REINFORCEMENT LEARNING BASED OF FIVE LEGS ROBOT FOR RESCUE OPERATIONS

Prihastono
Bhayangkara Surabaya University and PhD Student, Sepuluh Nopember Institute of Technology
INDONESIA
prihastono@ubhara.ac.id

Dr. Son Kuswadi
Electronic Engineering Polytechnic Institute of Surabaya
INDONESIA
sonk@eepis-its.edu

Prof. Dr. Achmad Jazidie
Sepuluh Nopember Institute of Technology
INDONESIA
jazidie@ee.its.ac.id

ABSTRACT

This research developed small scale prototype five legs robot which can be used for searching victims in tsunami disaster. The robot used dc servomotor as the actuator, ultrasonic range finder, magnetic compass, and limit switch as the sensor. Remote control is used to operate the robot, and wireless camera for the visualization. To implement intelligent control system in robot which not depend on the model, both of dynamic system model and dynamic environment model, this research use "behavior based" algorithm. Learning robot ability in navigation is developed using "reinforcement learning" algorithm. Learning processes are applied by interaction between system and environment with "reward" and "punishment" rule. In this case, the environment is everything around the robot and human is one of the environment. It will be very interesting if human and robot can interact each other. Everything that human want will be understood and then will be executed by robot Three street conditions are tested for robot's performance and the result will be discussed.

Key word: five legs robot, reinforcement learning, reward and punishment.

INTRODUCTION

Contribution of robotics technology to today's sophisticated tasks is an inevitable progress, leading to a gradual minimization of human share, mostly, due to saturation in improvements of human abilities or to complementation of human activities. Education and training are insufficient for dealing with the complex and exhaustive tasks (Zhiying, 2007). Thus, from the robotics point of view, the trend is to provide an intelligent versatile tool to be a complete substitution of human in risky operations and complement human operations when auxiliary intelligent dynamics are required for extra dexterity. As a part of this progress, Search and Rescue (SAR) is the one of the most crucial fields that needs robotics contribution.

Search and rescue (SAR) robotics can be defined as the utilization of robotics technology for human assistance in any phase of SAR operations (Inagaki, 1997). Robotic SAR devices have to work in extremely unstructured and technically challenging areas shaped by natural forces. One of the major requirements of rescue robot design is the flexibility of the design for different rescue usage in disaster areas of varying properties. Rescue robotic devices should be adaptable, robust, and predictive in control when facing different and changing needs. Intelligent, biologically inspired mobile robots and, in particular hexapod walking robots have turned out to be widely used robot types

Copyright © 2011 SAVAP International
www.savap.org.pk

beside serpentine mechanisms; providing effective, immediate, and reliable responses to many SAR operations.

Legged locomotion offers a significant potential for mobility over highly irregular natural rough terrains cut with ditches and high unpredictable in comparison to wheeled or tracked locomotion (Teruya, 2005), (Richard, 1998). Legs can provide the capabilities of stepping over obstacles or ditches, and maneuvering within confined areas of space. They can handle with softness, the unevenness of the terrain. Beside their main function in locomotion, legs are almost in every external process of animals. The articulated structures of legs serve as manipulators to pull, push, hold, etc. or as tactile sensors to explore the environment.

We should select the number of legs carefully by considering their locomotive environments, because each walking machine has peculiar merits. In case of heavy walking environments, hexapod walking may be more suitable than quadruped walking. If one of the six legs is broken down, continuation of the static walking may be ensured by the left five legs, while some parts of walking functions are reduced. In case of quadruped walking robot, it may not be possible to continue their static walking.

The research of legged robot was rapidly developed. It can be seen from recent ideas about new systems of mechanism of robot that take ideas from nature, called biology inspired robot. The proposed robot is mechanism five legs that inspired from sea star (phylum echynodermata). The robot, in which each of its leg may became front centre, make it possible to maneuver difficult done by the other legged robot. In addition, the robot can rotate in its centre of its body easily, that the other legged robot difficult to do so.

