is affected not only by the parameters set for the algorithm, but also by terrain complexity. Consequently, a more advanced testing session was performed using three types of terrains, Table 2.

The first type of terrain was a flat terrain without any obstacles; the second type of terrain involved only small obstacles; and the third type of terrain was covered with both small and large obstacles. To walk through the three types of terrains, the predefined m numbers of t actions were 3, 6, 9, 12 and 15. The predefined number of cycles to walk through the same terrain during learning was 2500.

2.4 Statistical Analysis

Each type of terrain was tested ten times by each predefined m number of actions. After calculating the covered total walked distance l_{walked} , the predefined m number of actions and the number of cycles to walk through the same terrain, the statistical analysis of these characteristics was performed using MatLab software. Maximum and minimum values were identified, the median, first quartile (Q1) and third quartile (Q3) was calculated for the number of cycles (i.e. locomotions) to reach the maximum total walking distance (i.e. select the optimized actions). The Jarque–Bera test (in MatLab software) was used to test the normal distribution of all parameters. The test returns the value of $h=1$ if it rejects the null hypothesis at the level of significance of 5%, and $h=0$ if does not.

3 Results

In this section, results for approach adopted to evaluate the hexapod's decisions by RL in three different complex terrains (see Table 2) are demonstrated. Each type of terrain was tested ten times by each predefined m number of actions and the Jarque–Bera test returned $h=1$ in all trials. The data were compared to identify the relationship between the predefined m number of t^* actions and the maximum total walking distance (l_{walked}), see Table 3 and Fig. 4. Using the RL method, the hexapod (i.e. model of the hexapod in MatLab software environment) is able to plan several actions in advance, see Table 3. The planned actions represent the planned changes in the state of the six-legged body to maintain static stable walking forward. The maximum total walking distance is achieved by the predefined number (m) of actions (t^*) after the learning process.

Table 3

Maximum total walked distances achieved by the predefined number of actions after the learning

Terrain	Predefined m number of actions t^*				
	3	6	9	12	15
	Maximum total walked distance [mm]				
1.	24	60	96	132	168
2.	24	49	64	79	94
3.	24	49	66	88	114

To determine the relationship between the number of cycles (i.e. locomotions) to reach the maximum total walking distance and the predefined m number of t^* actions, compare Table 4 and Fig. 5.

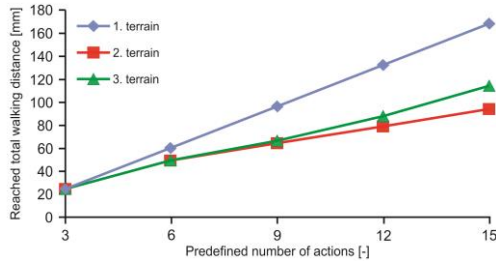


Figure 4

Diagram illustrating relationships between the maximum total walking distances and predefined numbers of actions in different terrains

Table 4

Numbers of cycles to reach the maximum total walking distance and to select the optimized actions in specific terrains

		Numbers of cycles to reach the maximum total walking distance														
		1 st terrain					2 nd terrain					3 rd terrain				
		Min	Max	Median	Q1	Q3	Min	Max	Median	Q1	Q3	Min	Max	Median	Q1	Q3
Predefined number m of actions	3	3	2037	45	12	59	26	1199	155	63	259	7	617	167	90	456
	6	76	2214	205	157	313	147	1686	331	239	696	108	1822	310	177	1294
	9	74	1268	211	118	730	169	2227	564	348	1109	223	1715	519	343	778
	12	59	1342	139	113	390	156	2373	485	267	1882	313	1633	1016	886	1539
	15	79	741	150	112	231	435	2025	938	816	1347	404	2058	1381	798	1715

It is clear that a higher value of the defined m number of actions increases the number of obstacles to be overcome, Table 2. Furthermore, the total walking distance is also increased, Table 3. When traversing the three different types of terrains, all obstacles (both small and large) were overcome.

