Prediction and Anticipation for Adaptive Artificial Limbs

Patrick M. Pilarski, Michael R. Dawson, Thomas Degris, Jason P. Carey, K. Ming Chan, Jacqueline S. Hebert, and Richard S. Sutton

PREDICTING THE FUTURE has long been regarded as a powerful means to improvement and success. The ability to make accurate and timely predictions enhances our ability to control our situation and our environment. Assistive robotics is one prominent area where foresight of this kind can bring improved quality of life. In this article, we present a new approach to acquiring and maintaining predictive knowledge during the online, ongoing operation of an assistive robot. The ability to learn accurate, temporally abstracted predictions is shown through two case studies—able-bodied subjects engaging in the myoelectric control of a robot arm, and an amputee participant's interactions with a myoelectric training robot. To our knowledge, this work is the first demonstration of a practical method for real-time prediction learning during myoelectric control. Our approach therefore represents a fundamental tool for addressing one major unsolved problem: amputee-specific adaptation during the ongoing operation of a prosthetic device. The findings in this article also contribute a first explicit look at prediction learning in prosthetics as an important goal in its own right, independent of its intended use within a specific controller or system. Our results suggest that real-time learning of predictions and anticipations is a significant step towards more intuitive myoelectric prostheses and other assistive robotic devices.

Myoelectric Prostheses

Assistive biomedical robots augment the abilities of amputees and other patients with impaired physical or cognitive function due to traumatic injury, disease, aging, or congenital complications. In this article we focus on one representative class of assistive robots: myoelectric artificial limbs. These prostheses monitor electromyographic (EMG) signals produced by muscle tissue in a patient's body, and use the recorded signals to control the movement of a robotic appendage with one or more actuators. Myoelectric limbs are therefore tightly coupled to a human user, with control processes that operate at high frequency and over extended periods of time. Commercially available devices include powered elbow, wrist, and hand assemblies from a number of suppliers (Fig. 1). Research into technologies and surgeries related to next-generation artificial limbs is also being carried out internationally at a number of institutions [1]–[10].

Fig. 1. **Example of a commercially available myoelectric prosthesis with multiple joints and functions.** The intuitive myoelectric control of multiple actuators and functions is a challenging problem.

Despite the potential for improved functional abilities, many patients reject the use of powered artificial limbs [1]–[3]. Recent studies point out three principal reasons why amputees reject myoelectric forearm prostheses. These include a lack of intuitive control, insufficient functionality, and insufficient feedback from the myoelectric device [1], [3]. As noted throughout the literature, the identified issues extend beyond the domain of forearm prostheses and are major barriers for upper-body myoelectric systems of all kinds.

Challenges facing myoelectric prosthesis users are currently being addressed through improved medical techniques, new prosthetic technologies, and advanced control paradigms. In the first approach, medical innovation with targeted motor and sensory reinnervation surgery is opening new ground for intuitive device control and feedback [8], [11]. In the second approach, prosthetic hardware is being enhanced with new sensors and actuators; state-of-the-art robotic limbs now begin to approach their biological counterparts in terms of their capacity for actuation [9], [10].

In this article we focus on the third approach to making myoelectric prosthesis use more accessible—improving the control system. The control system is a natural area for improvement, as it links sensors and actuators for both motor function and feedback. Starting with classical work in the 1960's, both industry and academia have presented a wide range of increasingly viable myoelectric control approaches.

Cite as: P. M. Pilarski, M. R. Dawson, T. Degris, J. P. Carey, K. M. Chan, J. S. Hebert, and R. S. Sutton, "Prediction and Anticipation for Adaptive Artificial Limbs," *IEEE Robotics and Automation Magazine*, Special Issue on Assistive Robotics, March 2013, In Press.

When compared to traditional body-powered hook and cable systems, myoelectric approaches represent a move toward the more natural, physiologically intuitive operation of a prosthetic device. Within the space of myoelectric control strategies, conventional proportional control is still considered the mainstay for clinically deployed prostheses—in this approach, the amplitude of one or more EMG signals is proportionally mapped to the input of one or more powered actuators.

Despite widespread use, conventional myoelectric control has a number of known limitations [4]. One primary challenge for conventional control is the growing actuation capabilities of current devices. Conventional control is further constrained by the limited number of signal sources in a residual amputated limb. This problem becomes more pronounced with higher levels of amputation; patients who have lost more function will require more complex assistive devices, but have fewer discrete sources from which to acquire control information [2], [4]. Even with advanced function switching, conventional approaches are only able to make use of a fraction of the movements available to next-generation devices. Intuitive, simultaneous actuation of multiple joints using myoelectric control remains a challenging unsolved problem [2].

Pattern Recognition

The dominant approach for improving myoelectric control has been the use of pattern recognition techniques. As reviewed by Scheme and Englehart [4], the state-of-the-art for myoelectric pattern recognition relies on sampling a number of training examples in the form of recorded signals, identifying relevant features within these signals, and then classifying these features into a set of control commands. This approach has been largely implemented in an *of ine* context, meaning that systems are developed and trained and then not changed during regular (non-calibration) use by an amputee. Demonstrated offline methods include support vector machines, linear discriminant analysis, artificial neural networks, and principal component analysis on time and frequency domain EMG information [3]–[5]. Offline pattern recognition approaches are straightforward to deploy, and have allowed amputees to successfully control both conventional and state-of-the-art myoelectric prostheses in real time [8]; the training time of these methods is also realistic for use by amputees.

Though less common, myoelectric pattern recognition systems can also be trained *online*, meaning that they continue to be changed during normal use by the patient. Examples within the domain of prosthetic control include the use of artificial neural networks [12] and reinforcement learning [7]. In both cases, feedback from the user or the control environment was used to continually update and adapt the device's control policy. Online adaptation is critical to robust long-term myoelectric prosthesis use by patients, and is currently an area of great clinical interest [4], [5]. As discussed by Scheme and Englehart, and Sensinger et al., there have been a number of initial studies on adapting control to changes in the electrical and physical aspects of EMG electrodes (e.g., positional shifts and conductivity differences), electromagnetic interference, and signal variation due to muscle fatigue [4], [5]. This work has made it clear that myoelectric control must take into account real-time changes to the control environment, a patient's physiology, and their prosthetic hardware. To be effective in practice, adaptive methods need to be robust, easily trained, and not a time burden to the patient. However, a robust, unsupervised approach to online adaptation has yet to be demonstrated [4].

Prediction in Adaptive Control

A key insight underpinning prior work in adaptive or robust systems is that accurate and up-to-date predictive knowledge is a strong basis for modulating control. Prediction has proved to be a powerful tool in many classical control situations (e.g., model predictive control). Although classical predictive controllers provide a noticeable improvement over non-predictive approaches, they often require extensive offline model design; as such, they have limited ability to adapt their predictions during online use.

Prediction is also at the heart of current offline machine learning and pattern recognition prosthesis control techniques. Given a context (e.g., moment-by-moment EMG signals), pattern recognition approaches use information extracted from their training process to identify (classify) the current situation as one example from a set of motor tasks. In other words, they perform a state-conditional prediction of the user's motor intent, which can then be mapped to a set of actuator commands. As a recent example, Pulliam et al. used a time-delayed artificial neural network, trained offline, to predict upper-arm joint trajectories from EMG data [13]. The aim of their work was to demonstrate a set of predictions that could be used to facilitate coordinated multi-joint prosthesis control.

