Tacit Learning for Emergence of Task-Related Behaviour through Signal Accumulation

Vincent Berenz[1], Fady Alnajjar[1], Mitsuhiro Hayashibe[2], and Shingo Shimoda[1]

[1] BSI-Toyota Collaboration Center, RIKEN, Nagoya, Japan
{vincent,fady,shimoda}@brain.riken.jp
[2] INRIA DEMAR Project and LIRMM, UMR5506, CNRS University of Montpellier, France
mitsuhiro.hayashibe@inria.fr

Abstract. Control of robotic joints movements requires the generation of appropriate torque and force patterns, coordinating the kinematically and dynamically complex multijoints systems. Control theory coupled with inverse and forward internal models are commonly used to map a desired endpoint trajectory into suitable force patterns. In this paper, we propose the use of tacit learning to successfully achieve similar tasks without using any kinematic model of the robotic system to be controlled. Our objective is to design a new control strategy that can achieve levels of adaptability similar to those observed in living organisms and be plausible from a neural control viewpoint. If the neural mechanisms used for mapping goals expressed in the task-space into control-space related command without using internal models remain largely unknown, many neural systems rely on data accumulation. The presented controller does not use any internal model and incorporates knowledge expressed in the task space using only the accumulation of data. Tested on a simulated two-link robot system, the controller showed flexibility by developing and updating its parameters through learning. This controller reduces the gap between reflexive motion based on simple accumulation of data and execution of voluntarily planned actions in a simple manner that does not require complex analysis of the dynamics of the system.

1 Introduction

Living organisms are confronted with changes of the environment. They create new behavioural patterns using body environment interactions when carrying out tasks under unknown situations. In spite of advances in machine learning and adaptive methods, such as artificial neural networks [1][2][3], reinforcement learning [4][5][6], adaptive controls [7][8] and fuzzy control [9][10], we have so far been unsuccessful in creating artificial systems that have such adaptability. In a previous paper [11] we introduced tacit learning. Tacit learning is an unsupervised learning method in which various behavioral structures spontaneously emerge through body-environment interactions subject to certain innate rules. Computations progress by accumulating the local activities of elements. This accumulation creates behaviors adapted to the environment. We investigate the possible role of such accumulation in the

P. Sinčák et al. (eds.), *Emergent Trends in Robotics and Intelligent Systems*,
Advances in Intelligent Systems and Computing 316, DOI: 10.1007/978-3-319-10783-7_3

spontaneous generation of adaptive behaviors based on reflex action. The use of reflex actions is particularly appealing to the control of robotics systems; contrary to methods of traditional control theory, it does not rely on the analytical resolution of systems characterized by highly non-linear dynamics.

Previously, we used the torque as the input to be reduced, the proposed controller being used for the emergence of bipedal walking behaviors of a 36DOF humanoid robot [12]. Tacit learning allowed to reduce the complexity of designing a walking behavior by allowing most of the joints to have unspecified target angles. The gait that emerged from the learning process was highly adapted to the environment in terms of efficiency, rhythm and robustness. Our robot succeeded in learning bipedal walking in a completely model-free fashion. Balance emerged within approximately 10 minutes of tacit learning in real environments. More recently, a novel optimal control paradigm in motor learning based on tacit learning was proposed and showed that simple tacit learning can realize simultaneously environmental adaptation and optimal control [13]. In a vertical reaching task, this method systematically produced motor synergies which would induce an efficient solution in a redundant task space.

In this paper, we use tacit learning to create behaviors based on accumulation of task-space feedback information. Learning refers here to the dynamic tuning of the controller. Task-space feedback information is used in many modern robot control systems to improve robustness. While the sensory information is important to improve the endpoint accuracy in the presence of uncertainty, most sensory control schemes require the exact knowledge of the Jacobian matrix from joint-space to task-space. This approach is inconvenient if the system is too complex for finding an analytical solution online, or if some of the kinematic parameters of the system are changing or unknown, for example when the robot manipulates a tool with unknown length. Several approaches approximating Jacobian controllers for set-point control of robots with uncertainties in both kinematics and dynamics have been proposed [14][15][16].

Based on the principle of tacit learning, we propose a simple controller that avoids the usage of the Jacobian altogether in a process that does not involve any model of the system being controlled. Our objective is to design a new control strategy that can achieve levels of adaptability similar to those observed in living organisms, be plausible from a neural control view point, and does not require complex analysis of the dynamics of the system.