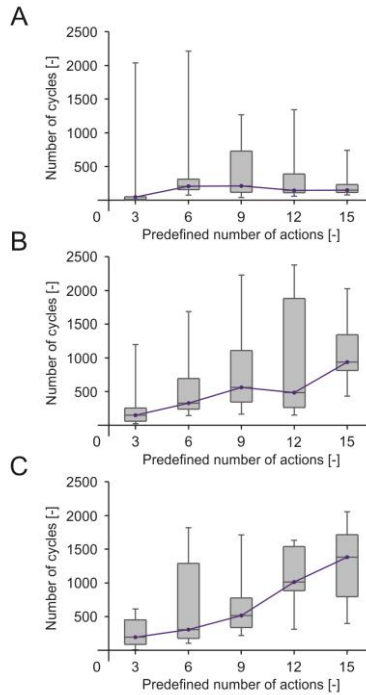


Figure 5

Diagram illustrating relationship between the number of cycles to reach the maximum total walking distance and the predefined numbers of actions in 1st terrain (A), 2nd terrain (B) and 3rd terrain (C)

4 Discussion

The data presented in Fig. 5B and Fig. 5C point out to the relationship between the predefined number of actions and the number of cycles (i.e. locomotions) to reach the maximum total walking distance and indicate the increase of the median of the number of cycles to reach the maximum total walking distance, helping to select the optimized actions, in the case of the 2nd and 3rd terrain. The data presented in Fig. 5A indicate a gradual increase and a subsequent decrease of the median of the number of cycles to reach the maximum total walking distance, and help select optimized actions, in the case of the 1st terrain (without obstacles). This is due to

the fact that the identified states (p, r) of the six-legged body and terrain are the same or similar during walking, and thus the optimized actions are the same and computation time of the learning process to find the optimized actions can be reduced. In all cases, the Jarque–Bera test confirmed that we can reject the hypothesis that the number of cycles to reach the maximum total walking distance has a normal distribution. In the case of the 1st and the 2nd terrain (including only small obstacles), strong asymmetrical (skewed) distributions of the number of cycles to reach the maximum total walked distance is found. The median of the number of cycles to reach the maximum total walking distance during learning is much lower than the maximum value, see Table 4. Thus, the maximum total walking distance is achieved by a lower number of the cycles. In the case of the 3rd terrain with large obstacles, the maximum total walking distance is achieved by a higher number of the cycles.

The data also indicate that the increase in the number of cycles to reach the maximum total walked distance is hampered (i.e. reduced) as the identified states (p, r) of the six-legged body and terrain are similar during walking in the case of the 1st and the 2nd terrain (including only small obstacles) and predefined greater number of actions. In the case of the 3rd terrain (with small and large obstacles), the rate of increase in the number of cycles is not constant or reduced since the terrain is complex and the identified states (p, r) of the six-legged body and terrain are not identical during walking. In general, moving over a more complex terrain reduces the number of actions and results in higher values of maximum number of cycles (i.e. locomotions). On the other hand, the maximum total walking distance is lower as terrain complexity reduces the total number of possible actions denoted by t .

In view of the adjustment of learning process described above, the approximate number of cycles (i.e. locomotions) sufficient for the subsequent selection of the optimal action t can reach up to 1500. The approximate number of cycles can be lower than 1000 in the case of a less complex terrain and the predefined $m=9$ number of t actions. With a more complex terrain, the approximate number of cycles can be lower than 1000 in the case of a lower predefined ($m=3$) number of t actions. In general, the use of a lower predefined ($m=3$) number of actions is appropriate for a complex terrain.