In this article, we present a new approach for acquiring and maintaining predictive knowledge during the real-time operation of a myoelectric control system. A unique feature of our approach is that it uses ongoing experience and observations to continuously refine a set of control-related predictions. These predictions are learned and maintained in real time, independent of a specific myoelectric decoding scheme or control approach. As such, the described techniques are applicable to both conventional control and existing pattern recognition approaches. We demonstrate online prediction learning in two experimental settings: 1) able-bodied subject trials involving the online myoelectric control of a humanoid robot limb and 2) trials involving control interactions between an upper-limb amputee participant and a myoelectric training robot. Our online prediction learning approach contributes a novel gateway to unsupervised, user-specific adaptation. It also provides an important tool for developing intuitive new control systems that could lead to improved acceptance of myoelectric prostheses by upper-limb amputees.

An Approach to Online Prediction Learning for Myoelectric Devices

Systems that perform online prediction and anticipation are in essence addressing the very natural question "what will happen next?" To be of benefit to an adaptive control system, this question must be posed in computational terms, and its answer must be continually updated online from real-time sensorimotor experience.

There are a number of predictive questions that have a clear bearing on myoelectric control. Examples of useful questions include: "what will be the average value of a grip sensor over the next few seconds?", "where will an actuator be in exactly 30 seconds?", or "how much total myoelectric activity will be recorded by an EMG sensor in the next 250ms?" These anticipatory questions have direct application to known problems for rehabilitation devices—issues like grip slippage detection, identifying user intent, safety, and natural multi-joint movement [1], [4]. It is important to note that questions of this kind are temporally extended in nature; they address the expected value of signals and sensors over protracted periods of time, or at the moment of some specific event. They are also context dependent, in that they rely on the current state and activity of the system. For example, predictions about future joint motion may depend strongly on whether an amputee is playing tennis or driving a car.

A predictive system should be able to express knowledge about the value of signals that will be observed in the near future—for instance, the expected value of sensor and actuator readings over a time frame ranging from the next millisecond to the next few seconds or minutes (consider the predictive system demonstrated by Pulliam et al., which learned to anticipate scalar joint angle signals through offline training [13]). A system should also be able to predict the outcome of events with no fixed length, or those that take a variable amount of time to return an outcome. Anticipations of this kind represent a common form of knowledge, but one which falls outside the learning capabilities of most standard pattern recognition approaches. To date there are few approaches able to learn this form of anticipatory knowledge for real-valued signals, and fewer still which can learn and continually update (adapt) this type of predictive representation during online, real-time operation.

Reinforcement learning (RL) is one form of machine learning that has demonstrated the ability to learn in an ongoing, incremental prediction and control setting [14]. An RL system uses interactions with its environments to build up expectations about future events. Specifically, it learns to estimate the value of a one-dimensional feedback signal termed *reward*; these estimates are often represented using a *value function*—a mapping from observations of the environment to expectations about future reward.

RL is viewed as an approach to artificial intelligence, natural intelligence, optimal control, and operations research. Since development in the 1980's, RL algorithms have come to be widely used in robotics, and have found the best known approximate solutions to many games; they have also become the standard model of reward processing in the brain [15].

Recent work has provided a straight-forward way to use RL for acquiring expectations and value functions pertaining to non-reward signals and observations [16]. These *general value functions* (GVFs) are proposed as way of asking and answering temporally extended questions about future sensorimotor experience. Predictive questions can be defined for different time scales, and may take into account different methods for weighting the importance of future observations. The anticipations learned using GVFs can also depend on numerous strategies for choosing control actions (policies), and can be defined for events with no fixed length [16]. Expectations comprising a GVF are acquired using standard RL techniques; this means that learning can occur in an incremental, online fashion, with constant demands in terms of both memory and computation. The approach we develop in this paper is to apply GVFs alongside myoelectric control.

Formalizing Predictions with GVFs

We use the standard RL framework of states ($\mathbf{s}$ $\mathcal{S}$), actions (a $\mathcal{A}$), time steps $t \geq 0$, and rewards (r) [14]. In our context, a GVF represents a question q about a scalar signal of interest, here denoted r for consistency; this question depends on a given probability function for choosing actions $\pi : \mathcal{S} \times \mathcal{A}$ $[0, 1]$ and a temporal continuation probability $\gamma :$ $\mathcal{S}$ $[0, 1]$. A question q may therefore be written as: "given state $\mathbf{s}$, what is the expected value of the cumulative sum of a signal r while following a policy π, and while continuing with a probability given by γ?" Formally, the value function $V_q(\mathbf{s})$ for our question is defined as follows, where actions are taken according to π and there exist state-dependent continuation probabilities $0 \leq \gamma(\mathbf{s}) \leq 1$ for all s $\mathcal{S}$:

$$V_q(\mathbf{s}) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \left(\prod_{i=1}^{k} \gamma(\mathbf{s}_{t+i}) \right) r(\mathbf{s}_{t+k+1}) \middle| \mathbf{s}_t = \mathbf{s} \right].$$

This defines the exact answer to a question when states are fully observable. In practice, a state is rarely fully observable or needs to be approximated to represent an answer to a question. We instead assume that we observe a vector of features $\mathbf{x}$ that depends on the current state $\mathbf{s}$ according to some state approximation function $\mathbf{x}$ $\mathsf{approx}(\mathbf{s})$. We can then present the approximate answer $\hat{V}_q(\mathbf{s})$ to a question q as a prediction P_q that is the linear combination of a (learned) weight vector $\mathbf{w}$ and the feature vector $\mathbf{x}$ at time t:

$$P_q = \hat{V}_q(\mathbf{s}) = \mathbf{w}_q^\top \mathbf{x}$$

For our work, the approximation function $\mathsf{approx}(\mathbf{s})$ was implemented using tile coding, as per Sutton and Barto [14]. Tile coding is a linear mathematical function that maps a real-valued signal space into a linear (binary) vector form that can be used for efficient computation and learning [14].

To facilitate incremental computation, in what follows we consider the exponentially discounted case of GVFs, where $0 \leq \gamma \leq 1$, and γ is the same for all states in the system. The value for state $\mathbf{s}$ is therefore the expected sum of an exponentially discounted signal r for each future timestep:

$$V_q(\mathbf{s}) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k r(\mathbf{s}_{t+k+1}) \middle| \mathbf{s}_t = \mathbf{s} \right].$$

The cumulative value inside this expectation is termed the *return*. For the purposes of post-hoc comparison, the true return R_q on a time step t may be computed by recording future experience over a window of T data points, where T is large enough that γ^T approaches zero. The difference

$P_q - R_q$ between the predicted and computed return values on any given time step t is then a measure of the absolute return prediction error on that time step. In what follows, we discuss the *time scale* of a prediction, meaning the time constant of the return predictions defined by $1/(1-\gamma)$. Predictions and errors are also presented in temporally normalized form to allow for visual comparison on a similar scale. For normalized return predictions ($\bar{P}_q$) and normalized mean absolute return errors (NMARE), return values are scaled according to the time constant, e.g., using $\bar{P}_q = P_q/(1-\gamma)$.

