# Correct equations for the dynamics of the cart-pole system

Răzvan V. Florian
Center for Cognitive and Neural Studies (Coneural)
Str. Saturn 24, 400504 Cluj-Napoca, Romania
florian@coneural.org

July 11, 2005; updated February 10, 2007

**Abstract**

The problem of balancing a pole on a moving cart is a widely used benchmark problem for testing reinforcement learning algorithms. The classic papers that introduced this problem contain mistakes in the equations that govern the dynamics of the cart-pole system, and these mistakes propagated in other studies that used the same problem as a benchmark. Here we provide the equations that describe correctly the dynamics of the system.

## 1 Introduction

The control of a cart-pole system is widely used as a benchmark problem for testing the efficiency of reinforcement learning algorithms. It seems to have been first used as a test problem in adaptive control by Michie and Chambers (1968*a*,*b*) and became a more famous problem since its use in the paper of Barto et al. (1983). Google Scholar[1] reports about 500 papers citing this paper, and about 100 papers containing the words "cart pole" or "cartpole". There are, however, two mistakes in the equations from Barto et al. (1983) that describe the dynamics of the cart pole. One mistake introduces a difference between the reported equations and the equations describing a correct physical model, and the other mistake is probably a typo. The mistakes propagated in other papers that followed the original paper of Barto (e.g., Anderson, 1986; Schmidhuber, 1990; Si and Wang, 2001). The existence of these mistakes does not affect the validity of the reinforcement learning algorithms presented in the papers using them, because the problematic equations still describe a complex dynamical system. However, we believe that it is useful to be corrected, for the sake of scientific rigor.

Here we provide the equations that describe correctly, from a physical point of view, the dynamics of the system.

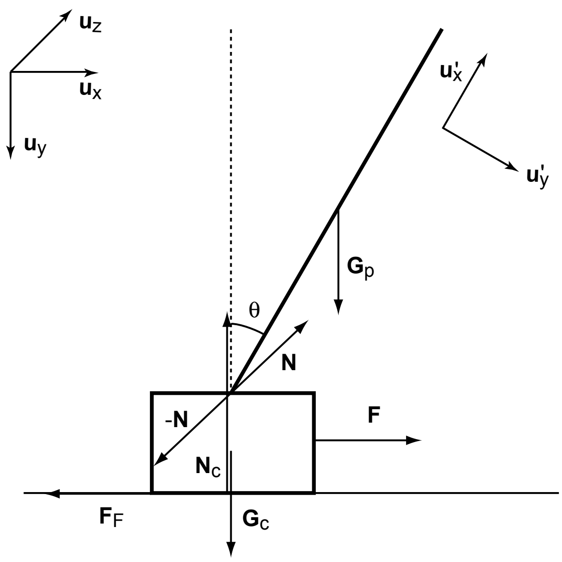


Figure 1: The cart-pole system.

## 2 The system

The studied system is a cart of which a rigid pole is hinged (see figure). The cart is free to move within the bounds of a one-dimensional track. The pole can move in the vertical plane parallel to the track. The controller can apply a force **F** to the cart, parallel to the track. The cart has a mass $m_c$; the pole has mass $m_p$ and length $2l$. We note with $x$ the position of the cart along the track. The angle between the pole and the vertical is $\theta$. The friction coefficient between the cart and the track is $\mu_c$; there also exists friction in the articulation connecting the pole to the cart that leads to a torque $\mu_p\,\dot{\theta}$.

## 3 The wrong equations and the mistakes

The equations reported in Barto et al. (1983) are:

$$\ddot{\theta} = \frac{g\sin\theta + \cos\theta\left[\dfrac{-F - m_p\, l\,\dot{\theta}^2\sin\theta + \mu_c\,\mathrm{sgn}(\dot{x})}{m_c + m_p}\right] - \dfrac{\mu_p\,\dot{\theta}}{m_p\, l}}{l\left[\dfrac{4}{3} - \dfrac{m_p\cos^2\theta}{m_c + m_p}\right]} \tag{1}$$

$$\ddot{x} = \frac{F + m_p\, l\,[\dot{\theta}^2\sin\theta - \ddot{\theta}\cos\theta] - \mu_c\,\mathrm{sgn}(\dot{x})}{m_c + m_p} \tag{2}$$

The mistake inherent in these equations is related to the friction force between the cart and the track, that is considered to be $\mu_c\,\mathrm{sgn}(\dot{x})$. The mistake is apparent even by

[1] http://scholar.google.com/

dimensional analysis, observing that $\mu_c\,\mathrm{sgn}(\dot{x})$ is adimensional instead of having the dimensionality of a force. In fact, the friction force is the product between the friction coefficient and the magnitude of the force normal to the track; and the force normal to the track is not constant, because the movement of the pole induces a variation in its magnitude.