The data suggesting the relationship between the predefined number of actions and the maximum total walking distance, Fig. 4, indicate an increase in the maximum total walking distance in all cases, i.e. the terrains are passable. In case of the 1st terrain (without obstacles), the curve of the increase in the maximum total walking distance is constant because the optimized t actions are the same. Thus, the type of moving the trunk and legs in the terrain is always the same. For static stable walking forward, the model of hexapod can use acyclic and/or cyclic gait. The type of gait depends on the results of the search for the most appropriate sequence of actions to achieve the maximum total walked distance by the predefined number of actions. In case of the 1st terrain, model of the hexapod

opted for cyclic gait, as is apparent from Fig. 4, where the walked distance after the individual actions remains the same. The decision making process of the hexapod shows that acyclic gait suits a very complex terrain and cyclic gait on the other hand a less complex terrain. The data in Fig. 4 also indicate that the maximum total walking distance is reduced due to a growing number of obstacles to be overcome (Tab. 2), especially in the 2nd terrain. The maximum total walking distance is reduced more significantly in the case of overcoming a greater number of small obstacles (2nd terrain) than in the case of smaller number of large obstacles (3rd terrain).

Proceeding from the above analysis, it was found that the results show the suitability of the adopted approach. The methods proposed for the study of the six-legged walking, as well as for describing a terrain with obstacles, allow us to quantitatively assess and evaluate the hexapod's decisions in complex terrain. It of course follows that, alternatively, other nonlinear methods adapted for hexapod's decisions in complex terrain could be used, e.g. the tensor product models [31], methods for delivery vehicle routing problem [32], nonlinear multivariable systems using recurrent cerebellar models [33], neural control mechanisms [16], etc. Application of these methods to the problem of hexapod's decisions may find its place in the context of follow-up studies. However, RL method was used as RL is widely accepted as a tool to understand the goal-directed behavior of real organisms (including six-legged insects) that learn and interact with their environment in real time. Thus the objective of RL is to select actions so as to maximize their long-term rewards [34]. Regarding the methods to describe the complexity of the terrain, no similar method to describe the complexity of terrain with respect to the hexapod geometry has been introduced before. In the past, methods for describing obstacles in a terrain were only mentioned [30]. Not only does our approach allow for the description of the complexity of a terrain, it also provides description of the size of the obstacle with respect to LW [2, 23]. Thus, the proposed method may complement existing methods designed for trajectory planning [5-8].

Conclusions

The proposed techniques for describing complexity of a terrain and hexapod's decisions in the terrain were described, tested, and verified in the article. In order to meet the demands for quantitative description of the terrain, the method describing lengths (horizontal sizes) of the obstacles in the direction of the hexapod's movement was designed. The technique for decision making process of hexapod based on RL selects the most suitable action for each state to overcome the longest (i.e. maximum) distance in the terrain which includes obstacles. The methods for the study of the decision making process of hexapods, as well as for the simulation of six-legged walking as well as describing terrains with obstacles may find application in bioengineering, robotics, military systems and other related fields.

In future works, the problem of vertical size of obstacles in terrain and the problem of accelerated motion in terms of its dynamics will be studied. The proposed methods rely on the movement of most insects since they move slowly, and the problem of vertical size of obstacles has been partially solved in [22]. However, vertical movement of the legs is determined exclusively by the height of the obstacles in direction of walking [30] and does not affect the actions (i.e. the strategy of leg coordination of hexapod and the transition from current-state to future-state) determined by the methods described in this article.

Acknowledgement

This work was done in the framework of CTU project SGS15/107/OHK4/1T/17. The authors would like to thank Andrej Madoran, BA, for the translation of this work.

