CREATIVE CONTROL FOR INTELLIGENT AUTONOMOUS MOBILE ROBOTS

XIAOQUN LIAO **MASOUD GHAFFARI**
SOUMA M. ALHAJ ALI **ERNEST L. HALL**
Center for Robotics Research, University of Cincinnati, Cincinnati, Ohio

ABSTRACT
For intelligent robots to accomplish tasks in an unstructured environment, the adaptive critic algorithm has been shown to provide useful approximations or even optimal control policies to non-linear systems. The purpose of this paper is to explore the use of new learning control methods defined as Creative Learning or Creative Control that goes beyond the adaptive critic method for unstructured environments. The creative controller like the adaptive critic controller has information stored in a dynamic database (DB), plus a dynamic task control center (TCC) that functions as a command center to decompose tasks into sub-tasks with different dynamic models and multi-criteria functions. The simulation results based on adaptive critic learning are discussed in this paper. The significance of this paper is to better understand the adaptive critic learning theory and move forward to develop more human-intelligence-like components into the intelligent robot controller. Moreover, it should extend to other applications.

1. INTRODUCTION

Hall (1985) defines an intelligent robot as one that responds to changes to its environment through sensors connected to a controller. Dynamic Programming (DP) is perhaps the most general approach for solving optimal control problems. Adaptive Critics Design (ACD) offers a unified approach to deal with the controller's nonlinearity, robustness, and reconfiguration for a system whose dynamics can be modeled by a general ordinary differential equation. ANN and Backpropagation (BP) made it possible for ACD implementation (Werbos, 1994). However, in order to develop "brain-like intelligent control" (Werbos, 1999), it is not enough to just have the adaptive critic portion. Here we proposed a novel algorithm, called Creative Learning (CL) to fill this gap as discussed in Section 4.

2. NEUROCONTROLLER AND INTELLIGENT CONTROL THEORY

In order to design intelligent robot controllers, one must also provide the robot with a means of responding to problems in both temporal and spatial context. It is the goal of the robot researcher to design a neural learning controller to utilize the available data from the repetition in robot operations. The neural learning controller in Figure 1, based on the recurrent network architecture, has the time-variant feature that once a trajectory is learned, it should learn a second one in a shorter time.

ANN can be used to obtain the system model identification that can be used to design the appropriate intelligent robot controller. Once the real system model is available, they can also be used directly in design of the controller

(Narendra, etc., 1990). Neural network approaches to robot control can be classified as: supervised control, direct inverse control, neural adaptive control, backpropagation of utility involves information flowing backward through time, and adaptive critic method that uses a critic evaluating robot performance during training.

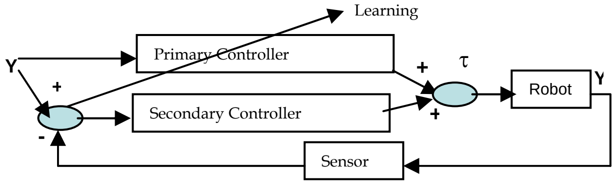

Figure 1 Recurrent neural learning controller

3. ADAPTIVE CRITIC LEARNING

Werbos (1995) summarized recent accomplishments in neurocontrol as a "brain-like" intelligent system. It should contain at least three major general-purpose adaptive components: (1) an Action or Motor system, (2) an "Emotional" or "Evaluation" system or "Critic" and (3) an "Expectations" or "System Identification" component.

"Critic" serves as a model of the external environment to be controlled, solving an optimal control problem over time may be classified as adaptive critic designs (ACD). ACD is a large family of designs which learn to perform utility maximization over time. In dynamic programming, normally the user provides the function U($\underline{X}$(t), $\underline{u}$(t)) , an interest rate r**,** and a stochastic model. Then the analyst tries to solve for another function J($\underline{X}$(t)), so as to satisfy some form of the Bellman equation shown in Equation (1) that underlies dynamic programming (Werbos, 1999):

$$J(\underline{X}(t)) = \max_{\underline{u}(t)}(U(\underline{X}(t),\underline{u}(t)) + < J(\underline{X}(t+1)) > /(1+r)) \qquad (1)$$

Where "<>" denotes expected value.

