# Dynamic Switching and Real-time Machine Learning for Improved Human Control of Assistive Biomedical Robots

Patrick M. Pilarski, Michael R. Dawson, Thomas Degris, Jason P. Carey, and Richard S. Sutton

*Abstract*— **A general problem for human-machine interaction occurs when a machine's controllable dimensions outnumber the control channels available to its human user. In this work, we examine one prominent example of this problem: amputee switching between the multiple functions of a powered artificial limb. We propose a dynamic switching approach that learns during ongoing interaction to anticipate user behaviour, thereby presenting the most effective control option for a given context or task. Switching predictions are learned in real time using temporal difference methods and reinforcement learning, and demonstrated within the context of a robotic arm and a multi-function myoelectric controller. We find that a learned, dynamic switching order is able to out-perform the best fixed (non-adaptive) switching regime on a standard prosthetic proficiency task, increasing the number of optimal switching suggestions by 23%, and decreasing the expected transition time between degrees of freedom by more than 14%. These preliminary results indicate that real-time machine learning, specifically online prediction and anticipation, may be an important tool for developing more robust and intuitive controllers for assistive biomedical robots. We expect these techniques will transfer well to near-term use by patients. Future work will describe clinical testing of this approach with a population of amputee patients.**

## I. INTRODUCTION

Within the context of human-robot interaction, it is often the case that the number of controllable degrees of freedom within a robot will dramatically exceed the number of control channels that can be easily manipulated by a human user. In such situations, it is challenging to form a link between the human and the robot that enables high levels of robot functionality while simultaneously providing an intuitive, learnable control scheme for the human user. In other words, effective control is frequently confounded by a disparity between the size or nature of the signal space on both sides of the human-machine interface.

This issue is especially prominent within the domain of assistive biomedical robots and rehabilitation robotics. In these systems, the human user is often tightly coupled to the robotic device by way of sensors or switches affixed directly to the subject's body. Approaches include internal or surface electromyography (EMG), Electroencephalography (EEG), force measurement, and direct neural recording. In each case, activity from the human body results in a set of electrical signals which can be processed and directed toward the control of an electromechanical assistive device. The more reduced the functioning of the user, the more limited are the possible sites to record these signals and the greater the required functionality for the assistive robot.

Powered myoelectric prostheses are a prototypical example of this situation [1]–[3]. In a myoelectric prosthesis, surface EMG signals are recorded from muscle tissue in an amputee's residual limb. These signals are interpreted to form control commands for a robotic appendage with one or more degrees of freedom (DoF). In some cases, information about the activity of the limb is also returned to the user by way of auditory, visual, or tactile feedback [3]. Despite many advances in novel pattern-recognition-based control schemes (e.g., see reviews by Oskoei and Hu [4] and Scheme and Englehart [5]), conventional myoelectric controllers typically control a single degree of freedom with a single residual muscle pair [5]. Unfortunately, as the amputation level increases, the number of muscle sites available for use as input signals to control schemes decreases [1], [2], [5].

In order to increase the number of DoF that an amputee can control, conventional controllers are often extended using a voluntary switch that allows the user to cycle their control through the available degrees of freedom on the prosthesis [1], [5]. For example, an upper-arm (transhumeral) amputee may need to shift their control between the wrist, elbow, or hand joints of their prosthesis; a lower-arm amputee might choose between the different grip styles available via their dexterous hand prosthesis. Switching of this kind may be implemented using mechanical toggles, linear displacement transducers, force sensitive resistors, co-contraction of two EMG channels, or a third dedicated EMG channel [2], [3]. Examples of commercial systems that employ switching-based control options include the bebionic hand (RSLSteeper), the Boston Digital Arm System (Liberating Technologies, Inc.), the UtahArm3+ (Motion Control, Inc.), the DynamicArm and ErgoArm (Otto Bock HealthCare), and the i-limb (Touch Bionics, Inc.).

Despite the promise of increased functionality, standard switching-based methods are acknowledged to be unacceptably slow and difficult to use [5]; patients have reported a number of barriers to their uptake of commercially available multifunction prostheses. As reported in a recent needs assessment study (Peerdeman et al., 2011), these concerns include the lack of functionality and non-intuitive nature of most conventional control schemes [6]. Ease of selecting different limb movements and grasp types has been identified as an important functional requirement for patient acceptance.

