

# Cognitive Development Through a Neurologically-Based Learning Framework

John Nassour and Gordon Cheng  
Institute for Cognitive Systems  
Technische Universität München  
Munich, Germany  
Email: {nassour, gordon}@tum.de

**Abstract**—We develop through continuous interactions with the world, learning from our experiences of successes as well as our failures. Neurological understanding of the involved mechanisms is beginning to emerge to a level where they can be validated on robots functioning in the real world. For instances: 1) the Anterior Cingulate Cortex “ACC” has shown to contribute to Cognitive control by modulating the error-related signals for both positive as well as negative past experiences, thus, acting as an early warning system “EWS”. The notion of Vigilance helps formulate such a mechanism in the manner we learn and develop, thus the way we make decisions. 2) the Orbitofrontal Cortex “OFC” is said to play a role in the way we learn by representing the effective value of reinforcements, thus, regulating decision-making and expectation. These neural mechanisms play an underpinning role in Cognitive Development. In this paper, we show that computational models of these mechanisms have been realized on a robot that can acquire and develop new skills (e.g. walking, throwing).

## I. INTRODUCTION

The notion of “cognition” aims to capture the capability of mental activities of human beings for abstracted information from real world. It refers furthermore to their representation and storage in memory. It includes various mental processes like perception, attention, reasoning, learning, recognition, decision, as well as task coordination. All cognitive models have in general the same objective, analyze how human think, reason, remember, perceive, and learn. By studying the behavioral consequences of the brain, cognitive neuroscience promises to delineate the connections between the brain anatomy and the functionality of the human mind that is studied in cognitive psychology.

The research on cognitive neuroscience adds a biologically-inspired intelligent dimension into robotics research. In contrast to traditional robotics control, which focuses on programming robots to solve one specific task in one environment (sense-act, sense-act, sense-act, ...), cognitive robotics control aims to generate intelligence and adaptive behaviors based on animal or human thinking and learning processes (sense-learn-predict-act, sense-learn-predict-act,...).

To be considered as a human partner, human-like robots and human interactive robots should be provided with sophisticated cognitive systems based on open-ended learning toward cognitive developmental robotics. Learning can be considered to be open-ended if it handles tasks that are unknown or even not well-defined previously [1]. Nevertheless, the

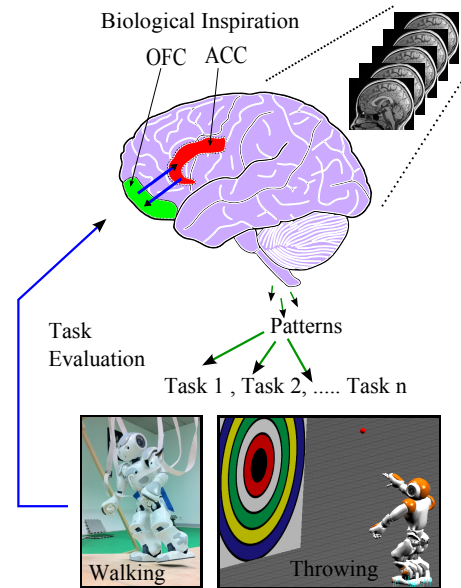


Fig. 1: The conceptual overview of our work. (“ACC” is the Interior Cingulate Cortex; “OFC” is the OrbitoFrontal Cortex.)

learning mechanisms based on fitness function minimization are limited to solve such tasks. Human develops and learns in open-ended way across its lifespan. A human child can learn tasks that he never did before, this can be referred to the mental and physical development. Traditional robots can percept and act only with the external environment, while robot based cognition can be intrinsically motivated by the internal environment, therefore, it can percept and act with the external and the internal environment to reach an intrinsic motivation (e.g. search for missing knowledge in the word-model and trigger a learning process when needed.) [2].

Human learns tasks from their own experiences by self-exploration and observation of others’ actions. The evaluation of the achieved task is driven by rewards. Human can improve their skills in order to gain more rewards. Neurobiological studies suggest that the orbitofrontal cortex (OFC) is related to reward dealing in the human brain [3]. Neurons of OFC are the key reward structure of the brain, where reward is coded in an adaptive and flexible way [4].