Learning Predictions with Temporal-Difference Methods

An application of GVFs to the myoelectric control setting is depicted in Fig. 2. Here sensorimotor signals from a classical myoelectric control environment (light grey boxes) are used as input to a function approximation routine; the resulting feature vector, and some subset of the input signals, are used to update a set of GVFs. These GVFs can then be used to predict the anticipated future value(s) for signals of interest.

For this work, GVFs were implemented as described by Sutton et al. [16]. Anticipations were then learned in an incremental fashion during online operation. This was done using temporal-difference (TD) learning, a standard technique from RL; for additional detail on TD learning and the TD(λ) algorithm, please see Sutton and Barto [14], and Sutton [17]. A procedural description of this approach follows.

At the beginning of each experiment, the weight vectors $\mathbf{w}_q$ were initialized to starting conditions. On every time step the system received a new set of observations s'. As learning progressed, each GVF was updated with the value of an instantaneous signal of interest, denoted r_q, using the TD(λ) algorithm. Predictions P_q for the signals of interest r_q, $q \quad \mathcal{Q}$, could then be sampled and recorded online using the linear combination $P_q \quad \mathbf{w}_q^T\mathbf{x}$. Learning proceeded as follows:

initialize: $\mathbf{w}, \mathbf{e}, \mathbf{s}, \mathbf{x}$
loop
  observe $\mathbf{s}$
  $\mathbf{x}' \leftarrow \mathsf{approx}(\mathbf{s})$
  for all questions $q \in \mathcal{Q}$ **do**
    $\delta \leftarrow r_q + \gamma_q \mathbf{w}_q^T\mathbf{x}' - \mathbf{w}_q^T\mathbf{x}$
    $\mathbf{e}_q \leftarrow \mathsf{min}(\lambda\mathbf{e}_q + \mathbf{x}, 1)$
    $\mathbf{w}_q \leftarrow \mathbf{w}_q + \alpha\delta\mathbf{e}_q$
    $\mathbf{x} \leftarrow \mathbf{x}'$
  end for
end loop

As shown in the procedure above, weight vectors $\mathbf{w}_q$ for each GVF were updated using both the state approximation $\mathbf{x}'$ and the signal of interest r_q. For each question q, a temporal-difference error signal (denoted δ) was computed using the signal r_q and the difference between the current and discounted future predictions for this signal ($\mathbf{w}_q^T\mathbf{x}$ and $\gamma_q\mathbf{w}_q^T\mathbf{x}'$, respectively). Next, an eligibility trace $\mathbf{e}_q$ of the current feature vector was updated in a replacing fashion, where $\mathbf{e}_q$ was a vector of the same size as $\mathbf{x}$. The trace $\mathbf{e}_q$ was used alongside the error signal δ to update the weight vector $\mathbf{w}_q$ for each GVF, with $\alpha > 0$ as a scalar step-size parameter and λ as the trace decay rate. Replacing traces were used as a technique to speed learning; for more detail, please see Sutton and Barto [14].

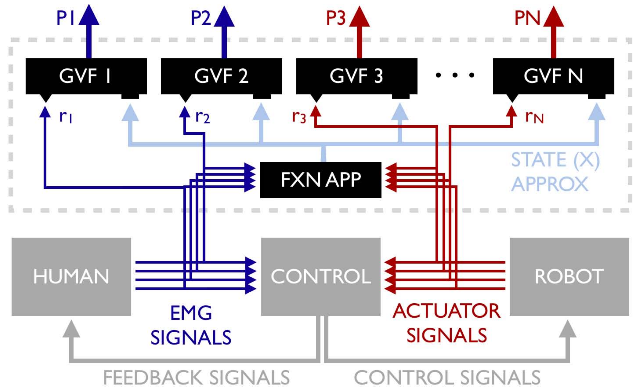

Fig. 2. **Schematic showing how general value functions (GVFs) predict the expected future value of signals** from the sensorimotor space of a myoelectric control system. Each GVF learns temporally extended predictions P_q about a specific signal of interest r_q. Predictions are learned with respect to the current state of the system, as represented by the feature vector $\mathbf{x}$. This feature vector is generated from the observed sensorimotor signal space using a function approximation routine, shown here as FXN APP'.

The per-time-step computational complexity of this procedure grows linearly with the size of the feature vector, making it suitable for real-time online learning. Linear computation and memory requirements are important for myoelectric control—when using the approach presented above, increasing the number of control sources or feedback signals leads to only a linear (and not exponential) increase in the computational demand of the learning system. Many GVFs can therefore be learned in parallel and during online real-time operation [18].

Case Study 1: Prediction During Online Control

As a first application example, we examined the ability of a GVF-based learning system to predict and anticipate human and robot signals during online interactions between able-bodied (non-amputee) subjects and a robotic device (Fig. 3). Specifically, we examined the anticipation of two signal types: user EMG signals, and the angular position of user-controlled elbow and hand joints of a robot limb. The robotic platform for these experiments was a Nao T14 robot torso (Aldebaran Robotics, France), shown alongside myoelectric recording equipment in Fig. 3. EMG signals used in device control and learning were obtained via a Bagnoli-8 (DS-B03) EMG System with DE-3.1 Double Differential Detection EMG sensors (Delsys, Boston, USA), and a NI USB-6216 BNC analog to digital converter (National Instruments Canada).

Testing was done with seven able-bodied subjects of varying age and gender. All were healthy individuals with no neurological or motor impairments. To generate a rich stream of sensorimotor data in an online, interactive setting, these participants worked with the robotic platform to complete a series of randomized actuation tasks. Participants actuated one of the robot's arms using conventional myoelectric control with linear proportional mapping. EMG signals were sampled and processed according to standard procedures from four input electrodes affixed to the biceps, deltoid, wrist flexors, and wrist extensors of a participant's dominant arm. Pairs of

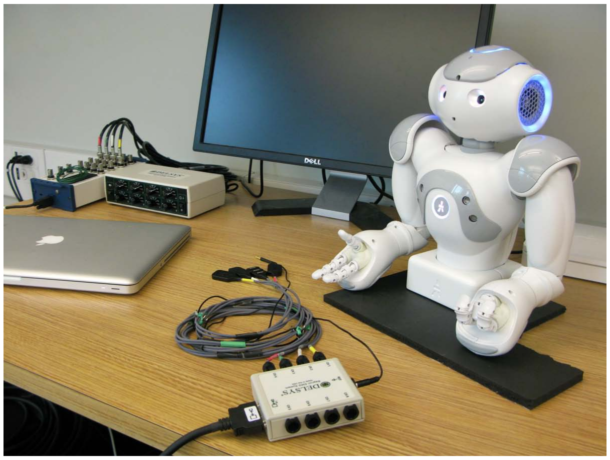

Fig. 3. **The experimental setup for able-bodied subject trials**, including an Aldebaran Nao T14 robotic platform, laptop computer, analog to digital converter, and Bagnoli-8 EMG recording system.