While new surgeries like targeted reinnervation are ex-

P. M. Pilarski and R. S. Sutton are with the Department of Computing Science, University of Alberta, Edmonton, Alberta, T6G 2E8, Canada; M. R. Dawson is with the Glenrose Rehabilitation Hospital, Edmonton, Alberta, T5G 0B7, Canada; T. Degris is with INRIA National Labs, Bordeaux, France; J. P. Carey is with the Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, T6G 2G8, Canada.
*Please direct correspondence to:* pilarski@ualberta.ca

Preprint submitted to Fourth IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics. Received January 27, 2012.

panding the potential for intuitive multi-function control through the creation of new physiologically mapped EMG recording sites [7], the majority of amputees still require some form of manual or multi-state switching to control the different degrees of freedom or grip styles of their artificial limb or hand [3], [5]. Even in the case of expert-designed switching schemes, switching patterns are normally set in-clinic and do not adapt to ongoing changes in the lifestyle, use patterns, or physiology of the amputee. Standard switching-based methods are acknowledged to be unacceptably slow and difficult to use [5]. In light of these limitations, more natural, adaptive approaches are now seen as a requirement for effective multifunction control [5].

With this in mind, the present work employs real-time pattern recognition and machine learning techniques to help remove switching-related barriers to patient acceptance and control. In order to reduce the switching time and number of switching actions required to actuate and smoothly control a multifunction prosthesis, we propose the use of online machine learning to anticipate and predict human control signals during their interaction with a myoelectric device. A key intuition is that there is a great degree of regularity in the movements that comprise many daily tasks; we demonstrate how it is possible to learn these regularities in terms of the control actions a user deploys during task completion. By accurately predicting the next DoF or mode to be actuated by a user in any given context, it is possible to present these functions as first choices to the user at the point of manual switching. This will enable control systems that move beyond fixed switching patterns to adaptively prioritize control options in terms of their immediate utility, streamlining the interface between a human and their assistive robot.

## II. REAL-TIME PREDICTION

Prediction is a key component of most advanced pattern recognition systems intended for use in myoelectric control [4], [5]. Being able to accurately predict limb movements, user intent, or grip type from an ongoing stream of patient control signals is integral to the classifications made by most pattern recognition approaches. Notable examples include linear discriminant analysis (LDA), support vector machines (SVM), artificial neural networks (ANN), and principal component analysis (PCA). Scheme and Englehart provide a comprehensive review of the state of the art in this domain [5]. However, while most myoelectric pattern recognition systems are capable of providing control decisions for realistic online operation (e.g., every 250ms or less), few have the ability to *continue* to learn online—i.e., to adapt their predictions during ongoing, real-time operation. Online adaptation is closely tied to the idea of robustness, and acknowledged to be a main focus for future progress in the domain of myoelectric control [5], [8].

One machine learning approach that has demonstrated its practical applicability in an ongoing, incremental prediction and control setting is *reinforcement learning* (RL) [9]. In RL, a learning system uses samples of interactions with its environments to build up expectations about future values of a scalar feedback signal termed *reward*. These expectations are represented using a *value function* that maps observations in the environment to predictions about future reward.

Recently, Sutton et al. proposed the idea of a *general value function* (GVF) [10]. This construct retains the key properties of RL, but extends the idea of the value function to represent predictive knowledge about arbitrary (i.e., non-reward) signals and observations. In essence, these GVFs and their associated learning techniques can be thought of as a method for asking and answering predictive questions about future sensorimotor signals. Moreover, GVFs can be learned in an incremental fashion using standard RL methods [9], with linear per-time-step complexity in both memory and computation. They are therefore well suited to real-time anticipatory predictions about user control signals and function switching behaviour.

Specific uses for GVFs include learning to predict the ongoing magnitude of a signal given the current state of the system (e.g., "what will be the average value of a particular sensor during the next two minutes") or predicting the magnitude of a sensor reading conditional on some distinct event (e.g., "given my current state, what will be the value of a particular sensor in exactly thirty seconds"). The questions posed by GVFs can also be temporally extended; that is, they can be predicted outcome measures with different time scales (e.g., milliseconds, seconds, minutes, hours) and with different ways of weighting the importance of future signals. Predictions can also depend on different policies or strategies for choosing control actions [10].