By observing its cortical activities, studies of the Anterior Cingulate Cortex (ACC) suggest that it is responsible to avoid

repeating mistakes [5]. This cortical area acts as an early warning system (EWS) that adjusts the behavior to avoid dangerous situations. It responds not only to the sources of errors (external error feedback), but also to earliest sources of error information available (internal error detection) [6]. EWS has shown to be affected by the tolerant to risks, psychological studies provide further evidences of people’s strategies into two classes as in taking or aversion risks [7]. A Developmental study of risky decisions suggests that risk taking behavior is related to human age and development [8].

“NeuroRobotics” research draw on human learning methods in order to improve the autonomy and the robustness of robots for their dealing with environment changes. In connection with these neurological studies, we proposed a learning method based on human learning from experiences (ACC) and inspired by the way the human brain code rewards (OFC), in order to allow a humanoid robot to learn a walking task. With the vigilance threshold concept that represents the tolerance to risk, the method guaranteed the balance between exploration and exploitation. Most task learning methods based on reward use predefined parameters in their reward function [9], [10], which cannot be obtained without previous experiences to achieve the desired task. Learning based adaptive reward don’t require any previous information about the reward, it is able to build the experience only based on the reward available information after starting from scratch.

Our approach has been implemented on the NAO humanoid robot, controlled by a bio-inspired neural controller based on a central pattern generator “CPG” [11], [12], see Figure. 1. The learning system adapts three intrinsic parameters in the CPG (the frequency of oscillation, and the motor neuron gain in pitch and roll) in order to walk on flat and sloped terrains.

## II. NEUROPSYCHOLOGICAL INSPIRATION

### A. ACC contribution in Cognitive Control

The ACC and neighbouring areas are involved in controlling and monitoring goal-directed-behavior to avoid repeating mistakes. Brown and Braver develop a computational model that shows how ACC not only detects errors, it may predict error likelihood before the error occurs [5]. The ACC is activated proportionally to the observed likelihood of the error. The error-likelihood hypothesis assumes the training signal that affects the ACC is acquired and dopaminergic. The phasic suppression of dopamine, which drives the error-related negativity (ERN) [13], [14], may play the role of a training signal that make ACC activation stronger for contexts with more frequent error. As a result of fMRI observation of subjects’ ACC, the ACC cells learn to respond with more activation for cues with high error likelihood. The results suggest that the ACC is involved in cognitive control through its risk-related cortical activity.

### B. OFC contribution in Motivational Control

The primate orbitofrontal cortex (OFC) is involved in the motivational control of goal-directed behaviour [15]. It has an essential role in controlling and correcting reward-related

and punishment-related behaviour [16]. The neurons of OFC are involved in the processing of motivational values of voluntary action rewards. OFC neurons increase their activities during the expectation of reward and after receiving it [15]. Subjects select more frequently rewards when they have to chose between different rewards at the same time, however, such frequently rewards can be ignored when more delectable rewards become available. It seems that motivational values are not fixed to defined rewards, unlike physical properties. The reward discrimination in some OFC neurons is based on the relative preference rather than the physical properties [15].

### C. OFC-ACC Connectivity During Decision-Making

Many neuroscience research study the primate brain regions involved in decision-making and other neurobiological mechanisms [17], [18]. The challenge was not only to detecting the brain regions that exhibit significantly during such mechanisms, but also to understand how different brain regions interact between each other. Cohen et al. designed a fMRI study that separates experimentally the neural activity related high-risk and low-risk choosing from other processes such as reward anticipation and evaluation during the general framework of decision making [17]. They showed that choosing high-risk over low-risk decision was related with increased activity in both ACC and OFC. It seems that OFC carries on reward associations for stimulus [19], and ACC contains mechanisms that control the selection of appropriate behaviors [20]. According to [17], no ACC activities were observed during low risk decision, while both ACC and OFC show a high activation when subjects made high-risk. However, this study was not able to distinguish whether ACC activation are related with small chance of a large reward or large chance of a failure. ACC and OFC exhibited similar patterns for activation and time courses and distinct patterns of functional connectivity. This suggests that they may play different and complementary roles in decision making [17].