A second mistake in Barto et al. (1983) is to consider the gravitational acceleration $g$ in these equations to be negative (in the paper of Barto is specified that $g = -9.8$ m/s$^2$). In fact, these equations are consistent with a positive value of $g$. Otherwise, the term $g\sin\theta$ would produce a negative $\ddot{\theta}$ for a small $\theta$, according to Eq. 1, meaning that the gravity pushes the pole towards the vertical position; this is, of course, wrong.

# 4 Correct equations

We consider that the cart acts with a reaction force $\mathbf{N}$ on the pole, at the articulation. According to the law of action-reaction, the pole will act on the cart with a force $-\mathbf{N}$.

By applying Newton's second law to the cart we get:

$$\mathbf{F}+\mathbf{F}_f+\mathbf{G}_c-\mathbf{N}+\mathbf{N}_c = m_c\,\mathbf{a}_c, \tag{3}$$

where $\mathbf{F_f}$ is the friction force between the cart and the track that acts on the cart, and $\mathbf{a}_c$ is the acceleration of the cart. We have $\mathbf{F} = F\,\mathbf{u}_x$; $\mathbf{F}_f = -F_f\,\mathbf{u}_x$; $\mathbf{G}_c = m_c\,g\,\mathbf{u}_y$; $\mathbf{N} = N_x\mathbf{u}_x - N_y\mathbf{u}_y$; $\mathbf{N}_c = -N_c\,\mathbf{u}_y$; $\mathbf{a}_c = \ddot{x}\,\mathbf{u}_x$; $\mathbf{u}_x$, $\mathbf{u}_y$ and $\mathbf{u}_z$ are the unit vectors of the laboratory frame of reference (see figure). Decomposing the previous equation on the $x$ and $y$ axis we get

$$F - F_f - N_x = m_c\,\ddot{x} \tag{4}$$

$$m_c\,g + N_y - N_c = 0. \tag{5}$$

According to the Coulomb model of friction, and assuming that the track limits the movement of the cart both downwards and upwards, the friction force is

$$F_f = \mu_c\,|N_c|\,\mathrm{sgn}(\dot{x}) = \mu_c\,N_c\,\mathrm{sgn}(N_c\,\dot{x}). \tag{6}$$

By applying Newton's second law to the linear movement of the pole we get:

$$\mathbf{N}+\mathbf{G}_p = m_p\,\mathbf{a}_p, \tag{7}$$

where $\mathbf{G}_p = m_p\,g\,\mathbf{u}_y$. The acceleration $\mathbf{a}_p$ of the center of mass of the pole is due to the composed effects of the acceleration of the cart it is attached to, and of the rotation of the pole with angular velocity $\boldsymbol{\omega} = \dot{\theta}\,\mathbf{u}_z$ and angular acceleration $\boldsymbol{\varepsilon} = \ddot{\theta}\,\mathbf{u}_z$:

$$\mathbf{a}_p = \mathbf{a}_c + \boldsymbol{\varepsilon}\times\mathbf{r}_p + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}_p), \tag{8}$$

where $\mathbf{r}_p = l\,(\sin\theta\,\mathbf{u}_x - \cos\theta\,\mathbf{u}_y)$ is the vector representing the position of the center of mass of the pole relative to the articulation around which the pole rotates. Thus, we get

$$\mathbf{a}_p = \ddot{x}\,\mathbf{u}_x + l\,\ddot{\theta}\,\mathbf{u}_z\times(\sin\theta\,\mathbf{u}_x - \cos\theta\,\mathbf{u}_y) + l\,\dot{\theta}^2\,\mathbf{u}_z\times[\mathbf{u}_z\times(\sin\theta\,\mathbf{u}_x - \cos\theta\,\mathbf{u}_y)]. \tag{9}$$