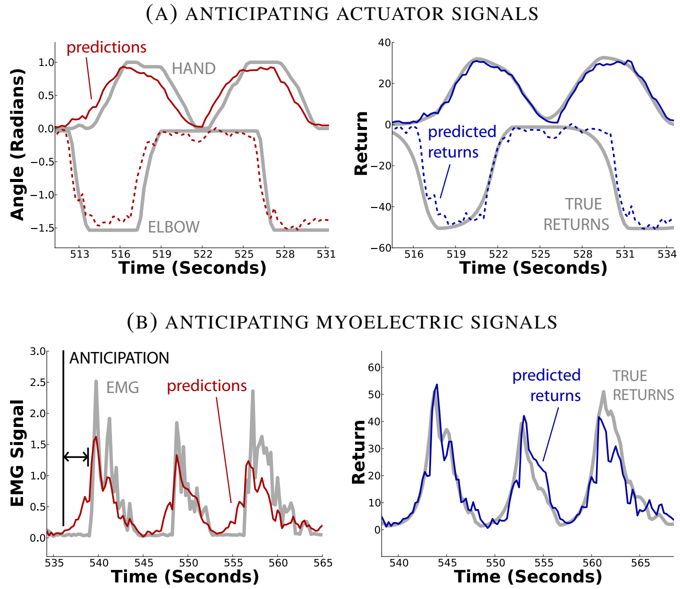

Fig. 4. **Examples of (a) actuator and (b) myoelectric signal prediction during able-bodied subject trials** after $\sim$10min of online learning. *Left:* normalized return predictions (red traces) precede the observed signal activity (grey lines) by 0.5–2.0s. *Right:* return predictions (blue traces) are consistent with the true return as computed post-hoc (grey lines).

processed signals were then mapped into velocity control commands for the robot's elbow roll actuator and hand open/close actuator. In each task, one arm of the Nao robot was moved to display a new gesture consisting of a static angular position on both hand and elbow actuators. Subjects were asked to make a corresponding gesture with the robot arm under their control. Once a subject maintained the correct position for a period of more than two seconds, a new (random) target configuration was displayed. Visual feedback to participants consisted of a front-on view of the robot system. Subjects performed multiple sessions of the randomized actuation task, with each session lasting between five and ten minutes. No subject-specific changes to the learning system were made; all subjects used exactly the same learning system with the same learning parameters, set in advance of the trials. To assess the real-time adaptation of learned predictions, additional testing was done via longer unstructured interaction sessions, some of which lasted over one hour and included tasks that produced moderate muscle fatigue.

As depicted in Fig. 2, the learning system observed the stream of data passing between the human, the myoelectric controller, and the robot arm. We created two GVFs for each of the different signals of interest r_q in the robotic system—one to predict temporally extended signal outcomes at a short time scale ($\sim$0.8s), and one to predict outcomes at a slightly longer time scale ($\sim$2.5s). As done in previous work [7], the learning system was presented with a signal space consisting of robot joint angles and processed EMG signals; at every time step, the function approximation routine mapped these signals into the binary feature vector $\mathbf{x}'$ used by the learning system. All signals were normalized between 0.0 and 1.0 according to their maximum possible ranges. Parameters used in the temporal-difference learning process were $\lambda = 0.999$, $\gamma = \ 0.97, 0.99\ $, and $\alpha = 0.033$. Weight vectors $\mathbf{e}$ and $\mathbf{w}$ for each GVF were initialized to 0. Learning updates occurred at 40Hz, and all processing related to learning, EMG acquisition, and robot control was done on a single MacBook Pro 2.53 GHz Intel Core 2 Duo laptop.

Results

We found that predictions learned using our GVF approach successfully anticipated measured signals after only short periods of online learning. Figure 4a (left) shows an example of joint angle prediction for the 0.8s time scale with one subject after $\sim$10min of learning. Here changes to the normalized return prediction signals $\bar{P}_q$ for both the hand (solid red trace) and elbow (dotted red trace) joints can be seen to occur in advance of changes to the measured actuator signals (wide grey lines). Predictions for both joints can be seen to precede actual joint activity by 0.5–2.0s. The system was also able to accurately predict myoelectric signals (Fig. 4b, left). Normalized EMG predictions (red line) rise visibly in advance of the actual myoelectric events (grey line), and changes to the processed myoelectric signal were anticipated up to 1500ms before change actually occurred. The accuracy of predictions for both actuator and myoelectric signals can be seen in Fig. 4a,b (right). For both slow and fast changes in the signal of interest, the return prediction (P_q, blue line) largely matched the true return as computed post-hoc (R_q, grey line), indicating similarity between learned predictions and computed returns.

As shown in Fig. 5, accurate predictions could be formed in 5min or less of real-time learning. These learning curves show the relationship between prediction error and training time, as averaged across multiple trial runs by one of the subjects. Learning progress for joint angle prediction and myoelectric signal prediction is shown in terms of the NMARE for the 0.8s time scale, averaged into 20 bins; error bars indicate the standard deviation (σ) over eight independent trials. After five minutes of learning, the average NMARE for both actuators was less than 0.2 radians (11.5°) for the 0.8s time scale and 0.3 radians (17.2°) for the 2.5s time scale. For the prediction of myoelectric signals, the average NMARE was less than 0.15V,

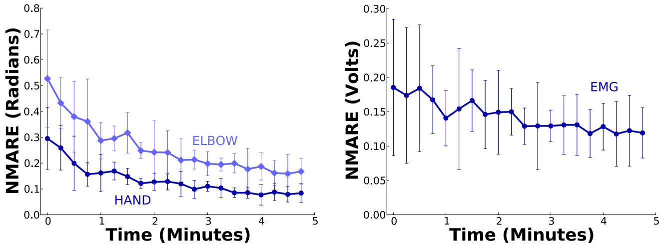

Fig. 5. **Learning performance during able-bodied subject trials.** Prediction learning curves are shown for elbow and hand actuator position (left), and for the average over all four myoelectric signals (right). Accuracy is reported in terms of normalized mean absolute return errors, averaged over eight trials.

or 3.0% of the maximum signal magnitude (5V). Prediction performance continued to improve over the course of an extended learning episode. These results are representative of the learning curves plotted for the other participants.

We also observed the adaptation of learned predictions in response to real-time changes. Perturbations to the task environment led to an immediate decrease in prediction accuracy, followed by a gradual recovery period. After a short learning period with one able-bodied subject, we transferred the four EMG electrodes to comparable locations on a second able-bodied subject. Within a period of less than 5min of use by the new subject the system had adapted its predictions about joint motion to reflect the new individual. Similar results were found for tasks involving gradual or sudden muscle fatigue; prediction accuracy was found to remain stable during periods of extended use (>60min of activity). Predictions were also found to recover their accuracy when a subject began holding a moderate weight with their controlling arm part way through a session. These observations reflect preliminary results, and further study is needed to verify the rate and amount of adaptation achievable in these situations.

Case Study 2: Amputee Control of a Robot Arm

Having demonstrated the applicability of GVFs for predicting and anticipating sensorimotor signals during able-bodied subject trials, we next assessed the ability of this approach to predict robot grip force, actuator position, actuator velocity, and other sensorimotor signals relating to an amputee's interactions with a robotic training prosthesis (Fig. 6).