References

- [1] Barfoot T., Earon E., D’Eleuterio G.: Experiments in Learning Distributed Control for a Hexapod Robot, *Robotics and Autonomous Systems*, Vol. 54, 2006, 864-872
- [2] Barfoot T., Earon E., D’Eleuterio G.: A Step in the Right Direction – Learning Hexapod Gaits through Reinforcement, *Proceedings of the International Symposium on Robotics (ISR)*, Quebec, Montreal, 2000; pp. 487-492
- [3] Parker G.B., Mills J.W.: Metachronal Wave Gait Generation for Hexapod Robots, *Proceedings of the World Automation Congress (WAC)*, USA, Anchorage, 1998; pp. 365-370
- [4] Youcef Z., Pierre C.: Control of the Trajectory of a Hexapod Robot based on Distributed Q-learning, *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE)*, France, Ajaccio, 2004; pp. 277-282
- [5] Touzet C.: Q-learning for Robots, *The Handbook of Brain Theory and Neural Networks*, Cambridge: MIT Press, 2003; pp. 934-937
- [6] Khriji L., Touati F., Benhmed K., Al-Yahmedi A.: Mobile Robot Navigation based on Q-Learning Technique, *International Journal of Advanced Robotic Systems*, 2011; Vol. 8, pp. 45-51
- [7] Porta J., Celaya E.: Efficient Gait Generation using Reinforcement learning, *Proceedings of the 4th International Conference on Climbing and Walking Robots (CLAWAR)*, Germany, Karlsruhe, 2001; pp. 411-418
- [8] Zeidan B.; Sakyasingha Dasgupta S., Wörgötter F., Manoonpong P.: Adaptive Landmark-based Navigation System Using Learning Techniques, *Lecture Notes in Computer Science*, 8575: 121-131

-
- [9] Porta J.: Rho-LEARNING: A Robotics-oriented Reinforcement Learning Algorithm, Technical Report IRI-DT-00-03, Institut de Robotica i Informatica Industrial, Barcelona, 2000
- [10] Espenschied K.S., Quinn R.D., Chiel H.J., Beer R.D.: Leg Coordination Mechanisms in Stick Insect Applied to Hexapod Robot Locomotion, *Adaptive Behaviour*, Vol. 1, 1993, pp. 455-468
- [11] Porta J., Celaya E.: Walking in Unstructured Natural Environments, *Proceedings of the European Workshop on Hazardous Robotics (HEROS)*, Spain, Barcelona, 1996; pp. 99-107
- [12] Aparna K., Geeta S.: Insect Inspired Hexapod Robot for Terrain Navigation, *Journal of Research in Engineering and Technology*, Vol. 2, 2013, pp. 63-69
- [13] Tedeschi F., Carbone G.: Design Issues for Hexapod Walking Robots, *Robotics*, Vol. 3, No. 2, 2014, pp. 181-206
- [14] Juang C., Chang Y., Hsiao C.: Evolving Gaits of a Hexapod Robot by Recurrent Neural Networks with Symbiotic Species-based Particle Swarm Optimization, *IEEE Transactions on Industrial Electronics*, 2011; 58: 3110-3119
- [15] Belter D., Skrzypczynski P.: A Biologically Inspired Approach to Feasible Gait Learning for a Hexapod Robot, *International Journal of Applied Mathematics and Computer Science*, 2010; Vol. 20, pp. 69-84
- [16] Goldschmidt D., Wörgötter F., Manoonpong P.: Biologically-inspired Adaptive Obstacle Negotiation Behavior of Hexapod Robots, *Frontiers in Neurobotics*, Vol. 8, 2014, pp 1-16
- [17] Irodova M., Sloan R.: Reinforcement Learning and Function Approximation, *Proceedings of the Eighteenth International Florida Artificial Intelligence Research Society Conference (FLAIRS 05)*, USA, Clearwater Beach, 2005; pp. 455-460
- [18] Sutton R.S., Barto A.G.: *Reinforcement Learning: An Introduction*, Cambridge: MIT Press, 1998
- [19] Kutilek P., Kacer J.: The Locomotion Control of the Concyclically Walking Carriage, *Cybernetics Letters*, Vol. 3, No. 1, 2005, p. 8
- [20] Hrdlicka I., Kutilek P.: Reinforcement Learning in Control Systems for Walking Hexapod Robots, *Cybernetics Letters*, Vol. 3, No. 1, 2005, p. 13
- [21] Belter D., Skrzypczynski P.: Integrated Motion Planning for a Hexapod Robot Walking on Rough Terrain, *Proceedings of the 18th World Congress of the International Federation of Automatic Control (IFAC 2011)*, Italy, Milano, 2011; pp. 6918-6923

