Available online at www.sciencedirect.com
SCIENCE DIRECT®
Neural Networks 19 (2006) 323–337
Neural Networks
www.elsevier.com/locate/neunet
2006 Special issue

Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model

Masato Ito [a,*], Kuniaki Noda [a], Yukiko Hoshino [a], Jun Tani [b,1]

[a] *Sony Intelligence Dynamics Laboratories, Inc., Takanawa Muse Building 4F, 3-14-13 Higashigotanda, Shinagawa-ku, Tokyo 141-0022, Japan*
[b] *Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan*

Abstract

This study presents experiments on the learning of object handling behaviors by a small humanoid robot using a dynamic neural network model, the recurrent neural network with parametric bias (RNNPB). The first experiment showed that after the robot learned different types of ball handling behaviors using human direct teaching, the robot was able to generate adequate ball handling motor sequences situated to the relative position between the robot's hands and the ball. The same scheme was applied to a block handling learning task where it was shown that the robot can switch among learned different block handling sequences, situated to the ways of interaction by human supporters. Our analysis showed that entrainment of the internal memory structures of the RNNPB through the interactions of the objects and the human supporters are the essential mechanisms for those observed situated behaviors of the robot

Keywords: Learning of object handling behavior; Dynamical systems approach; Recurrent neural network

1. Introduction

Learning object handling behavior by robots is a difficult problem since motor trajectories to achieve adequate handling behaviors could be diverse regarding various situations. Even when manipulating the same object, the motor time-development would be quite different depending on how the robot and the object are situated in the workspace. The current paper shows that a dynamic neural network model is effective in learning and generating such diverse and situational behaviors for object handling.

There are a substantial number of prior studies concerning the learning of object handling by robots. Recently, Bianco and Nolfi (2004) showed that a simulated robot arm can acquire object grasping behavior by evolving neural controllers. By evolving simple sensory-motor maps in layered networks, quite complex grasping behavior is generated dynamically even with a significant range of perturbations in position and direction of the object. However, it might be difficult to apply their evolutionary approach to a real robot task because it requires a substantial number of trials, which real robot situations cannot easily accommodate.

In some studies of reinforcement learning, behavior schemes are learned by combining predefined behavior primitives. For instance, for an object handling task, a robot learns to select among the predefined behavior primitives such as approaching, grabbing, carrying and releasing an object for each step appropriately. However, this approach can hardly be applied to a dynamic object handling behavior such as object grasping (Bianco & Nolfi, 2004) and juggling (Schaal, Sternad, & Atkeson, 1996) because it is difficult to divide the dynamic behavior scheme into a set of discrete behavior primitives manually. On the other hand, some researchers (Tani & Nolfi, 1999; Wolpert & Kawato, 1998) proposed models that can learn various behavioral skills from continuous sensory-motor flow without possessing any predefined behavior primitives. Recently, some of the authors proposed a neural network scheme, termed RNN with Parametric Bias (RNNPB) (Ito & Tani, 2004a; Tani, 2003) and applied it to the task of object manipulation by an arm-type robot (Tani, 2003). However, the task was quite simple since the object was manipulated only in a 2D workspace and the interaction dynamics between the arm and the object were quite limited.

In the current study, complex tasks of a ball and blocks manipulations utilizing a humanoid robot are considered. In order to let the robot acquire these task skills, an imitation

* Corresponding author. Tel.: +81 3 5448 4777; fax: +81 3 5448 7378.
E-mail addresses: masato@idl.sony.co.jp (M. Ito), noda@idl.sony.co.jp (K. Noda), yukiko@idl.sony.co.jp (Y. Hoshino), tani@brain.riken.go.jp (J. Tani).
[1] Tel.: +81 48 467 6467; fax: +81 48 467 7248.

0893-6080/$ - see front matter
doi:10.1016/j.neunet.2006.02.007

learning framework is introduced to avoid an unrealistic number of trial and error instances, which are often observed when applying reinforcement learning and genetic algorithms to complex behavior tasks. In our imitation learning method, manipulation of objects is directly taught by human supporters who guide the movements of the robot by grasping its arms. After repeated guidance and corresponding neuronal learning, the robot becomes able to generate the taught behavioral patterns with generalization. Although it is true that the introduction of direct teaching makes the task of imitation learning much easier (Billard, 2002), it has been reported that even chimpanzees cannot learn to imitate manipulatory actions by watching but can do so by direct teaching by human supporters (Myowa-Yamakoshi & Matsuzawa, 1999, 2000).