2 Tacit Learning Controllers

The general expression for a tacit controller is:

$$U = KX_c + Q \tag{1}$$

$$\dot{Q} = A \tag{2}$$

U is the control, Xc the state variable expressed in the control space, K is the proportional and derivative gain matrix, and A the effect to be minimized.

In this paper we extend this controller by expressing $\dot{Q}$, the tacit component of the controller, as:

$$\dot{Q} = W(p_1, p_2, \dots, p_m) \circ E \tag{3}$$

" $\circ$ " is the Hadamard product. E is a n × m matrix which rows vectors are all e, e being a vector of size m of errors expressed in the task space. The vector e has to construct such as e=0 for the desired state and n is the number of degrees of freedom of the system. W is also a n × m matrix. P is the set of parameters dynamically tuned by higher level algorithms and controllers such as obtaining linearization of the system described by equation (1). The parameters pj correspond for example to the norms of the row vectors of W or to the angles between them. The advantage of this approach is that values of pj can be directly related to features of the task-space, for example the speed of the end-point of a manipulator.

3 Two-Links Robotic System

3.1 Task and Control

This kind of controller can be applied to a wide range of robotic systems for which analytical solutions of a task expressed in the task-space is impractical. In this paper we provide control of a two-link robotic arm which kinematic configuration is unknown. We define the task as:

$$\begin{cases} H_{t\to\infty} = 0 \\ \dot{H}_{t\to\infty} = 0 \\ H_{|H|\to\delta} = \rho V \end{cases} \tag{4}$$

H is the vector defined by the end-point and the target. t is the time. δ is a small number and ρ is a real number. As shown in Fig.1, this corresponds to reaching the target at the angle specified by the vector V. Extended to the 6D space, this definition is simple but sufficient to specify a grasping task: it requires the robot to reach the target with the end-effect at a certain position and orientation, while leaving unspecified the trajectory taken by the end-effector and the configuration of the rest of the robotic system. By specifying a succession of tasks, this approach is also suitable to define the complete trajectory. For reasons explained later on, we refer to the line defined by the target and the vector V as the stability line.

We apply the equations 1-3 to this robotic system. We define e as:

$$e = \left[V \cdot H, rot_{\pi/2} V \cdot H\right] \tag{5}$$

H is the vector defined by the end-point and the target, '·' is the dot product and $rot_{\pi/2}$ the rotation matrix of the angle $\pi/2$. W is a 2×2 matrix and we will refer to its row vectors as w1 and w2. We define $p = [\phi, \alpha, a]$: ϕ is the angle between V and w1

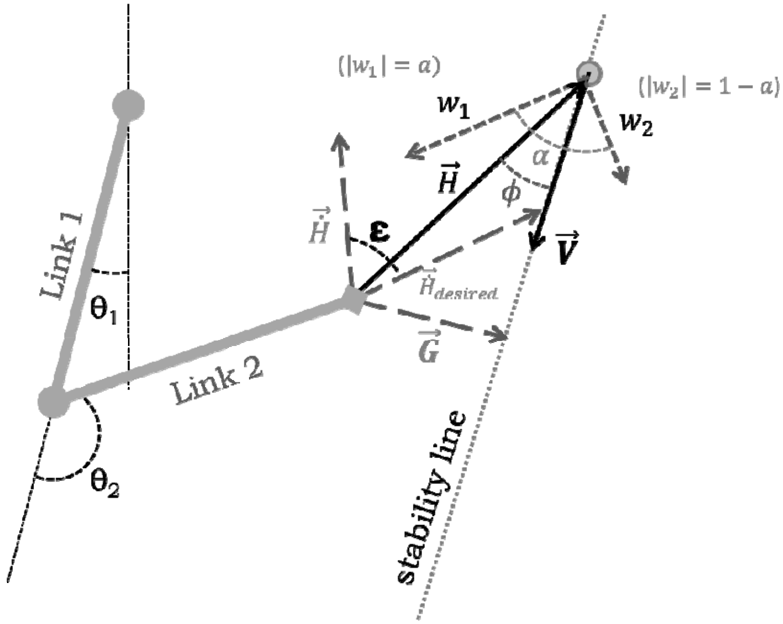

Fig. 1. The end-effector must reach the target following the line trajectory defined by the stability line. The controller achieves this by tuning φ (the orientation), α (the angles between w1 and w2), and a (the norms of w1 and w2) such as achieving ε=0 (see text for details).