THEORY AND SYSTEM DESIGN

The learning of five-legged walking is realized with the actual robot. This section introduces the system design of robot and application of reinforcement learning.

A. System Hardware Specification

The amount of joints on each leg will decide the Degree of Freedom (DOF) where in this case there will be 3 DOF for each leg.

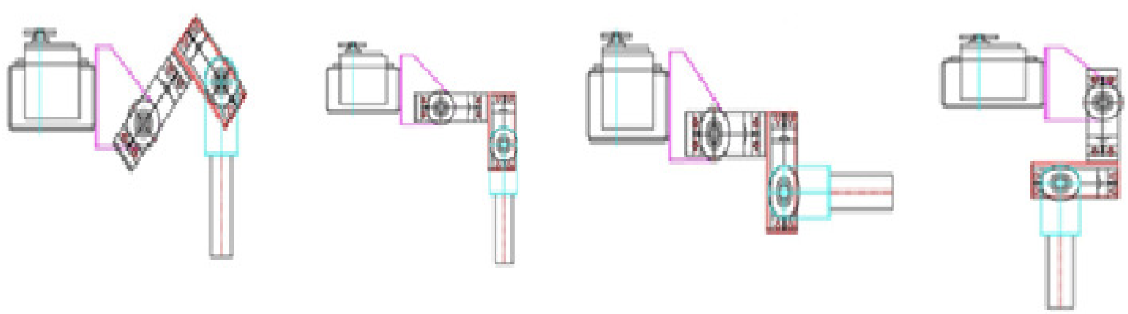


Figure 1 *Example of leg with 3 DOF*

The gait is an order for the lift and release activity on each legs. It's depends at the amount of the legs. The number of possible gate event can be formulated with : **N=(2k-1)!**

Angle joint on hinge base of the legs is referred to the point that can be reached by leg movement of joints that are most close to the body of robot (center of body). Because there are five legs, between the legs with one leg to the other leg is 72° away. So that each leg can move 72° (clockwise) and -72° (counter clockwise). But in reality, robot can only reach -50° to 50° because of mechanical robot legs require some space. Angle joint on hinge base of the legs is shown in Figure 2.

ISSN: 2223-9553
Volume 1, Issue 2, September 2011

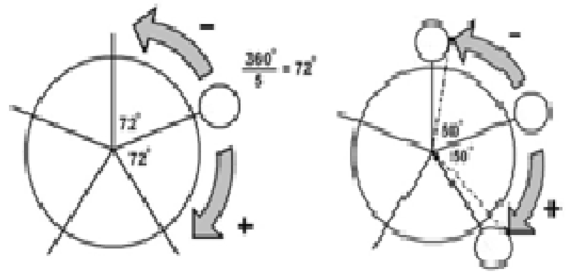


Figure 2. Angle joints which formed the base of leg and reduction of the base angle joints legs.

Velocity of robot is resulted by time delay every movement joint on the robot legs. This parameter is given on the value of the function of delay in the program microcontroller. Before calling this function, interrupt will be turned off, and time delay will be longer than expected.

The system is divided into several sections: actuator and sensor selection, mechanical and hardware building, and next reinforcement learning algorithm implementation in robot. The type of Actuator that we select is standard servo-motor (180 degree rotation) and we choose Hi-Tech servo-motor type HS-422 because it has several advantages compared with other servo motor. One of the benefits, it have a high enough torque (4.1kg.cm/56.93oz.in) and supported with the gear made of iron (iron dual-oilite bushings), making it possible to drive a significant burden. To detect the distance with robot obstacle (obstruction), Ping)))™ Ultrasonic Range Finder has been used which have the ability to measure the distance of objects as far as 3 cm to 300 cm. Limit switch is used to detect position of the robot. Limit switch installed in each leg. When the limit switches are touching objects, it will send a signal to the microcontroller to be processed further. We used two types of microcontroller, first is ATtiny2313 to control ultrasonic sensors and send the data series (USART) to ATmega16 handled limit switches and moved servo motor. Block diagram of system is shown in Figure 3.