### D. Risk Taking Behavior

Psychological studies show the probability of sampling with experience based learning in human is reduced with poor past outcomes [21], [22]. They show how adaptive sampling could lead to risk-averse as well as risk-seeking behaviors. Risk tendency may change according to the distribution of the uncertain alternatives and the information about foregone payoffs. Denrell et al. assume that the decision maker mostly samples the ambiguous substitution if its estimated value is positive. However, it explores eventually substitution with negative estimated value [21]. Erev and Barron have shown the role of adaptive sampling in modeling risk taking behavior for systems where decision is based on experience [23].

Based on previous studies in learning from mistakes, coding reward, and adaptive sampling for risk taking modulation; we introduce a learning mechanism that is able to learn and to evaluate humanoid tasks and to optimize its performance.

### III. SUCCESS-FAILURE LEARNING

The objective of this learning mechanism is to adapt parameters of a low-level controller and to detect their domain of viability. We designated by  $\Omega$  the state space of those influential parameters. The mechanism must be able to learn from negative feedback (failure) and positive feedback (success). Therefore it must have experience of success and other experience of failure in the state space  $\Omega$ . As each action vector  $\vec{v}$  from  $\Omega$  leads to either success or failure, the mechanism will evaluate whether this vector belongs to a success case or to a failure case. The decision mechanism “go” or “no-go” described in [24] works as an early warning system similar to that in the Anterior Cingulate Cortex [5], [25]. The learning architecture is then based on these two mechanisms and works as shown in Fig.2.

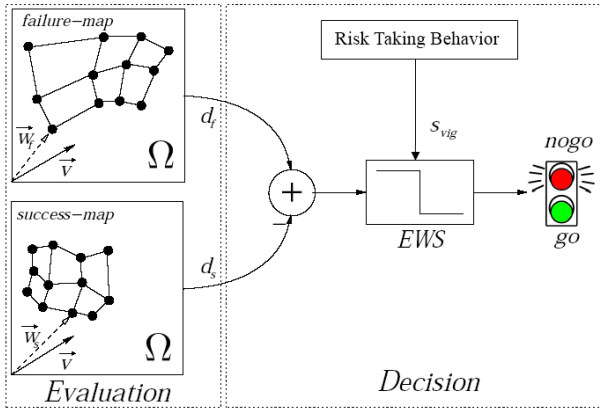


Fig. 2: Success-Failure Learning mechanism with evaluation and decision phases.

We have proposed previously a preliminary model for success-failure mechanism [26].

#### A. Evaluation Phase

To represent the knowledge in success and in failure, we define two independent neural networks that are well-known Self Organizing Maps (SOM), proposed by Kohonen [27]. Success map  $S_m$  learns in case of success trials, and failure map  $F_m$  learns in case of failure trials. During the learning, the two maps will be self-organized in the state space that will be therefore divided into three zones: 1) a zone of success represented by success map; 2) a zone of failure represented by failure map; and 3) a zone of conflict that corresponds to the overlapping between the two maps. The evaluation of any vector  $\vec{v}$  from space  $\Omega$  belonging to success or failure is defined by the distance between  $\vec{v}$  and each map. The distance of a vector with a map is the minimal Euclidean norm between this vector and the closest neuron’s weights vector in the state space (the winning neuron). For each  $\vec{v}$  we have therefore two distances: one to  $S_m$  called  $d_s$ , and another to  $F_m$  called  $d_f$ .  $d_s$  and  $d_f$  are then used for the decision process.

#### B. Decision mechanism

For a vector  $\vec{v}$ , the comparison between the distance with success map  $d_s$  and the distance with failure map  $d_f$  leads to

an expected result in the case where the vector was passed to the low level controller (trial). According to expected results, if it may lead to failure, then an Early Warning Signal (*EWS*) becomes active to avoid the passing into the lower level, and the decision will be “no-go”. When *EWS* is inactive the decision is “go”. The decision mechanism is affected by the threshold of vigilance  $s_{vig}$ .

#### C. Vigilance-Related Development

Psychological research studies suggest that some people are more tolerant to risk than others who are more cautious [28][29][30]. The vigilance is related to human learning in connection with decision making [31]. In the standard psychological assessment of risk taking, people are classed as risk seeking or risk averse [7].

Through the observation of particular areas located in cerebral cortex in the brain responsible for cognitive control, neuropsychological studies demonstrated a switching in human learning strategies around the age of twelve years. This switch from learning with positive feedback to learning with negative feedback probably comes from the combination of brain maturing and experience[8].