We have $\mathbf{u}_z\times\mathbf{u}_x = \mathbf{u}_y$ and $\mathbf{u}_z\times\mathbf{u}_y = -\mathbf{u}_x$. Hence,

$$\mathbf{a}_p = \ddot{x}\,\mathbf{u}_x + l\,\ddot{\theta}\,(\sin\theta\,\mathbf{u}_y + \cos\theta\,\mathbf{u}_x) - l\,\dot{\theta}^2(\sin\theta\,\mathbf{u}_x - \cos\theta\,\mathbf{u}_y). \tag{10}$$

An alternative way to get to the previous equation is to compute directly the acceleration component due to the angular velocity, $-l\,\dot{\theta}^2\,\mathbf{u}'_x$, and the acceleration component due to the angular acceleration, $l\,\ddot{\theta}\,\mathbf{u}'_y$, and expressing the unit vectors $\mathbf{u}'_x$ and $\mathbf{u}'_y$ of the frame of reference rotating with the pole in the laboratory frame of reference. By introducing Eq. 10 into 7 and decomposing on the $x$ and $y$ axis we get

$$N_x = m_p\,(\ddot{x} + l\,\ddot{\theta}\cos\theta - l\,\dot{\theta}^2\sin\theta) \tag{11}$$

$$m_p\,g - N_y = m_p\,(l\,\ddot{\theta}\sin\theta + l\,\dot{\theta}^2\cos\theta) \tag{12}$$

By applying Newton's second law to the rotational movement of the pole around the articulation (that moves with acceleration $\mathbf{a}_c$ relative to the laboratory frame of reference) we get:

$$\mathbf{M} = I\,\boldsymbol{\varepsilon} + \mathbf{r}_p \times \mathbf{a}_c, \tag{13}$$

where $\mathbf{M} = \mathbf{r}_p \times \mathbf{G}_p - \mu_p\,\dot{\theta}\,\mathbf{u}_z$ is the sum of the non-inertial torques acting on the pole relative to the articulation, $I = 4/3\,m_p\,l^2$ is the moment of inertia of the pole relative to the articulation, and $-\mathbf{r}_p \times \mathbf{a}_c$ can be interpreted as the torque generated by the inertial force caused by the acceleration of the cart. Hence, we get

$$m_p\,g\,l\sin\theta - \mu_p\,\dot{\theta} = 4/3\,m_p\,l^2\,\ddot{\theta} + m_p\,\ddot{x}\,l\cos\theta. \tag{14}$$

From Eq. 4 and 11 we get:

$$\ddot{x} = \frac{F + m_p\,l\,(\dot{\theta}^2\sin\theta - \ddot{\theta}\cos\theta) - F_f}{m_c + m_p}, \tag{15}$$

and by introducing this into Eq. 14 we get

$$\ddot{\theta} = \frac{g\sin\theta + \cos\theta\left[\dfrac{-F - m_p\,l\,\dot{\theta}^2\sin\theta + F_f}{m_c + m_p}\right] - \dfrac{\mu_p\,\dot{\theta}}{m_p\,l}}{l\left[\dfrac{4}{3} - \dfrac{m_p\cos^2\theta}{m_c + m_p}\right]}. \tag{16}$$

We see that the last two equations are the same as Barto's equations 1 and 2, with the difference that Barto used a form of the friction force $F_f$ that is wrong. Indeed, from Eq. 6, 5 and 12 we get

$$N_c = (m_c + m_p)\,g - m_p\,l\,(\ddot{\theta}\sin\theta + \dot{\theta}^2\cos\theta) \tag{17}$$

$$F_f = \mu_c\,[(m_c + m_p)\,g - m_p\,l\,(\ddot{\theta}\sin\theta + \dot{\theta}^2\cos\theta)]\,\mathrm{sgn}(N_c\,\dot{x}). \tag{18}$$

By introducing the previous equation in Eq. 15 and then in 16, we get

$$\ddot{\theta} = \frac{g\sin\theta + \cos\theta\left\{\dfrac{-F - m_p\,l\,\dot{\theta}^2\,[\sin\theta + \mu_c\,\mathrm{sgn}(N_c\,\dot{x})\cos\theta]}{m_c + m_p} + \mu_c\,g\,\mathrm{sgn}(N_c\,\dot{x})\right\} - \dfrac{\mu_p\,\dot{\theta}}{m_p\,l}}{l\left\{\dfrac{4}{3} - \dfrac{m_p\cos\theta}{m_c + m_p}[\cos\theta - \mu_c\,\mathrm{sgn}(N_c\,\dot{x})]\right\}}. \tag{19}$$