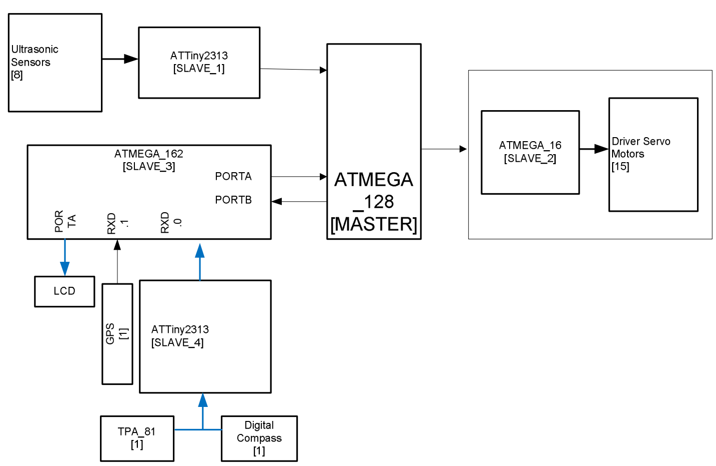


Figure 3. Block diagram of system.

In nature, five legs design can be found on the star sea. The design is inspiring writers to develop this type of robot. Mechanical design of robot developed with five-legged actuator. In every leg, there is 3 DOF as a reference for the movement of robot. Here's a picture viewed from the front of robot.

Copyright © 2011 SAVAP International
www.savap.org.pk

Robot kinematics mechanism

In the following is a robot kinematics analysis in brief. Fig. 4 shows robotic diagram and its notation.

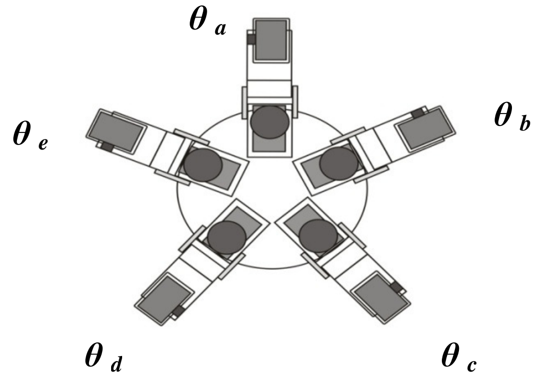


Figure 4. Robotic diagram and its notation used for kinematic analysis

Fig. 5 shows notation for kinematics development of one leg of the robot.

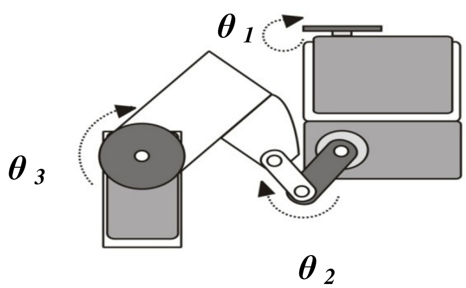


Figure 5. Notation for one legged

Since robot has five leg, and every leg has three angle of robot movement, then the forward kinematics of the overall robot is :

$$\boldsymbol{q}(x,y,z) = f\,(\theta_{a1},\theta_{a2},\theta_{a3},\theta_{b1},\theta_{b2},\theta_{b3},\theta_{c1},\theta_{c2},\theta_{c3},\ \theta_{d1},\theta_{d2},\theta_{d3},\ \theta_{e1},\theta_{e2},\theta_{e3}) \tag{1}$$

In this research we will not going into detail of derivation of kinematics equation, since we use different approach to control the movement of robot.