In the work that follows, we describe the application of GVFs to the problem of anticipating control switching events that occur during interactions between a human user and a biomedical assistive robot.

## III. EXPERIMENTAL SETUP

### A. Robotic Platform

The robotics platform used in this work was the *Myoelectric Training Tool* (MTT, shown in Fig. 1), a multi-function robot arm designed to prepare new amputees for the control of a commercial powered prosthesis [11]. Based on the AX-12 Smart Arm (Crustcrawler, Inc.), the MTT includes five degrees of freedom that mimic the functionality of commercial myoelectric prostheses. For this work, we used four of these actuators: hand open/close, wrist flexion/extension, elbow flexion/extension, and shoulder rotation. Feedback signals from the robotic arm were sampled at a rate of 50Hz, and consisted of the load (current), angular position, temperature, voltage, and angular velocity of each AX-12+ servomotor. The following information was also recorded for each actuator: an on/off signal indicating whether or not the joint was currently controlled by the user (*joint control state*), an on/off signal indicating whether or not the actuator was currently in motion (*joint activity*), and a time-decayed trace of the joint activity signal (*joint activity trace*; decay constant $\tau = 0.01$). Grip strength was measured via a force sensitive resistor bonded to the tip of the gripping device.

Preprint submitted to Fourth IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics. Received January 27, 2012.

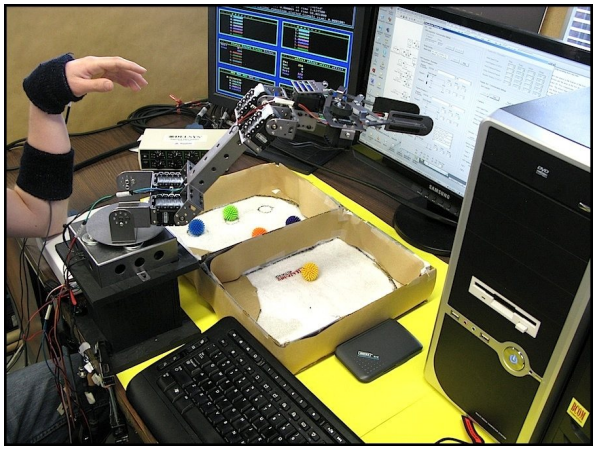

Fig. 1. **Able-bodied subject interacting interacting with the Myoelectric Training Tool (MTT);** experimental setup also includes a Bagnoli 8-channel EMG system, real-time control computer, and task workspace.

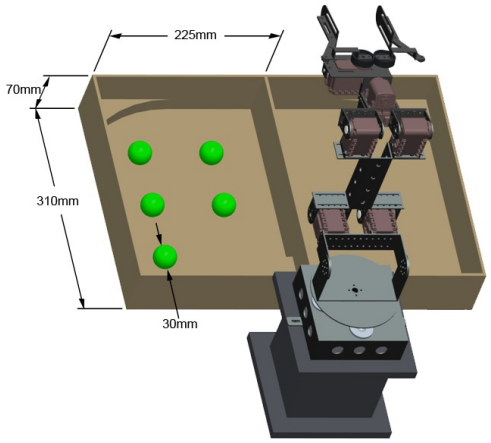


Fig. 2. **Schematic of the MTT and modified "box and blocks" test area.** Each trial involves controlling the MTT to transfer the five balls from the left compartment to the right compartment as quickly as possible.