## 5 Conclusion

In conclusion, we have provided dynamical equations for the cart-pole system that are correct from a physical point of view. They are:

$$N_c = (m_c + m_p)\, g - m_p\, l\, (\ddot{\theta} \sin\theta + \dot{\theta}^2 \cos\theta) \tag{20}$$

$$\ddot{\theta} = \frac{g \sin\theta + \cos\theta \left\{ \dfrac{-F - m_p\, l\, \dot{\theta}^2 \left[\sin\theta + \mu_c \operatorname{sgn}(N_c\, \dot{x}) \cos\theta\right]}{m_c + m_p} + \mu_c\, g \operatorname{sgn}(N_c\, \dot{x}) \right\} - \dfrac{\mu_p\, \dot{\theta}}{m_p\, l}}{l \left\{ \dfrac{4}{3} - \dfrac{m_p \cos\theta}{m_c + m_p} \left[\cos\theta - \mu_c \operatorname{sgn}(N_c\, \dot{x})\right] \right\}} \tag{21}$$

$$\ddot{x} = \frac{F + m_p\, l\, (\dot{\theta}^2 \sin\theta - \ddot{\theta} \cos\theta) - \mu_c\, N_c \operatorname{sgn}(N_c\, \dot{x})}{m_c + m_p}. \tag{22}$$

During a simulation of the system, at each timestep, we may assume that $N_c$ has the same sign as at the previous timestep (we may consider it to be positive at the beginning of the simulation) and compute $\ddot{\theta}$ according to Eq. 21. We then compute $N_c$ using the value of $\ddot{\theta}$ that we obtained, according to Eq. 20; if $N_c$ changes sign, we compute again $\ddot{\theta}$ taking into account the new sign. Finally, we compute $\ddot{x}$ according to 22. Usually, for common choices of the parameters, $N_c$ will be always positive, as the cart should not try to jump off the track.

If we neglect friction, the equations are

$$\ddot{\theta} = \frac{g \sin\theta + \cos\theta \left( \dfrac{-F - m_p\, l\, \dot{\theta}^2 \sin\theta}{m_c + m_p} \right)}{l \left( \dfrac{4}{3} - \dfrac{m_p \cos^2\theta}{m_c + m_p} \right)} \tag{23}$$

$$\ddot{x} = \frac{F + m_p\, l\, (\dot{\theta}^2 \sin\theta - \ddot{\theta} \cos\theta)}{m_c + m_p}. \tag{24}$$

In these equations, $g$ is positive, and not negative, as mistakenly indicated in the paper of Barto.

## 6 Acknowledgements

I thank to James Knight for spotting a typo in a previous version of this paper.

## References

Anderson, C. W. (1986), Learning and Problem Solving with Multilayer Connectionist Systems, PhD thesis, University of Massachussets, Department of Computer and Information Science.
http://www.cs.colostate.edu/~anderson/res/rl/chuck-diss.pdf

Barto, A. G., Sutton, R. S. and Anderson, C. (1983), 'Neuron-like adaptive elements that can solve difficult learning control problems', *IEEE Transactions on Systems,*

*Man, and Cybernetics* **13**, 834–846.
http://www.cs.ualberta.ca/~sutton/papers/barto-sutton-anderson-83.pdf.gz

Michie, D. and Chambers, R. A. (1968*a*), BOXES: An experiment in adaptive control, *in* E. Dale and D. Michie, eds, 'Machine Intelligence 2', Oliver and Boyd, Edinburgh, pp. 137–152.

Michie, D. and Chambers, R. A. (1968*b*), Boxes as a model of pattern formation, *in* C. H. Waddington, ed., 'Towards a theoretical biology', Vol. 1, Edinburgh University Press, Edinburgh, pp. 206–215.

Schmidhuber, J. (1990), Networks adjusting networks, *in* J. Kindermann and A. Linden, eds, 'Proceedings of Distributed Adaptive Neural Information Processing', Oldenbourg, pp. 197–208. Extended version: TR FKI-125-90 (revised), Institut für Informatik, TUM.
ftp://ftp.idsia.ch/pub/juergen/fki125.ps.gz

Si, J. and Wang, Y.-T. (2001), 'On-line learning control by association and reinforcement', *IEEE Transactions on Neural Networks* **12**(2), 264–276.
http://ebrains.la.asu.edu/~jennie/siandwang01.pdf