B. System Software

The sources of intelligence for primitively facilitated robotic agents exist in their reconceptualizing capability of what they physically sense from the environment. To implement this capability, we adopt an inductive learning method of conceptual formation for adaptively organizing a state space and develop a new algorithm for a robot to construct its task-relevant state space efficiently through matching encountering states with the similar situations in the past and through generalizing them. This research propose a methodology to dynamically increase the resolution of state spaces both adaptively and selectively by applying a concept formation technique from machine learning recursively to a record of a sensorimotor history of a learning agent. By connecting this with conventional reinforcement learning, we showed our algorithm can perform tasks without suffering from hidden state problems in an artificial maze environment, and also present its robustness even for a robot whose perceptual resources are quite restricted and/or bounded.

Copyright © 2011 SAVAP International

B.1.Reinforcement Learning

Reinforcement learning is learning what to do--how to map situations to actions--so as to maximize a numerical reward signal. The learner is not told which actions to take, as in most forms of machine learning, but instead must discover which actions yield the most reward by trying them. In the most interesting and challenging cases, actions may affect not only the immediate reward but also the next situation and, through that, all subsequent rewards. These two characteristics (trial-and-error search and delayed reward) are the two most important distinguishing features of reinforcement learning. Reinforcement learning is defined not by characterizing learning methods, but by characterizing a learning problem. The basic idea is simply to capture the most important aspects of the real problem facing a learning agent interacting with its environment to achieve a goal. The formulation is intended to include just these three aspects (sensation, action, and goal) in their simplest possible forms without trivializing any of them.

Reinforcement learning is different from supervised learning. Supervised learning is learning from examples provided by a knowledgeable external supervisor. In uncharted territory-- where one would expect learning to be most beneficial--an agent must be able to learn from its own experience. To obtain a lot of reward, a reinforcement learning agent must prefer actions that it has tried in the past and found to be effective in producing reward. When reinforcement learning involves supervised learning, it does so for specific reasons that determine which capabilities are critical and which are not. For learning research to make progress, important sub-problems have to be isolated and studied, but they should be sub-problems that play clear roles in complete, interactive, goal-seeking agents, even if all details of the complete agent cannot yet be filled in (Richard, 1998). The reinforcement learning algorithm is shown in Figure 6.

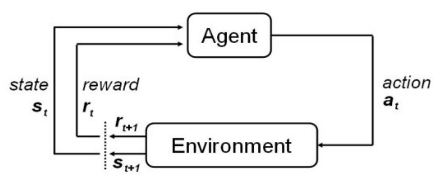


Figure 6. Reinforcement Learning Architecture.

This important thing for the learner and environment is:

- Agent : Learning and decide action
- Input : state (*st*) and reward (*rt*)
- Release : action (*at*)
- Goal : to maximize the amount of reward
- Environment:
    - All outside the agent (learner)
    - React to the action with a new state
    - Contains functions that generate reward or punishment

B.2.Related Work

Reinforcement learning (RL) has been used to train robot. Expressing such beliefs in a manner that can be naturally and effectively exploited by an reinforcement learning system can be a challenge. At the same time, it holds the promise of helping reinforcement learning to scale up to complex real-world tasks. Here is some research that use reinforcement learning to maximize gait on legged robot.

EXPERIMENT

The hip leg center position is in every corner of the fixed pentagon as a reference for the position and servo position at 0^0.

Copyright © 2011 SAVAP International
www.savap.org.pk

ISSN: 2223-9553
Volume 1, Issue 2, September 2011

Forward Movement

The center front leg work as a sweeper while the other four legs can move into 50^0 position, the leg movement combination that work as a balance are when the front leg is in opposite direction of the back leg

Backward Movement

The backward movement is actually the same as the forward movement, the difference is that four of the leg move in -50^0 direction while the front leg help push backward.

Askew Movement

There are two askew movement that can be done, which is the right askew and the left askew.