In our study for robot tasks learning by success and failure maps, we introduced the concept of vigilance in order to control the learning process in the two maps (success and failure) and manage the learning cycle while avoiding or taking risks according to the system’s needs. The vigilance is represented by a threshold  $s_{vig}$  that is used to adjust the early warning signal in the decision mechanism. This threshold describes the tolerance of risk.

In our previous work, the vigilance threshold was modulated according to the number of trials [32], see Figure 3. The

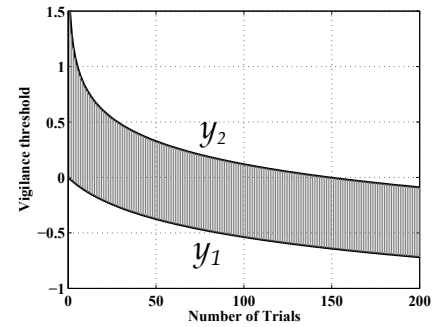


Fig. 3: Vigilance Model related to learning iterations,  $y_1 \leq s_{vig} \leq y_2$ . The risk behavior will change from prudence at the beginning of learning to adventure at the end [32].

vigilance was modulated to change the risk behavior from prudence at the beginning of learning to adventure at the end. An example of vigilance threshold modulation is given as following (see Fig.3):

$$y_1 \leq s_{vig} \leq y_2 \quad \begin{cases} y_1 = a_1 - b_1 * \log((x + c_1)^2) \\ y_2 = a_2 - b_2 * \log((x + c_2)^2) \end{cases} \quad (1)$$

The coefficients values are ( $a_1 = 0.9, a_2 = 1.47, b_1 = b_2 = 0.15, c_1 = c_2 = 20$ ) and were chosen after several attempts.

$y_1$  and  $y_2$  chosen curves ensure smooth change between the prudence and adventure above mentioned behaviors.

Vigilance modulation is an important approach that can drive the open end learning, it can be modulated in opposite way, start with taking risk then switch to risk avoiding behavior.

However, vigilance can be also modulated according to the distribution of the uncertain alternatives. In the next section, the vigilance threshold was adapted to drive the sampling process in the way that ensures success and failure maps learn and converge together.

#### D. Reward Coding

Most reinforcement learning based robotic walking studies use predefined constants to determine the maximum and the minimum reward or to determine the multiplier factors [9], [10]. Reward coding is a way that qualifies succeeded trials differently according to an optimized criterion. Each trial will have its own weighted reward representing the objective criterion to be optimized. During each learning step, neurons will get closer to trials with high rewards rather than to trials with low rewards. If the optimized criterion was the efficiency of learned task, the minimum value of the reward related multiplication factor  $r_{min}$  matches the trial with lowest efficiency. While the maximum value of the reward related multiplication factor  $r_{min}$  matches the trial with highest efficiency. This matching is done in adaptive way during the learning process, see Figure 4.

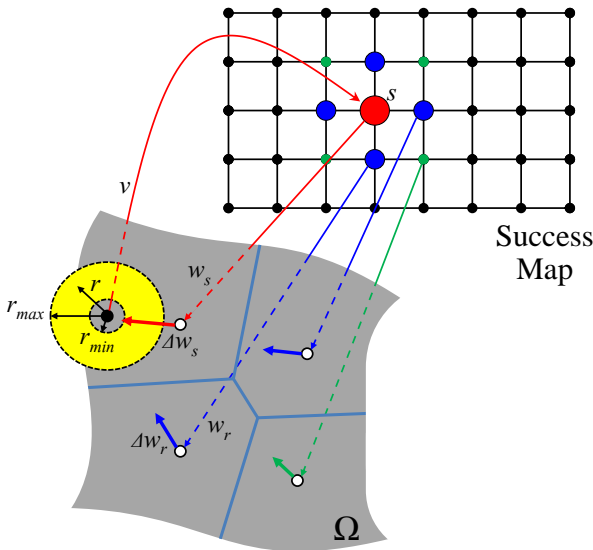


Fig. 4: Success map adaptation.  $r$  is a reward-related multiplication factor.  $r_{min}$  and  $r_{max}$  are the minimum and the maximum values for the multiplication factor.

After enough numbers of trials, this will result in a shift of the map into a spatial area associated with the highest rewards.