### B. Myoelectric Signal Acquisition and Control

EMG signals were acquired at 2kHz via a Bagnoli 8-Channel EMG system (DelSys, Inc.). Signals were digitized into a real-time control computer using a national instruments PCI-6259 data acquisition card. The EMG electrodes for the control channel were affixed over the left forearm extensors and flexors of an able-bodied subject, while the electrode for the EMG switching channel was placed over the subject's right forearm extensor; a reference electrode was affixed to the bone of the subject's left wrist. The hardware gain for all three channels was set to 1000 and the software gains for the hand open and close channels were 250 and 400 respectively. Mean absolute values (MAV) for all three EMG signals were recorded alongside the MTT data at every 50Hz robot sampling interval. A linear-proportional mapping, as used in conventional myoelectric controllers (see Parker et al., [2]), was then used to map the signal strength of the two control EMG signals to the angular velocity of a controlled servomotor on the MTT. The activity of the third EMG channel was used as a binary switch to manually cycle control between the four degrees of freedom on the MTT.

### C. The Modified Box & Blocks Task

To provide a rich stream of interactive data, we enlisted an able-bodied subject in a modified version of the standard clinical prosthesis proficiency test known as "box and blocks" [11], [12]. As shown in Figs. 1 and 2, the task environment consisted of a divided box with five coloured balls pre-located on the left side of a central barrier. The subject's objective on each trial was to move all five balls from the left box compartment to the right box compartment in as little time as possible. This requires the subject to repeatedly switch their control between the MTT's different actuators (Fig. 3). The subject controlled the robotic arm using the conventional EMG controller described above, where the signal strength of the left forearm extensor and flexor muscles were used to proportionally control the angular velocity of a given joint on the robotic arm. In order to control more than one DoF on the robotic arm, the subject also had a third electrode placed on the extensor of their right arm that they could activate in an on/off momentary fashion. This allowed them to cycle in one direction through a pre-defined switching list. Each time the subject switched to the next DoF in the list, an audio chime was sounded and a visual cue was displayed. This provided feedback to the user indicating that the switch was successfully activated (audio cue) and the DoF to which they had switched (visual cue). The switching order was chosen heuristically to be elbow flexion/extension, wrist flexion/extension, hand open/close, and shoulder rotation. As described by Dawson et al., this order was found to minimize the number of switching actions required for this particular task [11]. One or more manual switching actions were needed per change of actuator, resulting in a typical switching period between movements that lasted 1–3 seconds. The subject performed ten trials with informed consent and ethics approval by the Health Research Ethics board of the University of Alberta; the typical length for one trial episode was $\sim$2 minutes.

## IV. MACHINE LEARNING METHODS

### A. Online Learning Algorithm

As described above, general value functions are a versatile computational method for representing predictions about future signals and observations. For this work, we utilized GVFs to anticipate the control switching events that occur during interactions between the MTT and its human user. GVFs were formulated as described by Sutton et al. [10], and their values were learned in an online fashion following the temporal difference (TD) learning procedure presented as Algorithm 1. A step-by-step breakdown of Algorithm 1 follows; for additional detail on TD learning and the TD($\lambda$) algorithm, please see Sutton and Barto, [9].

At the start of the experiment, one GVF weight vector $\mathbf{w}_j$ was initialized for each joint $j$ in the target system (i.e., four GVFs, one for each of the shoulder, elbow, wrist, and hand servo motors). Over the course of learning, each GVF was

Preprint submitted to Fourth IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics. Received January 27, 2012.