In principle, any problem in decision or control theory can be classified as an optimization problem. Many ACDs solve the problem by approximating the function J. The most popular methods to estimate J in ACDs are heuristic dynamic programming (HDP), Dual Heuristic Programming (DHP) and Globalized DHP (GDHP) (Werbos, 1995) (Prokhorov, 1997). HDP and its ACD form have a critic network that estimates the function J (cost-to-go or strategic utility function) in the Bellman equation of dynamic programming, presented as follows:

$$J(t) = \sum_{k=0}^{\infty} \gamma^k U(t+k) \qquad (2)$$

Where γ is a discount factor (0<γ<1), and U(.) is the utility function or local cost. An alternative approach referred to as Dual Heuristic Programming (DHP) has been proposed. Here, the critic network approximates the derivatives of the future cost with respect to the state variable. It has been proved that DHP is capable of generating smoother derivatives and has shown improved performance when compared to HDP (Lendaris, 1999) (Venayagamoorthy, 2002). Werbos (1992) first proposed the idea how to do GDHP. Training the critic network in GDHP utilizes an error measure which is a combination of the error measures of HDP and DHP.

4. Creative Learning and Control

"Creative Learning" provides a model, for understanding the kind of intelligence that exists in biological brains. A creative control architecture as shown in Figure 2 is proposed in this paper accoding to the creative learning theory (Liao, 2003). In this proposed diagram, there are three important components: task control center, criteria (critic) knowledge database, and learning system. Adaptive critic learning method is a part of the creative learning algorithm. However, creative learning with decision-making capabilities is beyond the adaptive critic learning. The most important characteristics of the creative learning structure are: (1) Brain-like decision-making task control center, entails the capability of human brain decision-making; (2) Dynamic criteria database integrated into the critic-action framework, makes the adaptive critic controller reconfigurable and enables the flexibility of the network framework; (3) Multiple criteria, multi-layered structure; (4) Modeled and forecasted critic modules result in faster training network. A creative controller is designed to integrate the domain knowledge and task control center into the adaptive critic controller. It needs to be a well-defined structure such as in the autonomous mobile robot application as the test-bed for the creative controller.

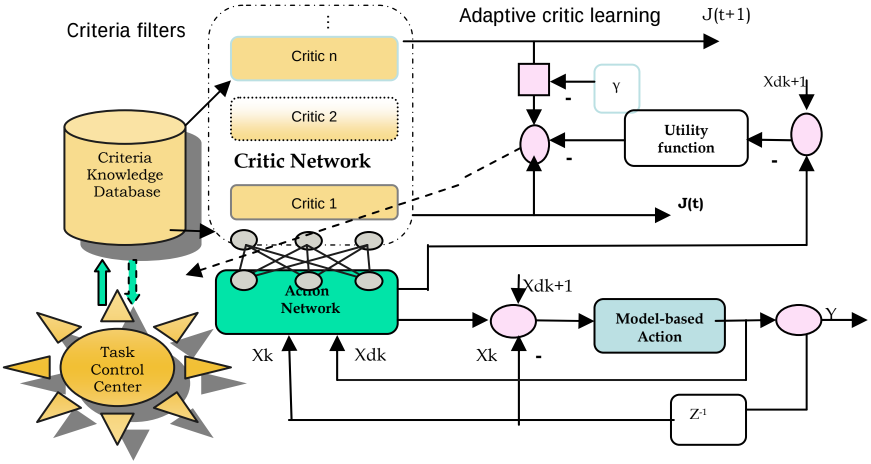

Figure 2 Proposed CL Algorithm Architecture

4.1 Task Control Center (TCC)

The task control center (TCC) can build task-level control systems for the creative learning system as shown in Figure 3. By "task-level", we mean the integration and coordination of perception, planning and real-time control to achieve a given set of goals (tasks). TCC provides a general task control framework, and it is intended to be used to control a wide variety of tasks and permits responsive actions based on mission commands, on interactions with other robots. Although TCC has no built-in control functions for particular tasks (such as robot path planning algorithms), it provides control functions, such as task decomposition, monitoring, and resource management, that are common to many applications. The particular task built-in rules in the dynamic database matches the constraints on particular control schemes or sub-tasks or environment allocated by TCC. The task control center acts as a decision-making system. It integrates domain knowledge or criteria into the database of the adaptive learning system. According to Simmons (1999), task control architecture for mobile robots provides a variety of control constructs that are commonly needed in mobile robot applications, and other autonomous mobile systems. Integrating TCC with adaptive critic learning system and interacting with the dynamic database, the creative learning system could provide both task-level and real-time control or learning within a single architectural framework.