By introducing the concept of adaptive coding of reward it will be possible to scale the quality of a trial according to

the quality in previous experiences even when starting from scratch. After learning, the optimal parameter is presented by the success map neuron that is close to the trial with maximum reward in training set.

The general diagram of the success-failure learning is presented in Figure 5.

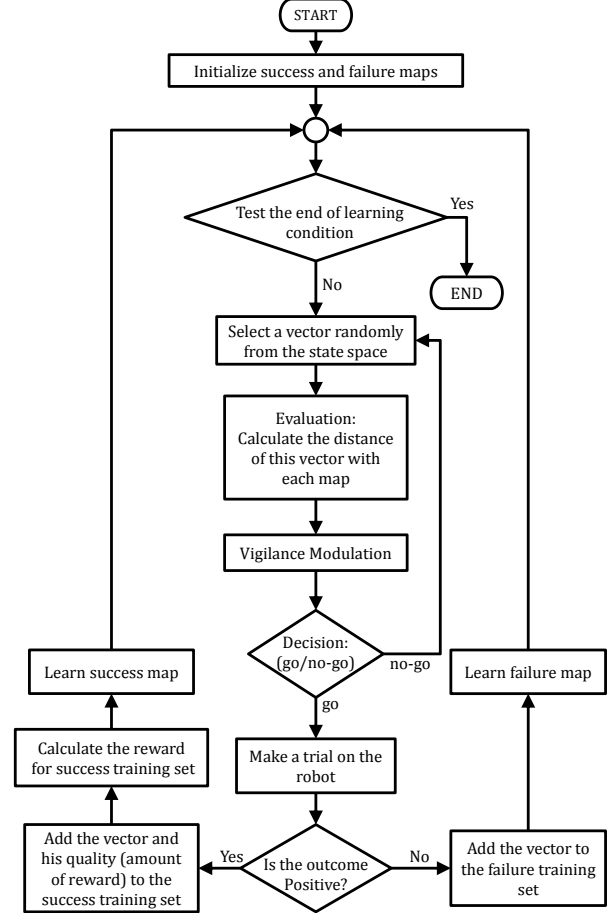


Fig. 5: Flow diagram for success-failure learning.

#### IV. LEARNING WITH ADAPTATION

We apply the learning mechanism proposed in the previous sections in order to learn efficient walking for a bipedal humanoid robot, NAO. We used success-failure learning to learn in a space of intrinsic parameters of the CPG controller (the frequency of oscillation, and the motor neuron gain in pitch and roll), the basic CPG model is presented in [32]. The optimization of walking efficiency was studied in term of energy as in [33]. The efficiency with which a muscle operates is defined in [33] by

$$efficiency = \frac{mechanical\ work\ done}{metabolic\ energy\ consumed} \quad (2)$$

This study is also generalized from a muscle to whole body movements like walking, and running [34], [35]. With inspiration from biomechanical studies, the efficiency of walking for a humanoid robot can be described in a similar fashion.

Our objective is to simultaneously learn and optimize walking with success-failure on-line learning. The robot learns to walk a 1.5[m] trajectory with start and end lines. In case of succeeded trials, the trainer sends a reward signal to the robot by caressing the head equipped with electrostatic sensors. The walking efficiency is calculated for each trial as:

$$\eta = \frac{E_k}{E_e} \quad (3)$$

Where  $E_k$  is the kinetic energy of a trial,  $E_e$  is the required electric energy for the entire trajectory.

The introduction of the efficiency for success map learning will shift the neurons of this map into the area in which the walking efficiency is high. Figure 6 shows the reward coding for success map in the beginning of learning (after four successful trials), and at the end of learning. Each sphere corresponds to a succeeded trial whose diameter represents the reward of this trial in the success map. The interest of using this technique is to make success-failure learning search for new trials in the space area where walking efficiency in term of energy is high. In other words, this leads to learn and optimize in a defined space. Figure 7 shows success maps after learning to walk on flat terrain with and without the technique of reward coding adaptively. In Figure 7(a), the success map learns all successful trials with the same opportunity, i.e. with the same reward. In Figure 7(b), the success map learns successful trials in accordance with their adaptive rewards. Trials with high reward influence success map neurons more than trials with low reward. Therefore, the success map will be attracted to the area where reward is high. This is influenced by the differences between highest and lowest rewards (scaling range limits:  $[r_{min}, r_{max}]$ ). In this study,  $r_{min}$  and  $r_{max}$  are set to 0.1 and 2.5.