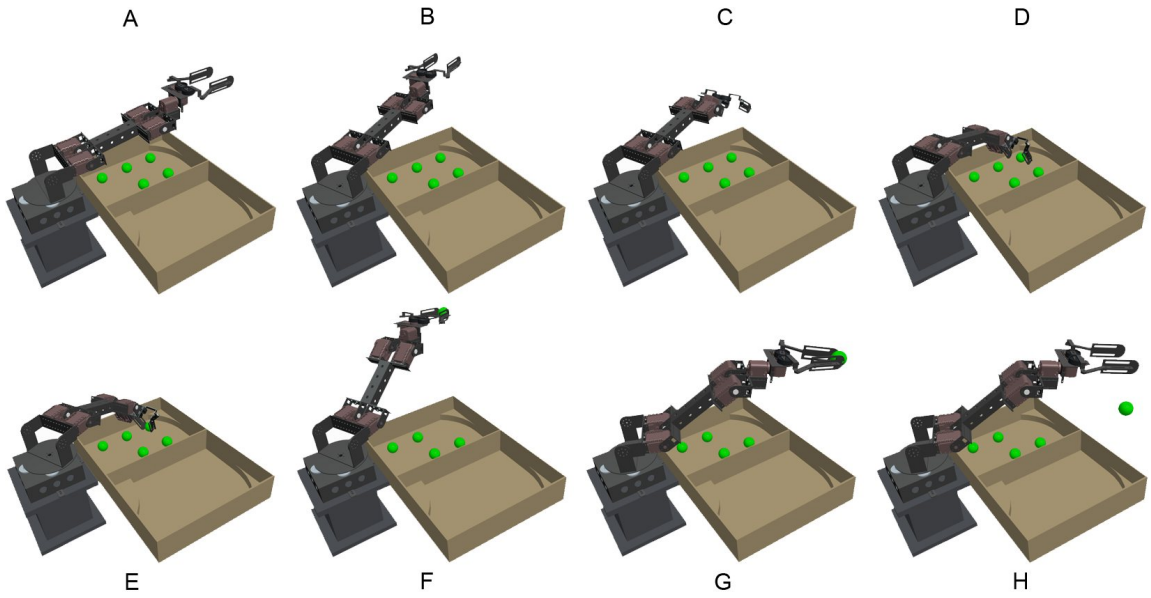


Fig. 3. **A sequence of actions from the modified box and blocks task.** For this example, the changes in joint control needed for moving a single ball to the target region include shoulder rotation (A→B), wrist flexion (B→C), elbow flexion (C→D), hand grip (D→E), elbow extension (E→F), shoulder rotation (F→G), and hand release (G→H). Manual function switching between each motion typically lasts 1–3 seconds or more.

updated according to the value of an instantaneous signal of interest $r_j$. Here $r_j$ was the joint activity signal indicating whether the associated actuator $j$ was moving ($r_j = 1$) or stationary ($r_j = 0$) on the current time step. The learned GVFs therefore represented a set of temporally extended expectations of these instantaneous motion signals—in effect, predictions about the next joint to be moved by the user.

Each step of the learning process proceeded as follows. On every time step (i.e., approximately every 20ms), the system received a new set of observations, denoted '$\mathbf{s}$', from the MTT and associated EMG system, along with the four signals of interest ($r_j$); $\mathbf{s}$ included all 32 signals shown in Table I. To facilitate efficient computation, $\mathbf{s}$ was processed using *linear function approximation*—a mathematical approach which allows a large, continuous state space or observation space to be represented in a linear form that can be readily used in computation and learning [9]. In this case, an approximation routine $\mathbf{x}' \leftarrow \mathsf{approx}(\mathbf{s})$ was used to map $\mathbf{s}$ into a binary feature vector $\mathbf{x}'$. For clarity, the details of this approximation method are presented in Sec. IV-B, below.

The weight vectors $\mathbf{w}_j$ for each GVF were then updated using this new information ($\mathbf{x}'$ and $r_j$). For each joint $j$, a temporal difference error signal (denoted $\delta$) was formed using the joint's current activity signal $r_j$ and the difference between the current and future predicted values for this signal (computed from the weight vector using the linear combinations $\mathbf{w}_j^T\mathbf{x}$ and $\gamma\mathbf{w}_j^T\mathbf{x}'$, respectively). Next, a trace $\mathbf{e}_j$ of the current feature vector was updated in a replacing fashion, where $\mathbf{e}_j$ was an eligibility trace vector with $\lambda$ as the corresponding eligibility trace decay rate. For more detail on replacing eligibility traces, please refer to Singh and Sutton [13] and Sutton and Barto [9]. This trace $\mathbf{e}_j$ was used alongside the error signal $\delta$ to update the weight vector $\mathbf{w}_j$ for each GVF. Here $\alpha > 0$ was a scalar step-size parameter.