4.2 Dynamic Knowledge Database (DKD)

Dynamic knowledge database serves as a “coupler” between a task control center and a nonlinear learning system. The purpose of the coupler is to provide the criteria functions for the adaptive critic learning system and filter the task strategies commanded by the task control center. The proposed dynamic database contains a copy of the model (or identification). Action and critic networks are utilized to control the plant under nominal operation, as well as make copies of a set of HDP or DHP parameters (or scenarios) previously adapted to deal with a plant in a known dynamic environment. It also stores copies of all the partial derivatives required when updating the neural networks using backpropagation through time (Yen, etc., 2002).

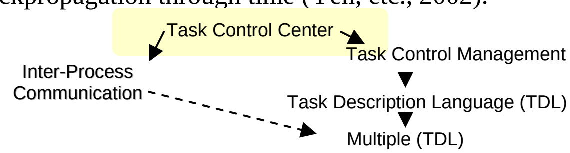

Figure 3 The structure of task control center

The data stored in the dynamic database can be uploaded to support offline or online training of the dynamic plant and provide a model for identification of nonlinear dynamic environment with its modeling function. Another function module of the database management is designed to analyze the data stored in the database including the sub-task optima, pre-existing models of the network and newly added models. The task program module is used to communicate with the task control center. The functional structure of the proposed dynamic database is shown in Figure 4 and 5.

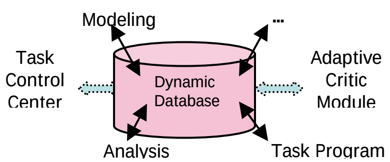

Figure 4 Functional structure of dynamic database

Database fields	
MODEL_ID	Action model ID
CRITERIA_FUN	Criteria function
Adaptive Critic Training Parameters	
INPUT_CRITIC	Input to critic network
DELT_J	J(t+1)-J(t)
…	…

Figure 5 Semantic database structure

4.3 Creative Control Mobile Robot Example

Suppose a mobile robot is used for urban rescue as shown in Figure 6. Since it is in an urban environment, it must use the established roadways. Along the road ways it can follow pathways. However, at intersections, it can choose various paths to go to the next block. Therefore, it must use different criteria at the corners. The overall goal is to arrive at the rescue site with minimum distance or time in order to rescue people. This example requires the use of both continuous and discrete tracking, a database of known information and multiple criteria optimization. It is necessary to add a large number of real-world issues including position estimation, perception, obstacles avoidance, etc.

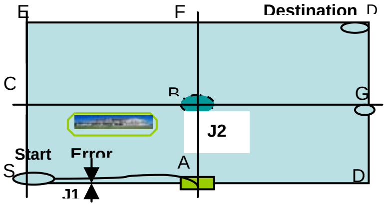

Figure 6 Simple urban rescue sites

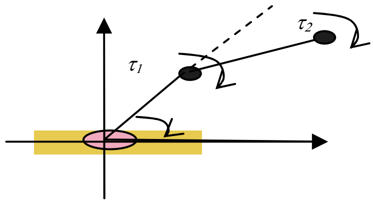

Figure 7 Two-link robot arm manipulator

As discussed in Section 4.1, the task control center (TCC) acts a decision-making command center. It takes perception information from sensors and other inputs to the creative controller and derives the criteria functions. It's appropriate to assume that J1 and J2 are the criteria functions that the task control center will transfer to the learning system at the beginning of the mission from the Start point (S) to Destination (D). The simplest style HDP in ACDs can be used. More discussion is presented in Liao's dissertation (Liao, 2003).