The askew movement is also the same with the forward movement, the only difference is that while doing the left askew, the front and the back left leg have a $< 50^0$ movement. In the other hand, while doing the right askew, the front and the back right leg have a $< 50^0$ movement.

Turnback Movement

This turnback movement consist of two movements, which is cw and ccw that controlled by the rotation of the leg, whether is in 50^0 or -50^0 direction. The turn movement for the turn right movement are the front leg turn around -50^0 folowed by the right front leg, right back leg, left back leg, right front leg, and continued by all of the hip servo at 0^0 position.

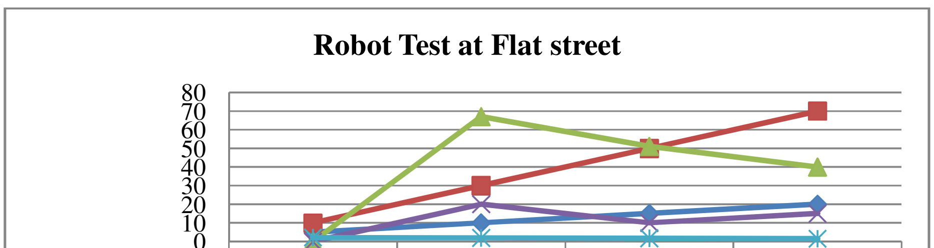
The performance of robot are tested in three conditions (flat street, bumpy and sloped street) by comparing the analyzed parameters :

1. Speed
2. Slope Error
3. Current consumption

RESULT AND DISCUSSION

4.1 Robot in flat street

The robot is tested to run on flat street (in this case is on wooden floor). The results are displayed on Fig.7 below.



	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Delay	5	10	15	20
Angle	10	30	50	70
Time (s/m)	0	67	51	40
Error Slope	0	20	10	15
Current	2.1	1.9	1.7	1.5

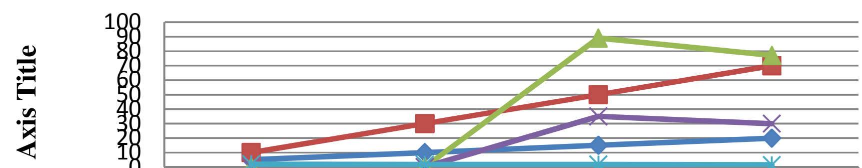
Figure 7. Comparison of parameters at flat street.

Copyright © 2011 SAVAP International

As Time Analysis (from speed variables) it can be found that the smallest time is at parameter 20/70° and the biggest Time is at parameter 5/10° (At 5/10° robot unmoved). As Slope Error Analysis, from fig. 7 above can be found that the smallest slope error is found at parameter 15/50°**,** and the biggest slope error is found at parameter 25/70°**.** As Current Consumption Analysis, the smallest Current Consumption is found at parameter 20/70°, and the biggest current consumption is found at parameter 5/10°

4.2 Robot in Bumpy street

In second test, The robot is tested to run on bumpy street (in this case is on grass). The results are displayed on Fig. 8 below.:



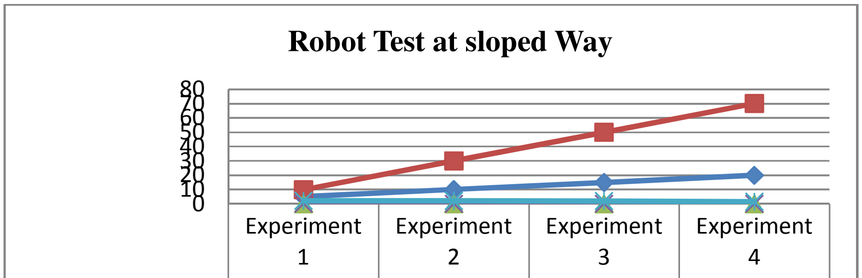
	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Delay	5	10	15	20
Angle	10	30	50	70
Time (s/m)	0	0	89	77
Error Slope	0	0	35	30
Current	2.1	1.8	1.8	1.5