**Algorithm 1** Learning General Value Functions with TD($\lambda$)

**initialize:** $\mathbf{w}, \mathbf{e}, \mathbf{s}, \mathbf{x}$
**repeat:**
  **observe** $\mathbf{s}$
  $\mathbf{x}' \leftarrow \mathsf{approx}(\mathbf{s})$
  **for** all joints $j$ **do**
    **observe** joint activity signal $r_j$
    $\delta \leftarrow r_j + \gamma\mathbf{w}_j^T\mathbf{x}' - \mathbf{w}_j^T\mathbf{x}$
    $\mathbf{e}_j \leftarrow \mathsf{min}(\lambda\mathbf{e}_j + \mathbf{x}, 1)$
    $\mathbf{w}_j \leftarrow \mathbf{w}_j + \alpha\delta\mathbf{e}_j$
    $\mathbf{x} \leftarrow \mathbf{x}'$

The prediction of future joint activity $p_j$ at any given time is sampled using the linear combination: $p_j \leftarrow \mathbf{w}_j^T\mathbf{x}$

Learned predictions $p_j$ for each given signal $r_j$ were sampled from the weight vector $\mathbf{w}_j$ and current feature vector $\mathbf{x}$ according to $p_j = \mathbf{w}_j^T\mathbf{x}$. Ranking the magnitude of these predictions at a given instant produced a context-dependent (dynamic) switching order.

A discounting factor of $\gamma = 0.992$ was used to compute $\delta$; each prediction $p_j$ therefore represented a summed exponentially decayed outcome with a half-life of $\sim 125$ time steps, or roughly 2.5s of data. Parameters used in the learning process were $\lambda = 0.999$ and $\alpha = 0.3/n$, where $n$ is the number of active features in the state approximation vector $\mathbf{x}$; vectors $\mathbf{e}_j$ and $\mathbf{w}_j$ were all initialized to 0.

### *B. Function Approximation*

The linear approximation method $\mathsf{approx}(\mathbf{s})$ was implemented using a tile coding function, as per Sutton and Barto [9]. This mapped the 32-dimensional real-valued signal space $\mathbf{s}$ into multiple five-dimensional binary arrays with

TABLE I
FIVE-DIMENSIONAL SIGNAL COMBINATIONS USED IN FUNCTION APPROXIMATION

**(A) Array Dimensions 1 through 4**

| | |
|---|---|
| Shoulder Servo Position | Elbow Servo Position |
| Wrist Servo Position | Hand Servo Position |

**(B) Array Dimension 5**

*One of:*

| | |
|---|---|
| ShoulderServoVelocity | ShoulderServoLoad |
| ShoulderServoVoltage | ShoulderServoTemperature |
| ElbowServoVelocity | ElbowServoLoad |
| ElbowServoVoltage | ElbowServoTemperature |
| WristServoVelocity | WristServoLoad |
| WristServoVoltage | WristServoTemperature |
| HandServoVelocity | HandServoLoad |
| HandServoVoltage | HandServoTemperature |
| HandForceSensor | EmgSwitchMav |
| Emg1Mav | Emg2Mav |
| HandControlState | WristControlState |
| ElbowControlState | ShoulderControlState |
| HandActivityTrace | WristActivityTrace |
| ElbowActivityTrace | ShoulderActivityTrace |

a discretization level of six bins per array axis. As shown in Table IA, the four primary axes of each array were the position of the shoulder, elbow, wrist, and hand servo; the fifth axis of each array was one of the remaining 28 signals in Table IB. One array was constructed for each signal in Table IB, leading to a total of a total of 28 arrays. This process was done for six overlapping tilings, each shifted from the origin by a random amount. The result was concatenated with one active baseline unit into a single binary vector consisting of 1,306,369 features; exactly $n = 169$ features were active at any given time, one for each tiling of each array, and one for the baseline feature. All signals in Table I were normalized between 0.0 and 1.0 according to their known ranges.

### C. Evaluation Methods

Ten trials of the modified box and blocks task were recorded, each lasting approximately two minutes (∼6k time steps). Ten fold cross-validation was then performed on the recorded data, with the learning system being trained incrementally on nine of the trials and then tested for its prediction accuracy on previously unseen data from the tenth hold-out trial. This process was repeated with each of the ten possible hold-out trials. We also assessed the efficacy of learning from multiple passes through the training data, whereby prediction accuracy on the testing data was evaluated following two or more iterations though the testing data. Learning was subject to real-time computation constraints, with the requirement that all learning and prediction related computation for each time step be completed in 20ms or less; in practice, all learning related computation for eight