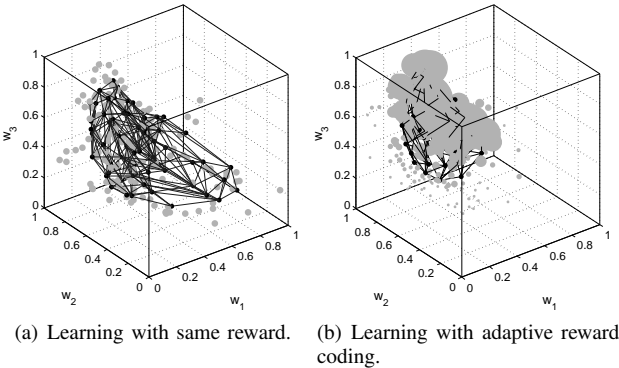


Fig. 7: The effect of reward coding on success map. Success map after learning with the same reward for all successful trials (a). Success map after learning with adaptive reward coding (b). Gray spots represent successful trials reward. Note that the map on the right moves into the area where rewards are high (representing high efficiency).

Regarding the learning frameworks with and without the application of reward coding adaptively shown in Figure 7,

performance was increased by 60%. This was calculated by the ratio of the highest efficiencies matched by neurons of both success maps. The ratio of the lowest efficiency of the neurons of success maps has increased by 40%. In order to provide sufficient precision in the network for our task, we have empirically selected a  $5 \times 5 \times 5$  dimensional network space to represent the success and failure maps. Learning occurred with 500 trials for each case. Computationally, all the processing of this learning framework in simulations as well as on the real robot can be performed in real-time, thus making our approach feasible for training on the real robot. Within the same cycle, joint angle commands are calculated in real-time and sent to joint motor circuit boards of NAO every 10[ms]. This is done inside a high priority thread on the robot. Physically each trial requires about 3 minutes, which includes learning and the experimental set up. A complete learning session in the robot usually takes about one week. Both Learning frameworks shown in Figure 7 start from scratch. Adaptive sampling driven by vigilance threshold ensures to have the same size of training sets to learn success map and failure map. (A video shows NAO humanoid robot achieving the walking task is available on: <http://web.ics.ei.tum.de/~nassour/naowalking.wmv>.)

## V. CONCLUSION

This paper has brought several ideas from different bodies of research. Research in machine learning, neuroscience, psychology, and robotics are involved in cognitive development. Understandings within human brain research help provide human-like learning mechanisms that can be implemented on robots. Our neurologically grounded learning framework imitates part of the functionality of the anterior cingulate cortex involved in learning from mistakes, and the orbitofrontal cortex involved in reward coding adaptively. This Success-Failure learning cycle forms an important part of our cognitive development architecture in order for robots to learn and acquire different physical and mental skills.

In this paper, we showed how vigilance is modulated differently as our robot develops over time under different conditions during decision making. These vigilance modulations can be said to play an important role throughout a lifetime of human as well as robot developments.

## REFERENCES

- [1] R. Gomes, "Towards open ended learning: budgets, model selection, and representation," Ph.D. dissertation, California Institute of Technology, 2011.
- [2] M. Oubbati and G. Palm, "Self-motivated learning robot," *9th Epigenetic Robotics International Conference (Epirob09): Modeling Cognitive Development in Robotic Systems, Venice, Italy, November 12-14, 2009*.
- [3] S. D. Iversen and M. Mishkin, "Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity," *Experimental Brain Research*, vol. 11, pp. 376–386, 1970.
- [4] S. Kobayashi, O. P. de Carvalho, and W. Schultz, "Adaptation of reward sensitivity in orbitofrontal neurons," *The Journal of Neuroscience*, vol. 30, no. 2, pp. 534–544, January 2010.
- [5] J. W. Brown and T. S. Braver, "Learned predictions of error likelihood in the anterior cingulate cortex," *Science*, vol. 307, pp. 1118–1121, 2005.
- [6] R. B. Mars, M. G. Coles, M. J. Grol, C. B. Holroyd, S. Nieuwenhuis, W. Hulstijn, and I. Toni, "Neural dynamics of error processing in medial frontal cortex," *Neuroimage*, vol. 28, no. 4, pp. 1007–1013, December 2005.