+ w2, α is the angle between w1 and w2, and a is the norm of w1. W can be determined based on p:

$$w_1 = a\left(rot_{\phi+\alpha/2}V\right) \tag{6}$$

$$w_2 = (1-a)\left(rot_{\phi-\alpha/2}V\right) \tag{7}$$

For this system, equation 3 can be written:

$$\dot{Q}_c = \begin{bmatrix} a\left(rot_{\phi+\alpha/2}V\right) \\ (1-a)\left(rot_{\phi-\alpha/2}V\right) \end{bmatrix} \circ \begin{bmatrix} V \cdot H, rot_{\frac{\pi}{2}}V \cdot H \\ V \cdot H, rot_{\frac{\pi}{2}}V \cdot H \end{bmatrix} \tag{8}$$

Once the values of ϕ, α, and a are set, all the parameters of the controller expressed by equations 1 and 3 are determined and the system can be controlled. Based on equations 4 and 5, an informal explanation of the proposed controller can be given. The target, w1 and w2 define two lines in the task space. If α is zero, then these two lines coincide. If the end-point is on this line, then $\dot{Q}$ is null and the system stabilizes. Furthermore, if ϕ is also zero, these two lines will also coincide with the stability line on which the system will stabilize. If α is not zero, then $\dot{Q} = 0$ can be achieved only when the end-point reaches the target. This induced instability will result in the end-effector moving. ϕ, α, and a can be dynamically tuned so that this movement will follow the stability line toward the target.

We consider that the desired end-point velocity is $\dot{H}_{desired} = H + G$, G being the vector between the end-point and the stability line. a is modelled as a virtual dynamic parameter controlled such as reaching $\dot{H} \approx \dot{H}_{desired}$:

$$\ddot{a} = -k_p a - k_d \dot{a} + k_t \int \varepsilon dt \quad (9)$$

k_p, k_d, and k_t are proportional, derivative, and tacit gains. ε is the signed angle between $\dot{H}$ and $\dot{H}_{desired}$.

α is tuned such as decreasing the speed of the endpoint when $\dot{H}$ is not converging toward $\dot{H}_{desired}$:

$$\alpha = \alpha_0 + \frac{1}{\tau_0} \times \int_{t-\tau_0}^{t} sign(\varepsilon\dot{\varepsilon})dt \quad (10)$$

α0 and τ_0 are parameters of the system and sign is the function defined as:

$$\begin{cases} sign(x) = 0 & if\ x \leq 0 \\ sign(x) = 1 & if\ x > 0 \end{cases} \quad (11)$$

Finally ϕ, tuned as the endpoint, is always positioned between the two lines defined by the target, w1 and w2 by setting it to the angle between V and H.

3.2 Simulation Results

The controller was tested in a simulation created using Open Dynamic Engine [17] using the configuration:

$$l1 = 0.5[m]\ l2 = 0.4[m] \quad (12)$$

$$m1 = 0.5[kg]\ m2 = 0.4[kg] \quad (13)$$

None of these configuration parameters are known to the controller. The gains of the controller were set to: kp=10 kd=0.6 and kt=1.0 for the controller presented in equation 1, and kp=10 kd=2.0 and kt=20 for the controller presented in equation 8. We set α0 =π/12 and τ0=1[s].

The arms moved in the 2D sagittal plane subjected to gravity and were required to reach a target at [0.5,0.0] in the Cartesian reference coordinate centered on the base of the arm. The reaching vector V was the vector defined by the target and the initial positions of the end-point. The starting position of the robot is θ1=0 and θ2=0, corresponding to straight joints pointing downward (see Figure 2). Using the proposed controller, the endpoint reaches the target following the stability line. Figure 2 shows in blue circles the trajectory of the end-point. The configuration of the robotic system is also shown for the starting position θ1=0 and θ2=0), the final position when the end-point reaches the target and for an intermediate time (in lighter grey). Figure 3 shows the evolution in time of some of the parameters of the system