4.4 Creative Controller and Simulation Results

As an ANN robot controller, the block diagram of the creative controller can be presented as shown in Figure 8. A creative controller is integrated into the intelligent robot learning system. As an expermental study, we used a two-link robot arm maniputlator shown in Figure 7 as one simulation example for adaptive critic learning. The dynamics of the two-link robot arm manipulater is shown as Equation 3. The simulation results are shown in Figure 9. When training with the adaptive crtic system, it is more stable and the tracking errors are smaller. The furthur results present a comparison with the adaptive critic learning techiques and more case studies for the system shown in Eq. 3 (Liao, 2003).

$$M(q)\ddot{q} + V(q,\dot{q}) + F(q,\dot{q}) + G(q) + \tau_d = \tau \qquad (3)$$

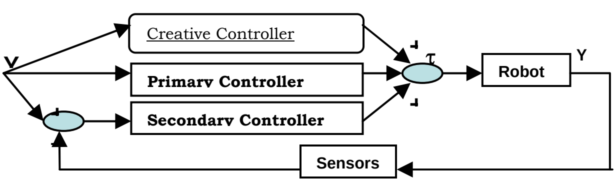

Figure 8 Creative controller structure

5. CONCLUSION

A creative learning theory proposed in this paper includes all the components in the adaptive critic family, which generalizes adaptive critic learning by modifying the learning rates and utilizing multiple critics and providing a model-based action database. A creative controller can be used to explore an unpredictable environment, and permit the discovery of unknown problems. By learning the domain knowledge, the system should be able to obtain the global optima and escape local optima. The creative controller for intelligent robots like the adaptive critic controller has information stored in a dynamic database, plus a dynamic task control center that functions as a command center to decompose tasks into sub-tasks with different dynamic models and criteria functions. Improvements such as those shown in Figure 9 have been achieved.

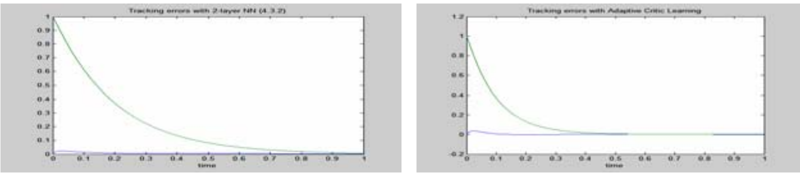

Figure 9 (a) Tracking error with NN (b)Tracking error with AC (t=1sec)

REFERENCES

Hall, E.L., and Hall, B. C. (1985), *Robotics: A User-Friendly Introduction*, pp. 1-8, Saunders College Publishing, Holt, Rienhart and Wilson, Orlando FL.

Liao, X., Creative Learning for Intelligent Autonomous Mobile Robots, PhD Dissertation, University of Cincinnati, 2003.

Lendaris, G.G., Shannon, T.T. and Rustan, A. (1999), A Comparison of Training Algorithms for DHP Adaptive Critic Neurocontrol, Neural Networks, 1999. IJCNN '99. International Joint Conference on, pp: 2265 -2270, v. 4.

Narendra, K.S., and Parthasarathy, K., (1990) Identification and control of dynamical systems using Neural networks, *IEEE Transactions on Neural Networks*, vol. **1(1),** pp. 4-27.

Prokhorov, D., and Wunsch, D., (1997) Adaptive Critic Designs, *Neural Networks, IEEE Transactions on*, v. 8, n.5, pp.997-1007.

Simmons, R., Task Control Architecture, <http://www.cs.cmu.edu/afs/cs/project/TCA/www/TCA-history.html

Venayagamoorthy, G. K., Harley, R. G. and Wunsch, D. C., (2002) Comparison of Heuristic Dynamic Programming and Dual Heuristic Programming Adaptive Critics for Neurocontrol of a Turbogenerator, *IEEE Transactions on Neural Networks*, v. 13, no. 3, pp764-773.

Werbos, P. (1994), *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting*, Wiley.

Werbos, P.J. (1999), Tutorial on Neurocontrol, Control Theory and Related Techniques: From Backpropagation to Brain-Like Intelligent Systems, the Twelth International Conference on Mathematical and Computer Modelling and Scientific Computing (12th ICMCM & SC).

Werbos, P., (1995) Optimal neurocontrol: practical benefits, new results and biological evidence, Wescon Conference Record, p.580-585.

Werbos, P.J. (1992), Approximate Dynamic Programming for Real-Time Control and Neural Modeling, In White D, A and D.A. Sofge, editors, *Handbook of Intelligent Control*, pages 493--525. Van Nostrand Reinhold.

Yen, G.G., and Lima, P. G., (2002) Dynamic Database Approach for Fault Tolerant Control Using Dual Heuristic Programming, Proceedings of the American Control Conference, Anchorage, pp.5080-5085.