Figure 8. Graphic comparison of parameters at bumpy street

As a time analysis (or Speed analysis), the smallest time is at parameter 25/70°, and the Biggest Time at parameter 15/50° (robot unmoved at Parameter 5/10° & 10/30°). From slope error analysis, the smallest slope error is found at parameter 20/70°, and the biggest slope error found at parameter 25/70°**.** For Current Consumption Analysis, the smallest current consumption is found at parameter 25/70°, and the Biggest Current Consumption found at parameter 5/10° (see Fig. 8 above).

4.3 Robot in sloped street

In the third test, the robot is tested to run on sloped street (in this case is on sloped wooden floor). As a time analysis (or speed test), the measurable time only found at parameter 25 / 90,**.** (at parameter 5/10°, 10/30°, 15/50°, 20/70°, the robot unmoved). For slope error analysis, slope error is found at parameter 25 / 70° (at parameter 5/10°, 10/30°, 15/50°, 20/70°, the robot unmoved). And for current consumption analysis, the smallest current consumption is found at parameter 15/50° & 20 / 70°, and the biggest current consumption is found at parameter 25/70°**.** The results are displayed in Fig. 9 below.



	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Delay	5	10	15	20
Angle	10	30	50	70
Time (s/m)	0	0	0	0
Error Slope	0	0	0	0
Current	2.1	2	1.9	1.5

Figure 9. Comparison of parameter at sloped street

CONCLUSION

Conclusions are found after carried out tests and analysis on reinforced learning based 5 legs robot.. At flat street, the best performance is found if robot walks at delay 20ms and angle 70^0 (not too fast). At bumpy street, the best performance is found if robot walks at delay 40ms and angle 80^0 (slow). At sloped street, the best performance is found if robot walks at delay 50ms and angle 70^0 (very slow). The performance of robot's speed is caused by delay (the bigger delay, speed decreased). The performance of robot's slope error is caused by motor speed (the faster motor, the bigger error found) and the angle degree of hip (smallest angle, the bigger error). The current consumption of robot is caused by motor's speed (the faster motor, current consumption is bigger) and the load supported by motor (the bigger load, the bigger current consumption).

REFERENCES

1. Dusko M. Katic, Aleksandar D.Rodic, abd Miomir K.Vukobratovic. *Reinforcement Learning Control Algorithm for Humanoid Robot Walking*. International Journal of Information and Systems Science Volume 4, Number 2, (2008) Pages 256-267.
2. Ernest J P Earon, Tim D Barfoot, and Gabriele M T D'Eleuterio. From the Sea to the Sidewalk: *The Evolution of Hexapod Walking Gaits by a Genetic Algorithm*. Proceedings of the Third ICES International Conference on Evolvable Systems: From Biology to Hardware (2000), LNCS 1801, pp. 51–60.
3. G. Clark Haynes and Alfred A. Rizzi. *Gait Regulation and Feedback on a Robotic Climbing Hexapod*. Proceedings of Robotics: Science and Systems, August, (2006).
4. Hajime Kimura, Toru Yamashita, Shigenobu Kobayashi. *Reinforcement Learning of Walking Behavior for a Four-Legged Robot*. Proceedings of the 40th IEEE, Conference on Decision and Control Orlando, Florida USA, December (2001).
5. Hataitep Wongsuwarn, and Djitt Laowattana. *Neuro-Fuzzy Algorithm for a Biped Robotic System* International Journal of Applied Mathematics and Computer Sciences Volume 3 Number 4, (2006) pp.195-201.

6. Inagaki, K. *Gait study for hexapod walking with disabled leg. Intelligent Robot and Systems*, 1997. IROS '97. Proceedings of the 1997 IEEE/RSJ International Conference on Volume 1, 7-11 Sept (1997). pp: 408-413.