Imitation learning by watching may require human specific cognitive functions to solve the corresponding problems (Dautenhahn & Nehaniv, 2002; Nehaniv & Dautenhahn, 2001) with joint attention mechanisms (Baron-Cohen, 1996; Moore & Corkum, 1994), which our current robots as well as chimpanzees do not have.

The current study also investigates the issues surrounding interactive and cooperative behavior generation involving robots and human supporters. Interactive generation has been addressed in the research of human–robot cooperation. In the field of conventional engineering robotics, many have studied cooperative tasking such as carrying an object (Yokoyama et al., 2003) or dancing with a human (Kosuge, Hayashi, Hirata, & Tobiyama, 2003). In those studies, robots are controlled to keep desired states within the global task models, where the human assistance is incorporated. When a human supporter pushes or pulls an object, the robot can interactively behave by keeping its state trajectories within the predesigned ones. However, in this approach, the controller of the robot has to be designed strictly as incorporated with the global task model. On the other hand, Ogata, Masago, Sugano, and Tani (2003) studied the cooperative robot–human navigation learning task without having such explicit task models. In their task, both of the human subjects and the robot learns to move to goal locations through repeated trials where the task skills of the robot are implicitly represented in the learned neural network. One of the crucial problems in interactive generation is how to coordinate the interactions between a robot's movements and the supporter's intentions of guidance. In order to accept guidance by human users, the robot's behavior generation has to be flexible enough to adapt to such external changes. On the other hand, the behavior generation has to be sufficiently robust in order to perform object handling behaviors stablely against various perturbations. Therefore, interactive generation involving human supporters requires a good balance between robustness and flexibility for adaptive behavior of the robot.

One specific goal of the current study is to show possible neuronal mechanisms that enable the robot to generate behavior adaptively corresponding to various situational changes of the robot, the object, and the human supporter. For this purpose, it is considered that reflex-type behavior generations for acquiring a simple sensory-motor mapping may not be sufficient since the recognition of situational changes in our task may require contextual information processing. In order to recognize current situations in a contextual manner, certain internal models might be required. The internal model, here, does not mean the global model of the task, but it refers to the capability to anticipate encountering sensory flow in the future by regressing sensory-motor flow of current and past time in a contextual manner. Much neuroscience research has identified that certain parts of prefrontal regions play an essential role in recognizing context switching. The Wisconsin card sorting task (WCST) (Milner, 1963) is one of the most popular schemes to investigate such mechanisms for the switching of cognitive sets. The subject is presented with cards of specific shapes, colors, and numbers. Then the subject has to sort the cards into different piles without having been explicitly given the current criteria for correct sorting. The subjects are then given feedback regarding the correctness of their current sorting results, which leads them to the correct sorting. Various neuro-imaging studies have indicated that the switching takes places with error monitoring in the anterior cingulate cortex (ACC) (Ito, Stuphorn, Brown, & Schall, 2003) and the resultant executive controls in the posterior parts of the bilateral inferior frontal sulcus (Nakahara, Hayashi, Konishi, & Miyashita, 2002). Although context switching in object manipulatory behavior and in the WCST dealing with cognition of abstract rules might be qualitatively different, they might share the same basic information flow of error-monitoring with anticipation and resultant executive control for switching.

In the current paper, our previously described scheme of the RNNPB (Ito & Tani, 2004b; Tani, 2003) is utilized as one possible neuronal network model to implement context switching. The ultimate challenge of the study is to clarify the essential mechanism of context switching for the task of object handling from the dynamical systems perspectives (Beer, 1995; Gelder, 1998). The dynamical structures that appear in the tight coupling among the body, the object and the internal neuronal processes will be explained by means of attractor dynamics and their parameter bifurcation characteristics.

2. Mechanism, model and algorithm

In order to achieve learning and the resultant interactive generation of learned behavior, a dynamic neural network model of RNNPB (Ito & Tani, 2004b; Tani, 2003) is utilized. In the following section, the basic cognitive modes of the RNNPB are introduced.

2.1. The basic mechanism

The following explains the basic idea for three different cognitive operational modes for a robot, which include learning, object handling, and object handling with human supporters. First, in the learning phase, sensory-motor patterns of guided behaviors are embedded in the RNNPB in the form of attractor dynamics. The attractor represents the essential spatio-temporal structure of the target behavior. Moreover,