Two trials were conducted with an amputee participant. The subject was a twenty-year-old male with a left transhumeral amputation, injured in a work accident sixteen months prior to the first trial. Six months prior to the first trial the subject underwent surgical revision of his limb, involving Targeted Muscle Reinnervation (TMR) as described by Dumanian et al. [11], as well as Targeted Sensory Reinnervation (TSR). The motor reinnervation procedure involved rerouting of the median nerve to innervate the medial biceps remnant muscle, and the distal radial nerve to the lateral triceps muscle, in order to provide additional myoelectric signal control sites for myoelectric prosthesis control. Sensory reinneration involved identifying specific median nerve fascicles with high sensory content, and coapting these fascicles to the intercostal brachial cutaneous sensory nerve. In a similar fashion, ulnar sensory

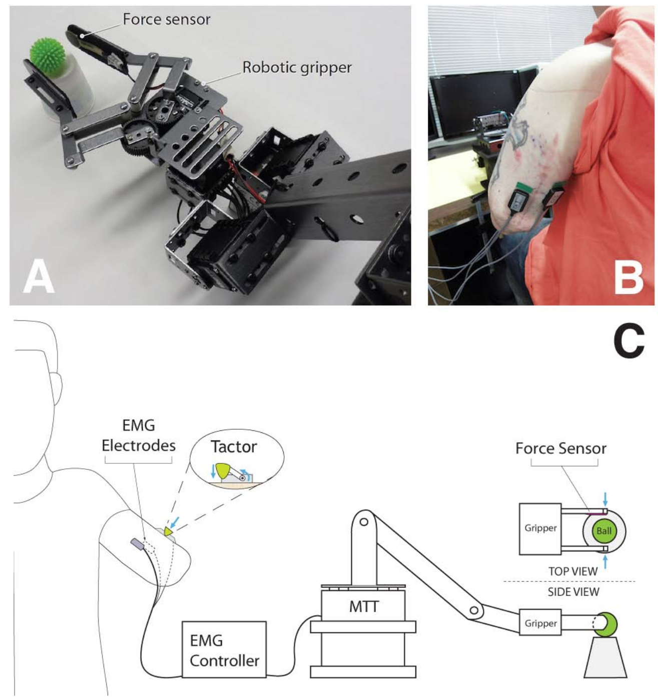

Fig. 6. **Experimental setup for amputee trials**: (a) multi-joint robot arm and force sensor, (b) participant with Bagnoli-8 EMG system, and (c) schematic of the Myoelectric Training Tool (MTT) and tactor feedback system.

fascicles with high sensory content were coapted to the sensory branch of the axillary nerve, and the remainder of the ulnar nerve trunk was rerouted to the motor branch of the brachialis muscle. Reinnervation resulted in a widely distributed discrete representation of digital sensation on the upper arm. A second trial occurred nine months after the first, wherein we observed noticeable changes to the subject's nerve reinnervation and improvements in his ability to operate a myoelectric device.

The robot platform used by this subject was the Myoelectric Training Tool (MTT), a clinical system designed to help new amputees prepare for powered prosthesis use [6]. The MTT includes a five-degree-of-freedom robot arm that mimics the functionality of commercial myoelectric prostheses. Signals from the robotic arm included the load, position, and velocity of each servomotor. A sensory feedback system was also used alongside the MTT in order to convey the sense of touch to the subject; robot grip force was measured and communicated to the subject via experimental tactors (micro servomotors) placed over the reinnervated skin of his upper arm. The subject controlled the robot arm via conventional myoelectric control with linear proportional mapping. Myoelectric signals from three of the subject's reinnervated muscles and two of the subject's native muscles were acquired, processed, and used to both control and switch between the various functions of the robot arm according to standard practice for the MTT [6].

During his two visits, the participant was asked to perform a variety of different actuation and sensation tasks with the MTT. Testing spanned multiple trials and multiple days, and included periods where the subject could control the MTT in an unstructured manner. Specific tasks included using both reinnervated motor control and sensory feedback to grip, manipulate, and discriminate between a series of small compressible objects in the absence of visual and auditory feedback. The subject also completed multiple trials of a modified clinical proficiency test known as *box-and-blocks*—a common procedure for assessing upper limb function. This test involved the amputee participant sequentially controlling four actuators on the robot arm to transfer a series of small objects between the two compartments of a divided workspace (Fig. 7). For this task, EMG signals from one pair of native muscles were used to actuate the robot's elbow, and signals from a reinneverated muscle pair were used to sequentially actuate the three remaining degrees-of-freedom on the robot arm (hand open/close, wrist flexion/extension, and shoulder rotation). Switching between joints was controlled using a voluntary EMG switch located on the subject's third reinnervated muscle group. For the box-and-blocks task, the subject was given normal visual and auditory feedback about the system, and could view displayed information about his EMG signals and currently selected robot actuators. More detail on the MTT box-and-blocks task can be found in work by Dawson, Fahimi, and Carey [6], and Pilarski et al. [19].

Data from these trials provided a rich sensorimotor space to evaluate our prediction learning methods. The data stream included EMG onsets ($\ll$1s in length), transient actuator motion and velocity readings ($<$1s) and periods of sustained motion or gripping actions (1–10s). We again created two GVFs for each signal of interest, one with a time scale of $\sim$0.67s and one with a time scale of $\sim$2.0s. Subsets of the available signals were given as input to the function approximation routine, as outlined above, and as described in related work [7], [19]. Learning updates occurred at 50Hz. All other parameters of the learning system, function approximation system, and computational hardware remained the same as in the previous case study. Using this configuration, the average computation time needed to update all GVFs and retrieve predictions at each time step was approximately 1ms.

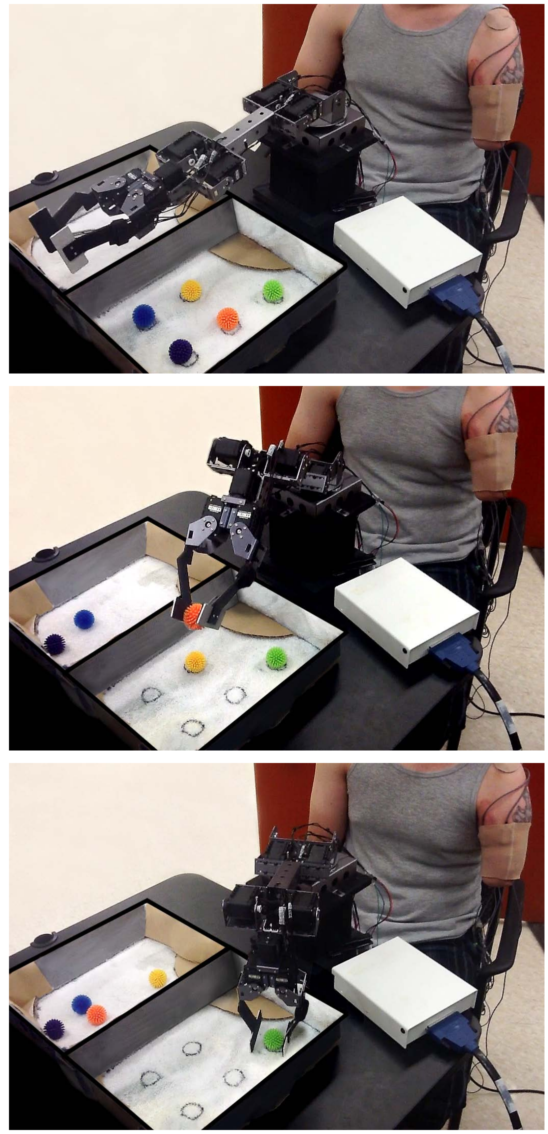

Fig. 7. **The modified box-and-blocks task**, wherein an amputee controls the multiple joints of the robot arm to move a series of small objects across the central barrier of a divided workspace.