7. Inagaki, K.; Kobayashi, H. *Adaptive wave gait for hexapod synchronized walking*. Robotics and Automation. Proceedings IEEE International Conference on 8-13 May 1994 vol.2 (1994). pp: 1326 – 1331.

8. Jun Morimoto, Gordon Cheng, Christopher G. Atkeson, and Garth Zeglin. *A Simple Reinforcement Learning Algorithm For Biped Walking.*Proceedings of the 2004 IEEE, International Conference on Robotics & Automation, New Orleans, LA, April (2004) pp.3030-3035.

9. KangKang Yin, Kevin Loken, Michiel van de Panne, *SIMBICON: Simple Biped Locomotion Control,* Proceedings of the 2007 SIGGRAPH conference Volume 26 , Issue 3,July (2007).

10. M.A. Lewis, G.A. Bekey. *Gait Adaptation in a Quadruped Robot*. Journal Autonomous Robot, Volume 12, Number 3, May (2002), pp. 301-312.

11. Masaki Ogino; Yutaka Katoh; Masahiro Aono; Minoru Asada; Koh Hosoda. *Reinforcement Learning of Humanoid Rhythmic Walking Parameters based on Visual Information.* Journal of Advanced Robotics, Volume 18, Number 7, (2004) , pp. 677-697.

12. Mustafa Suphi Erden, Kemal Leblebicioglu. *Free gait generation with reinforcement learning for a six-legged robot*. Robotics and Autonomous Systems 56, (2008). pp.199–212.

13. Nate Kohl and Peter Stone. *Policy Gradient Reinforcement Learning for Fast. Quadruped Locomotion*. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), New Orleans, LA, May (2004). pp. 2619-2624

14. Richard S. Sutton, Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, ISBN 0262193981, (1998).

15. Sonia Chernova, Manuela Veloso. *An Evolutionary Approach To Gait Learning For Four-Legged Robot.* In Proceedings of IROS'04 (International Conference on Intelligent Robot and Systems 2004), Sendai, Japan, September (2004).

16. T.D. Barfoot, E.J.P. Earon, G.M.T. D'Eleuterio. *Experiments in learning distributed control for a hexapod robo*t. Journal of Robotics and Autonomous Systems 54, (2006). pp. 864-872.

17. Tak Fai Yik and Claude Sammut. *Trial-and-Error Learning of a Biped Gait Constrained by Qualitative Reasoning*. In M. Srinivasan & M. Dunbabin (Eds.), Australasian Conference on Robotics and Automation, Brisbane. (2007).

18. Teruya Makoto, Yamada Koji, Endo Satoshi. *Studies on Forward Motion of Five Legs Robot (Nippon Kikai Gakkai Robotikusu, Mekatoronikusu Koenkai Koen Ronbunshu)* . Journal Code: L0318B VOL.2005; NO.; PAGE. 2P1-S-065 (2005).

19. W. Salatian, Keon Young Yi, Yuan F. Zheng. *Reinforcement Learning for a Biped Robot to Climb Sloping Surfaces*, Journal of Robotic Systems 14(4), 283–296 (1997).

20. Youcef Zennir, Pierre Couturier. *Multiactor approach and hexapod robot learning*. Proceedings 2005 IEEE International Symposium on Computational Intelligence in Journal of Robotics and Automation June 27-30, (2005), Espoo, Finland. pp.665-671.

21. Zennir Y., Couturier P., Bétemps M. *Distributed Reinforcement Learning of Six-Legged Robot to Walk.* The Fourth International Conference on Control and Automation (ICCA'03), 10-12 June, (2003), Montreal, Canada

22. Zhiying Wang, Xilun Ding, Alberto Rovetta. *Structure Design and Locomotion Analysis of a Novel Robot for Lunar Exploration*. 12th IFToMM World Congress, Besançon (France), June18-21 (2007).

Copyright © 2011 SAVAP International