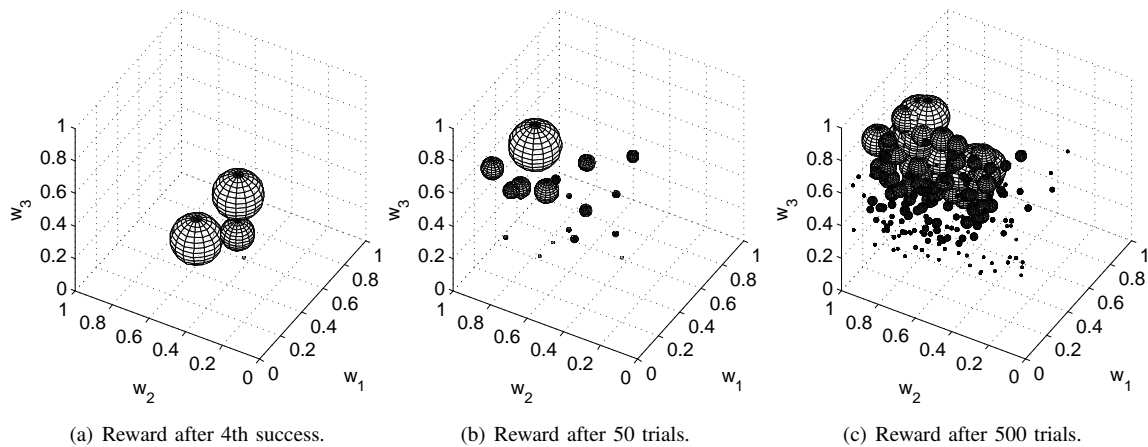


Fig. 6: Succeeded trials' reward related to walking efficiency for learning success map. Where  $w_1$  is motor neuron gain in pitch,  $w_2$  is motor neuron gain in roll, and  $w_3$  is  $\sigma_s$  that is related to the oscillation frequency.