- [22] Hauser K., Bretl T., Latombe J. C., Harada K., Wilcox B.: Motion Planning for Legged Robots on Varied Terrain, *International Journal of Robotics Research*, Vol. 27, 2008, pp. 1325-1349
- [23] Görner M., Chilian A., Hirschmüller H.: Towards an Autonomous Walking Robot for Planetary Surfaces, *Proceedings of the 10th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS 2010)*, Japan, Sapporo, 2010; pp. 170-177
- [24] Rebula J. R., Neuhaus P. D., Bonnlander B. V., Johnson M. J., Pratt J. E.: A Controller for the LittleDog Quadruped Walking on Rough Terrain, *Proceedings of the 2007 IEEE International Conference on Robotics and Automation (ICRA 2007)*, Italy, Rome, 2007; pp. 1467-1473
- [25] Palis F., Rusin V., Schmucker U., Schneider A., Zavgorodniy Y.: Walking Robot with Articulated Body and Force Controlled Legs, *Proceedings of the Research on Adaptive Motion in Animals and Machines (AMAM 2005)*, Germany, Ilmenau, 2005; pp. 1-6
- [26] Pettersson L.: Terrain Analysis as a Design Tool for Autonomous Vehicles in Difficult Terrain, *Proceedings of the Second NordDesign*, Sweden, Stockholm, 1998; pp. 1-10
- [27] Celaya E., Porta J.: Force-based Control of a Six-legged Robot on Abrupt Terrain using the Subsumption Architecture, *Proceedings of the International Conference on Advanced Robotics (ICAR '95)*, Spain, Sant Feliu de Guixols, 1995; pp. 413-419
- [28] Kianercy A., Galstyan A.: Dynamics of Boltzmann Q-learning in Two-Player Two-Action Games, *Physical Review E*, Vol. 85, No. 4, 2012, pp. 1-10
- [29] Angelov P., Sadeghi-Tehran P., Ramezani R.: An Approach to Automatic Real-Time Novelty Detection, Object Identification, and Tracking in Video Streams based on Recursive Density Estimation and Evolving Takagi-Sugeno Fuzzy Systems, *International Journal of Intelligent Systems*, Vol. 26, No. 3, 2011, pp. 189-205
- [30] Kesper P., Grinke E., Hesse F., Wörgötter F., Manoonpong P.: Obstacle/Gap Detection and Terrain Classification of Walking Robots based on a 2D Laser Range Finder, *Proceedings of the 16th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines (CLAWAR 2013)*, Australia, Sydney, 2013, pp. 419-426
- [31] Precup R., Dragos C., Preitl S., Radac M., Petriu E.: Novel Tensor Product Models for Automatic Transmission System Control, *IEEE Systems Journal*, Vol. 6, No. 3, 2012, pp. 488-498

- [32] Khmelev A., Kochetov Y.: A Hybrid Local Search for the Split Delivery Vehicle Routing Problem, *International Journal of Artificial Intelligence*, Vol. 13, No. 1, 2015, pp. 147-164
- [33] Chen C. H., Chung C. C., Chao F., Rudas I. J., Intelligent Robust Control for Uncertain Nonlinear Multivariable Systems using Recurrent Cerebellar Model Neural Networks, *Acta Polytechnica Hungarica*, Vol. 12, No. 5, 2015, pp. 7-33
- [34] Ludvig E. A.: Reinforcement Learning in Animals, *The Encyclopedia of the Sciences of Learning*, New York: Springer, 2012, pp. 2799-2802