Results

As shown in the results that follow, the system was able to successfully anticipate events initiated by the amputee subject, including joint angle changes, joint velocity changes, and grip force fluctuations. Accurate predictions were observed after 5–10min of real-time sampling and learning. To specifically evaluate GVF performance with respect to iterative offline learning, data was also recorded and divided into independent training and testing sets. Iterative training was found to further improve the accuracy of the system's predictions.

Figure 8 shows the accuracy and rate of improvement during offline learning for both hand actuator and grip force predictions on data from an object manipulation task. This is shown for the online learning case (a single pass through 16min of sensorimotor training data) and for 2–10 additional offline learning iterations. Accuracy was evaluated on both the training data and the previously unseen testing set. Each point in Fig. 8 represents the average NMARE over the training or testing data; results are shown for the 0.67s time scale (bottom trace) and the 2.0s time scale (top trace). As presented in Fig. 8a, iterative training produced a constant downward trend in joint angle prediction error (NMARE, with the angular signal presented in terms of servo control steps). The typical observed range of this actuator was $\sim$200 servo steps. Error on testing data after only one iteration—the online operation scenario—was found to be less than $7\ 0\%$ of this total range (Fig. 8a, right, first data point). The asymptotic error after multiple training passes was less than $4\ 0\%$. As shown in Fig. 8b, similar results were observed for grip force predictions. The typical range found for the grip force sensor was 0–1.5V; error on testing data after online training was less than $6\ 0\%$ of this range (Fig. 8b, right, first data point), with an asymptotic error after multiple training iterations of less than $4\ 5\%$. Online and asymptotic error is expected to further decrease with additional data, up to limits imposed by learning system generalization and the frequency of previously unseen events in the sensorimotor data.

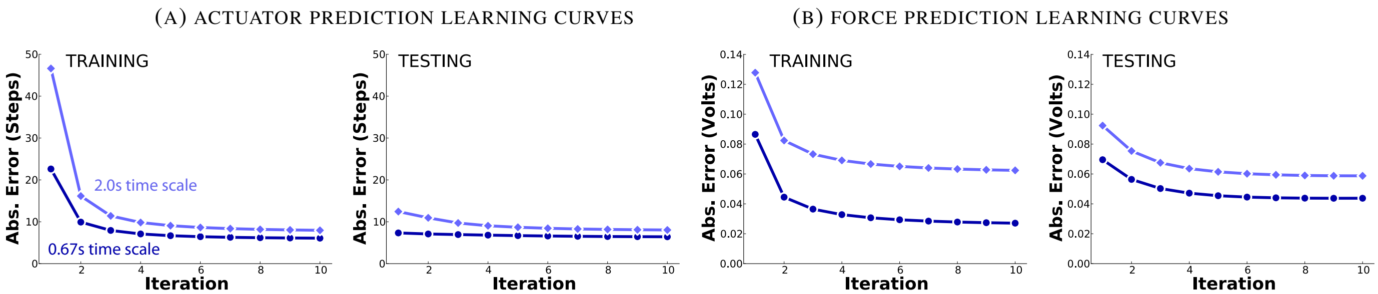

Fig. 8. **Learning performance on data from amputee-robot interaction.** Shown for (a) actuator joint angle prediction and (b) force prediction during a grasping task in terms of the average training and testing NMARE over ten learning iterations. *Bottom trace:* 0.67s time scale. *Top trace*: 2.0s time scale.

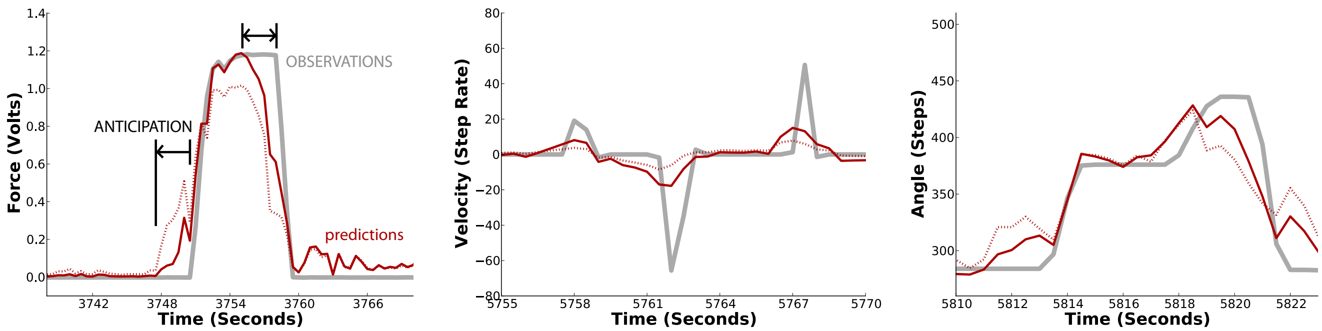

Fig. 9. **Prediction of grip force, actuator velocity, and actuator position signals from amputee testing data**. Shown after 3 learning iterations for time scales of 0.67s (solid lines) and 2.0s (dotted lines). Normalized return predictions (red lines) anticipate measured activity (grey lines) by 1–3 seconds.

Figure 9 shows a representative example of grip force, elbow velocity, and elbow joint angle predictions after three offline learning passes through recorded training data from object manipulation testing and the blocks-and-box task. Here the normalized predictions $\bar{P}_q$ at two time scales, 0.67s (solid red lines) and 2.0s (dotted red lines), are compared to the measured force, velocity, and positional sensor readings (grey lines); plotted data is binned into 25 timestep intervals. As was found in the able-bodied study, predicted signals anticipate changes to the actual measured signals by approximately 1–3s. Both rapid and slowly changing signals could be successfully predicted. The computational accuracy of these predictions was verified by comparing return predictions P_q to the true computed returns R_q, as done in the previous case study. The results in Fig. 9 are representative of our observations for the other signals and tasks in this domain.

As shown by the velocity prediction example in Fig. 9, the extended nature of exponentially discounted GVF predictions was found to affect the learning of events with a short duration relative to a GVF's time scale. At a prediction time scale of 0.67s, the rise and fall of transient velocity events could be accurately anticipated by more than 1s; however, normalized predictions were smoother than the measured signal, as they took into account velocity data before and after the depicted motion event. A time scale of 0.25s or less was found to effectively capture short-duration contours in our velocity data.