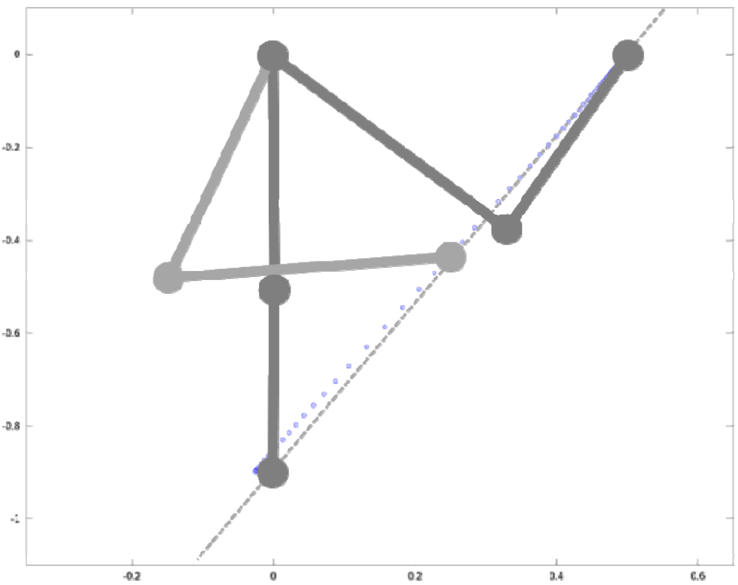

Fig. 2. The trajectory of the end-point (in blue). Postures of the system are shown at the beginning of the experiment, at the end of the experiment when the end-point reaches the target, and at an intermediary time. The stability line is represented by the black dashed line.

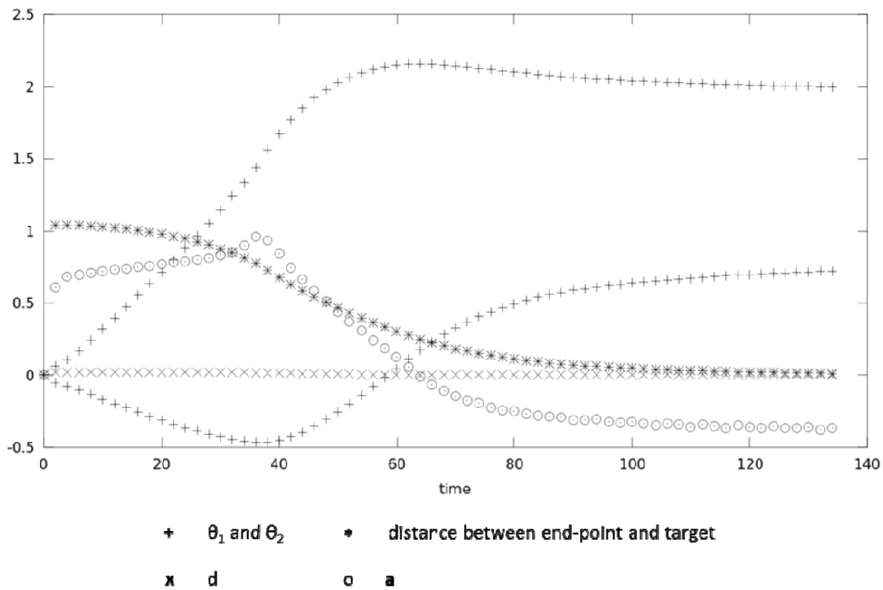

Fig. 3. Evolution of time in some parameters of the system. a is dynamically tuned to achieve a linear motion of the end-point, compensating the high nonlinearity displayed by the values of θ1 and θ2. It can be seen that d, the distance between the positions of the end-point and the desired line trajectory, remains very close to zero.

with time. The distance between the end-point and the target decreases while the distance between the end-point and the stability remains very close to zero. The parameter a is dynamically tuned such as finding values of θ1 and θ2 corresponding to the desired line trajectory of the end-point. The controller manages to compensate for the nonlinearities of the system, despite not using any kinematic parameter.