- [7] X. Wang, D. Kruger, and A. Wilke, "Towards the development of an evolutionarily valid domain-specific risk-taking scale," *Evolutionary Psychology*, vol. 5, no. 3, pp. 555–568, 2007.
- [8] L. Van Leijenhorst, P. M. Westenberg, and E. A. Crone, "A developmental study of risky decisions on the cake gambling task: Age and gender analyses of probability estimation and reward evaluation," *Developmental Neuropsychology*, vol. 33, no. 2, pp. 179–196, 2008.
- [9] T.-H. S. Li, Y.-T. Su, S.-W. Lai, and J.-J. Hu, "Walking motion generation, synthesis, and control for biped robot by using pgrl, lpi, and fuzzy logic," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 41, no. 3, pp. 736–748, 2011.
- [10] G. Endo, J. Morimoto, T. Matsubara, J. Nakanishi, and G. Cheng, "Learning cpg-based biped locomotion with a policy gradient method: Application to a humanoid robot," *International Journal of Robotics Research*, vol. 27, pp. 213–228, February 2008.
- [11] J. Nassour and G. Cheng, "Biologically-inspired neural controller based on adaptive reward learning," *Front. Comput. Neurosci. Conference Abstract: Bernstein Conference*, 2012.
- [12] F. B. John Nassour, Vincent Hugel and G. Cheng, "Qualitative adaptive reward learning with success failure maps: Applied to humanoid robot walking," *IEEE Transactions on Neural Networks and Learning Systems [In press]*, 2012.
- [13] W. J. Gehring, M. G. H. Coles, D. E. Meyer, and E. Donchin, "The error-related negativity: An event-related potential accompanying errors," *Journal of Psychophysiology*, vol. 27, p. S34, 1990.
- [14] W. J. Gehring, Y. Liu, J. M. Orr, and J. Carp, *Oxford handbook of event-related potential components*. New York: Oxford University Press., 2012, ch. The error-related negativity (ERN/Ne), pp. 231–291.
- [15] L. Tremblay and W. Schultz, "Relative reward preference in primate orbitofrontal cortex," *Nature*, vol. 398, no. 6729, pp. 704–8, April 1999.
- [16] E. T. Rolls, "The orbitofrontal cortex," *Philosophical Transactions of the Royal Society B*, vol. 351, pp. 1433–1444, 1996.
- [17] M. Cohen, A. Heller, and C. Ranganath, "Functional connectivity with anterior cingulate and orbitofrontal cortices during decision-making," *Cognitive Brain Research*, vol. 23, no. 1, pp. 61 – 70, 2005.
- [18] E. Bicho, W. Erlhagen, L. Louro, and E. C. e Silva, "Neuro-cognitive mechanisms of decision making in joint action: A human-robot interaction study," *Human Movement Science*, vol. 30, no. 5, pp. 846 – 868, 2011.
- [19] E. T. Rolls, "The orbitofrontal cortex and reward," *Cerebral Cortex*, vol. 10, no. 3, pp. 284–294, 2000.
- [20] V. Van Veen, J. D. Cohen, M. M. Botvinick, V. A. Stenger, and C. S. Carter, "Anterior cingulate cortex, conflict monitoring, and levels of processing," *NeuroImage*, vol. 14, no. 6, pp. 1302–1308, 2001. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/17604651>
- [21] J. Denrell, "Adaptive learning and risk taking," *Psychol Rev.*, vol. 114, no. 1, pp. 177–187, 2007.
- [22] J. G. March, "Learning to be risk averse," *Psychological Review*, vol. 103, no. 2, pp. 309–319, Apr 1996.
- [23] G. Erev, Ido; Barron, "On adaptation, maximization, and reinforcement learning among cognitive strategies," *Psychological Review*, vol. 112, no. 4, pp. 912–931, Oct 2005.
- [24] K. Matsumoto, W. Suzuki, and K. Tanaka, "Neuronal correlates of goal-based motor selection in the prefrontal cortex," *Science*, vol. 301, no. 5630, pp. 229–32, 2003.
- [25] J. W. Brown and T. S. Braver, "A computational model of risk, conflict, and individual difference effects in the anterior cingulate cortex," *Brain research*, vol. 1202, pp. 99 –108, 2008.
- [26] J. Nassour, P. Henaff, F. B. Ouezdou, and G. Cheng, "Experience-based learning mechanism for neural controller adaptation: Application to walking biped robots," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 11-15, St. Louis, MO, USA, 2009*, pp. 2616–2621.
- [27] T. Kohonen, *Self-Organizing Maps*, ser. Springer Series in Information Sciences. Berlin, Heidelberg: Springer, 1995, vol. 30.
- [28] J.-L. van Gelder, R. E. de Vries, and J. van der Pligt, "Evaluating a dual-process model of risk: affect and cognition as determinants of risky choice," *Journal of Behavioral Decision Making*, vol. 22, no. 1, pp. 45–61, 2009.
- [29] B. Pawlowski, R. Atwal, and R. I. M. Dunbar, "Sex differences in everyday risk-taking behavior in humans," *Evolutionary Psychology*, vol. 6, no. 1, pp. 29–42, 2008.
- [30] P. Horvath and M. Zuckerman, "Sensation seeking, risk appraisal, and risky behavior," *Personality and Individual Differences*, vol. 14, no. 1, pp. 41–52, 1993.
- [31] H. Ahn and R. W. Picard, "Affective-cognitive learning and decision making: A motivational reward framework for affective agent," in *The 1st International Conference on Affective Computing and Intelligent Interaction. October 22-24, 2005*.
- [32] J. Nassour, P. Hénaff, F. B. Ouezdou, and G. Cheng, "A study of adaptive locomotive behaviors of a biped robot: patterns generation and classification," in *Proceedings of the 11th international conference on Simulation of adaptive behavior: from animals to animats*, ser. SAB'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 313–324.
- [33] B. Abernethy, *The biophysical foundations of human movement*, 2nd ed. HUMAN KINETICS, 2005.
- [34] N. Berryman, M. Gayda, A. Nigam, M. Juneau, L. Bherer, and L. Bosquet, "Comparison of the metabolic energy cost of overground and treadmill walking in older adults," *European Journal of Applied Physiology*, pp. 1–8, 24 August 2011.
- [35] R. Margaria, *Biomechanics and energetics of muscular exercise*, first edition ed. Oxford University Press, USA, November 18 1976.