Discussion

Our case studies illustrate the potential of online prediction learning within the setting of myoelectric control. Using GVFs, it was possible to learn accurate temporally abstracted predictions about a human's interactions with a robotic device. This approach was also well suited to practical real-time implementation; learning updates and predictions were made under real-time computation constraints, and accurate predictions could be achieved within 5–10min of online learning. In addition, few application-specific changes and tuning operations were needed to shift between the different case studies and signal types presented in this work. Learning was performed in both an online setting (learning during ongoing experience) and an offline setting (learning from recorded data); offline and online predictions shared the same incremental learning framework. Together these results demonstrate the generality of the proposed approach—it would be straightforward to apply our methods to other assistive robotic domains.

The results in Fig. 9 highlight the relationship between prediction time scales and the temporal characteristics of sensorimotor events. By choosing prediction time scales appropriate to our events and questions of interest, we found that the learning system could accurately anticipate both slow and rapid changes within the sensorimotor stream. Predictions in this work were formed using an exponentially discounted weighting of future signals (i.e., a constant γ_q). Though not explored here, more complex methods for temporal extension are possible within the GVF framework [18].

Results from both case studies demonstrate the ability of a GVF learning approach to maintain and improve the accuracy of predictions over the course of online learning. Each new observation seen by the system was used to immediately refine and update existing predictive knowledge about the robot and its domain. With a suitable choice of function approximation methods, weight values learned for one situation could contribute to learning about previously unseen situations with similar state characteristics. One preliminary example was our ability to transfer electrodes between two able-bodied subjects in the middle of a learning session. The capacity to actively maintain prediction accuracy suggests a method for adapting a device to ongoing changes that occur during the daily life of an amputee—sweat, fatigue, altered use patterns, or physical changes to the assistive device and the positions of its myoelectric sensors. In essence, our methods enable a form of continuing domain adaptation. Online prediction learning may therefore provide a basis for rapidly adapting and calibrating factory-made devices for use by specific individuals.

While our approach performed well during structured and unstructured laboratory testing, its performance during complex real-life activities still needs to be demonstrated. Subtle changes to user intent and motor context may be difficult to discriminate in some real-life activities. As such, online learning during daily life may require additional external input signals (e.g., from the environment, the human, or the robot limb) in order to fully capture the context in which human actions are occurring. Our prediction and anticipation methods are not tied to a specific control approach, decoding scheme, or surgical setting; as such, they are suitable for both reinnervation patients and amputees using conventional myoelectric control. However, there may prove to be fundamental differences in the sensorimotor signals recorded in these different domains. Such differences could impact function approximation choices and the learning speed of the prediction methods. Further amputee studies are necessary to investigate these differences.

Prediction and Anticipation in Practice

There are several areas where the online prediction approach presented in this article promises to significantly improve amputee experiences with myoelectric control. Enhanced proprioceptive feedback to patients is one area where temporally extended predictions hold potential benefit. Proprioception is an important goal, as it allows an amputee to know where their prosthetic device is in space without visual confirmation; force and position feedback are cited as being requisites for patient acceptance [1]. With this objective in mind, tactile and auditory user feedback systems may benefit from incorporating temporally extended predictions about the motor consequences of a user's current commands. GVFs are able to provide context-dependent predictions of this type in a timely fashion.

A second area of impact for online prediction learning is enhanced failure forecasting. Grip slippage prevention is one representative example [1]. Advance knowledge of force, velocity, and position signals from a robotic prehensor could be used to estimate when a grasped object is about to unintentionally slip. The controller could then respond and prevent the failure by adjusting grip force and grip stiffness, or alert the user to the impending slip using tactile feedback. The use of foresight also allows a controller to avoid unintentional collisions between a prosthesis and its environment, and to anticipate mechanical, thermal, and electrical damage to the device. Sudden changes in prediction accuracy may also help detect and identify alterations in a system, for example transient changes to EMG signals, shifts in electrode positing, and other challenges to clinical robustness as identified by Scheme and Englehart [4].

A final area of utility for the approach presented in this article is facilitating the intuitive control of multifunction myoelectric devices. As discussed above, one of the major barriers to intuitive myoelectric control is a disparity between the number of EMG recording sites on an amputee's body and the control complexity of their artificial limb [2], [4]. This discrepancy between sensor and actuator spaces will only widen as limb technology becomes more advanced. GVFs provide a mechanism to improve multifunction control through the anticipation of a user's intent. Extended predictions about a user's control behaviour and their motor objectives can be used to prioritize control options for the user, modulate actuator stiffness and compliance, or directly coordinate natural timings for the simultaneous movement of multiple joints. The use of GVF predictions to streamline an amputee's control interactions in this way is the subject of ongoing studies by our group, and our preliminary results indicate tangible benefits for switching-based prosthetic control [19].

Final Thoughts: A Basis in the Brain

The view that online, adaptive predictions are important to control has an additional basis in human biology and motor learning [20]. As described by Flanagan et al. and Wolpert et al., predictions are learned by human subjects before they gain control competency [20], [21]. There is a strong relationship between sensorimotor prediction and control in the human brain, with anticipated future consequences being viewed as a fundamental component for generating and improving control [21]. Online prediction error, as represented by the human dopamine system, is also believed to play a key role in adaptively regulating behaviour [22], and has been successfully modelled by the same reinforcement learning algorithms we have used in this article [15]. Finally, it has been suggested that motor awareness and the feeling of "executing an action" stem from anticipatory predictions in the brain, as opposed to actual muscle movement [23]. These observations from human motor control further suggest that online prediction and anticipation could positively impact the functionality, intuitiveness, and feedback of multifunction myoelectric devices.

CONCLUSION

This article demonstrated an online approach to acquiring and continuously updating a set of predictions and anticipations about a human user and their assistive biomedical device. The ability to learn accurate, temporally abstracted predictions was shown through two case studies: able-bodied subjects engaging in the myoelectric control of a humanoid robot arm, and an amputee participant controlling an experimental myoelectric training robot. In both scenarios, it was possible for a real-time machine learning system to provide advance knowledge about signals such as actuator position, grip force, and myoelectric activity. Our approach was able to learn accurate predictions in both settings, and could do so without the need to significantly alter the learning system for each specific domain or signal of interest. To our knowledge, this work is the first demonstration of a practical method for online, real-time prediction learning in the context of myoelectric control. As such, our approach is a fundamental tool for addressing one major unsolved problem in the domain of artificial limbs: amputee-specific adaptation during ongoing use. The findings in this article provide a starting point for research into long-term control adaptation. They also contribute a first explicit look at online prediction learning as an important goal in its own right, independent of its integration within a specific controller, myoelectric decoder, or prosthetic device. Prediction and anticipation hold promise for improving myoelectric control. Future work will extend the present study to complex real-world activities, explore the use of prediction learning to supplement existing myoelectric control architectures, and comprehensively evaluate prediction-based control adaptation with a population of amputees.

Acknowledgments

The authors gratefully acknowledge support from the Alberta Innovates Centre for Machine Learning, the Natural Sciences and Engineering Research Council, Alberta Innovates – Technology Futures, and the Glenrose Rehabilitation Hospital. We also thank Joseph Modayil and Adam White for a number of very useful discussions, and Michael Stobbe for his technical assistance and the photographs used in Figure 1. Informed subject consent was acquired in both case studies as per ethics approval by the University of Alberta Health Research Ethics Board.