4 Discussion and Future Work

Our goal is to provide control solutions to robotic tasks for which 1) analytical solutions have high complexity, for example grasping an object while walking, or 2) the environment cannot be modelled, which is an issue for the development of rehabilitation of the robotic system that adapts to each patient. To reach this goal, we are now working towards several directions:

4.1 Stability

The stability of the proposed controller depends highly on the initial configuration of the system as well as on the gain values chosen. Stability analysis is required to design a systematic method for selecting the value of these gain values in a fashion that guarantees stability. First results in regards of the stability analysis of tacit controllers have already been published [18] and further characterization is on-going. Extension to higher dimensionality

We are currently working on extending the methodology presented in this paper to the six dimensional space. While in 2D the task is represented by a line expressed in the task space, in a higher dimension the task will be represented by a combination of lines and planes and virtual dynamic variables also being tuned by the tacit learning controller to control the transformation of these geometric entities.

4.2 Motor Control

The neural mechanisms used for mapping goals expressed in the task-space into control-space related commands without using internal models remain largely unknown. But many neural systems rely on data accumulation: the presented controller incorporates knowledge expressed in the task-space using only the accumulation of data and is plausible from a neural control viewpoint.

5 Conclusion

In this work, the system achieved to control the trajectory of the end-point such as following a line trajectory. The controller manages to compensate for the nonlinearities of the system, despite not using any kinematic parameter. The presented controller does not use any internal model and incorporates knowledge expressed in the task-space using only the accumulation of data. By providing control solutions that are not based on the Jacobian matrix for solving the relationships between task and control space, we target to reach levels of adaptability similar to the one observed in living organisms. Such an adaptability level will provide advantages in the control of complex over an actuated robotic system or robotic system for which the environment cannot be modelled.

References

1. Minsky, M.L., Papert, S.A.: Perceptron. MIT Press, Cambridge (1969)
2. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by backpropagating errors. Nature 323(6088), 533–536 (1986)
3. Kuniyoshi, Y., Yorozu, Y., Suzuki, S., Sangawa, S., Ohmura, Y., Terada, K., Nagakubo, A.: Emergence and development of embodied cognition: A constructivist approach using robots. Prog. Brain Res. 164, 425–445 (2007)
4. Barto, A.G., Sutton, R.S., Anderson, C.W.: Neuron-like adaptive elements that can solve difficult learning control problems. IEEE Trans. Syst., Man, Cybern. SMC-13(5), 834–846 (1983)
5. Doya, K.: Reinforcement learning in continuous time and space. Neural Comput. 12(1), 219–245 (2000)
6. Tedrake, R., Zhang, T.W., Seung, H.S.: Stochastic policy gradient reinforcement learning on a simple 3D biped. In: Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., pp. 2849–2854 (2004)
7. Astrom, K.J., Wittenmark, B.: Adaptive Control. Addison-Wesley, Reading (1989)
8. Slotin, J.E., Li, W.: Applied Nonlinear Control. Prentice-Hall, Englewood Cliffs (1991)
9. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. Syst. Man, Cybern. SMC-3(1), 28–44 (1973)
10. Juang, J.G.: Fuzzy neural network control CMAC of a biped walking robot. IEEE Trans. Syst., Man, Cybern. B, Cybern. 30(4), 594–601 (2000)
11. Shimoda, S., Kimura, H.: Bio-mimetic Approach to Tacit Learning based on Compound Control. IEEE Transactions on Systems, Man, and Cybernetics- Part B 40(1), 77–90 (2010)
12. Shimoda, S., Kimura, H.: Adaptability of tacit learning in bipedal locomotion. IEEE Transactions on Autonomous Mental Development 5(2), 152–161 (2013)
13. Hayashibe, M., Shimoda, S.: Emergence of Motor Synergy in Vertical Reaching Task via Tacit Learning. In: International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4985–4988 (2013)
14. Cheah, C.C., Hirano, M., Kawamura, S., Arimoto, S.: Approximate Jacobian control for robots with uncertain kinematics and dynamics. IEEE Transactions on Robotics and Automation 19(4), 692–702 (2003)
15. Dixon, W.E.: Adaptive regulation of amplitude limited robot manipulators with uncertain kinematics and dynamics. IEEE Transactions on Automatic Control 52(3), 488–493 (2007)
16. Ozawa, R., Oobayashi, Y.: Adaptive task space PD control via implicit use of visual information. In: Int. Sym. Robot Control, pp. 209–214 (2009)
17. Smith, R.: Open Dynamics Engine, `http://www.ode.org/`
18. Shimoda, S., Yoshihara, Y., Fujimoto, K., Yamamoto, T., Maeda, I., Kimura, H.: Stability analysis of tacit learning based on environmental signal accumulation. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 7-12 (2012)