References

[1] B. Peerdeman, D. Boere, H. Witteveen, R. Huis in t Veld, H. Hermens, S. Stramigioli, H. Rietman, P. Veltink, and S. Misra, "Myoelectric forearm prostheses: State of the art from a user-centered perspective," *J. Rehabil. Res. Dev.*, vol. 48, no. 6, pp. 719–738, 2011.

[2] T. W. Williams, "Guest editorial: Progress on stabilizing and controlling powered upper-limb prostheses," *J. Rehabil. Res. Dev.*, vol. 48, no. 6, pp. ix–xix, 2011.

[3] S. Micera, J. Carpaneto, and S. Raspopovic, "Control of hand prostheses using peripheral information," *IEEE Rev. Biomed. Eng.*, vol. 3, pp. 48–68, 2010.

[4] E. Scheme and K. B. Englehart,"Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *J. Rehabil. Res. Dev.*, vol. 48, no. 6, pp. 643–660, 2011.

[5] J. Sensinger, B. Lock, and T. Kuiken, "Adaptive pattern recognition of myoelectric signals: exploration of conceptual framework and practical algorithms," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 17, no. 3, pp. 270–278, 2009.

[6] M. R. Dawson, F. Fahimi, and J. P. Carey, "The development of a myoelectric training tool for above-elbow amputees," *Open Biomed. Eng. J.*, vol. 6, pp. 5–15, 2012.

[7] P. M. Pilarski, M. R. Dawson, T. Degris, F. Fahimi, J. P. Carey, and R. S. Sutton, "Online human training of a myoelectric prosthesis controller via actor-critic reinforcement learning," in *Proc. 2011 IEEE Int. Conf. Rehabilitation Robotics (ICORR)*, Zurich, Switzerland, pp. 134–140.

[8] T. A. Kuiken, G. Li, B. A. Lock, R. D. Lipschutz, L. A. Miller, K. A. Stubblefield, and K. B. Englehart, "Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms," *JAMA*, vol. 301, no. 6, pp. 619–628, 2009.

[9] L. Resnik, M. R. Meucci, S. Lieberman-Klinger, C. Fantini, D. L. Kelty, R. Disla, and N. Sasson, "Advanced upper limb prosthetic devices: implications for upper limb prosthetic rehabilitation," *Arch. Phys. Med. Rehabil.*, vol. 93, no. 4, pp. 710–717, 2012.

[10] M. S. Johannes, J. D. Bigelow, J. M. Burck, S. D. Harshbarger, M. V. Kozlowski, and T. Van Doren, "An overview of the developmental process for the modular prosthetic limb," *Johns Hopkins APL Tech. Dig.*, vol. 30, no. 3, pp. 207–216, 2011.

[11] G. A. Dumanian, J. H. Ko, K. D. O'Shaughnessy, P. S. Kim, C. J. Wilson, and T. A. Kuiken, "Targeted reinnervation for transhumeral amputees: current surgical technique and update on results," *Plast. Reconstr. Surg.*, vol. 124, no. 3, pp. 863–869, 2009.

[12] D. Nishikawa, W. Yu, H. Yokoi, and Y. Kakazu, "On-line learning method for EMG prosthetic hand control," *Electron. Comm. Jpn. Pt. III*, vol. 84, no. 10, pp. 35–46, 2001.

[13] C. Pulliam, J. Lambrecht, and R. F. Kirsch, "Electromyogram-based neural network control of transhumeral prostheses," *J. Rehabil. Res. Dev.*, vol. 48, no. 6, pp. 739–754, 2011.

[14] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press, 1998.

[15] B. Seymour, J. P. O'Doherty, P. Dayan, M. Koltzenburg, A. K. Jones, R. J. Dolan, K. J. Friston, and R. S. Frackowiak, "Temporal difference models describe higher-order learning in humans," *Nature*, vol. 429, no. 6992, pp. 664–667, 2004.

[16] R. S. Sutton, J. Modayil, M. Delp, T. Degris, P. M. Pilarski, A. White, and D. Precup, "Horde: a scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction," in *Proc. 10th Int. Conf. Autonomous Agents and Multiagent Systems (AAMAS)*, Taipei, Taiwan, 2011, pp. 761–768.

[17] R. S. Sutton, "Learning to predict by the methods of temporal differences," *Mach. Learn.*, vol. 3, no. 1, pp. 9–44, 1988.

[18] J. Modayil, A. White, and R. S. Sutton, "Multi-timescale nexting in a reinforcement learning robot," in *Proc. Int. Conf. Simulation of Adaptive Behaviour (SAB)*, Odense, Denmark, 2012, pp. 299–309.

[19] P. M. Pilarski, M. R. Dawson, T. Degris, J. P. Carey, and R. S. Sutton,"Dynamic switching and real-time machine learning for improved human Control of assistive biomedical robots," in *Proc. 4th IEEE RAS & EMBS Int. Conf. Biomedical Robotics and Biomechatronics (BioRob)*, Roma, Italy, 2012, pp. 296–302.

[20] D. M. Wolpert, Z. Ghahramani, and J. R. Flanagan, "Perspectives and problems in motor learning," *Trends Cogn. Sci.*, vol. 5, no. 11, pp. 487–494, 2001.

[21] J. R. Flanagan, P. Vetter, R. S. Johansson, and D. M. Wolpert, "Prediction precedes control in motor learning," *Current Biology*, vol. 13, no. 2, pp. 146–150, 2003.

[22] J. Zacks, C. Kurby, M. Eisenberg, and N. Haroutunian, "Prediction error associated with the perceptual segmentation of naturalistic events," *J. Cogn. Neurosci.*, vol. 23, no. 12, pp. 4057–4066, 2011.

[23] M. Desmurget, K. Reilly, N. Richard, A. Szathmari, C. Mottolese, and A. Sirigu, "Movement intention after parietal cortex stimulation in humans," *Science*, vol. 324, no. 5928, pp. 811–813, 2009.

Authors

Patrick M. Pilarski, Department of Computing Science, University of Alberta, Edmonton, AB, Canada. E-mail: pilarski@ualberta.ca.
* *Corresponding author.*

Michael Rory Dawson, Glenrose Rehabilitation Hospital, Edmonton, AB, Canada. E-mail: mrd1@ualberta.ca

Thomas Degris, INRIA Bordeaux Sud-Ouest, Talence Cedex, France. E-mail: thomas.degris@inria.fr

Jason P. Carey, Department of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada. E-mail: jason.carey@ualberta.ca

K. Ming Chan, Division of Physical Medicine & Rehabilitation, University of Alberta, Edmonton, AB, Canada. E-mail: kming@ualberta.ca

Jacqueline S. Hebert, Division of Physical Medicine & Rehabilitation, University of Alberta, Edmonton, AB, Canada. E-mail: jhebert@ualberta.ca

Richard S. Sutton, Department of Computing Science, University of Alberta, Edmonton, AB, Canada. E-mail: rsutton@ualberta.ca