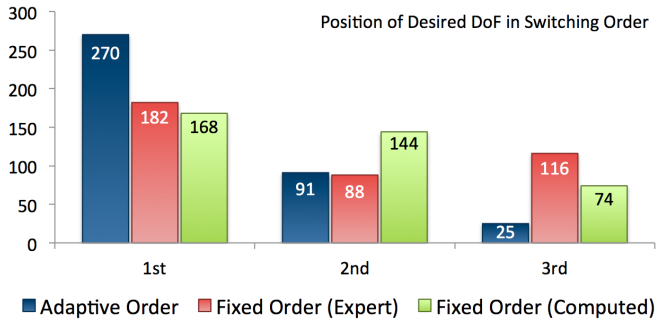


Fig. 4. **Key result: learned predictions correctly anticipate the next desired DoF to be selected by the user.** The dynamic switching order generated by the machine learning system (blue) is compared to the expert-designed switching cycle (red) and the best fixed order as computed post-hoc for this data (green). Bars show the number of instances over all ten testing periods in which the next DoF moved by the user was ranked first, second, third in a switching cycle at the termination of the previous movement.

simultaneous predictions was completed within 1–5ms, with processing being done on a standard MacBook Pro 2.53 GHz Intel Core 2 Duo laptop. For evaluation purposes, true prediction values $p_j^*$ were computed post-hoc by calculating the return observed by summing the exponentially discounted signal value $r_j$ at each time step in the future. To allow for visual timestep-by-timestep comparison of predictions and observed signals on the same vertical axis, predictions were also evaluated in temporally normalized form, with values scaled according according to $p_j/(1-\gamma)$.

## V. RESULTS & DISCUSSION

A key result for this work is that the learning system was able to correctly learn in real time to anticipate the next joint to be actuated by the human subject. This result is shown in Fig. 4. Here vertical blue bars show the number of instances over all ten testing periods in which the next DoF moved by the user was ranked first, second, third by the machine learning system in terms of the absolute magnitude of predictions at the moment the user's previous movement was terminated, with the currently activated DoF being placed last in the ranking. In effect, at the termination of each movement, the system formed a new, dynamic switching order based on the current information being observed from the experimental environment and its predictions about the next function to be used by the subject.

In 70% of the 386 total switching cases recorded during the ten testing trials, the DoF selected by the user was ranked first in the dynamic switching order generated by the machine learning system; this is 23% higher than was observed for the fixed, expert-designed switching scheme described in Sec. III-C. The average switching time observed during this experiment was 1.09s, 1.75s, and 2.21s for transitions involving one, two and three switching actions respectively. Using these values as a measure of switching cost gives an expected transition time of 1.32s for the adaptive switching order, as compared to 1.57s for the fixed, expert-designed switching order (Fig. 4, red bars), and 1.55s for the best possible fixed order, as computed post-hoc for this data (Fig.

Preprint submitted to Fourth IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics. Received January 27, 2012.

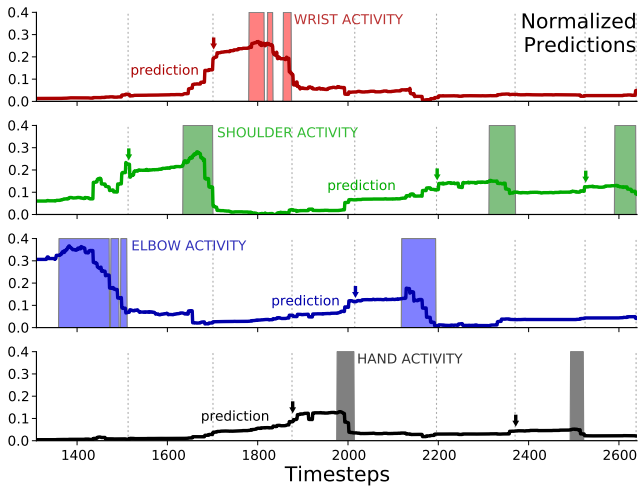


Fig. 5. **Example of learned joint activity predictions as compared to actual joint activity** on previously unseen testing data. Solid lines indicate the normalized magnitude of the predictions $p_j$ for the future movement of each joint, and filled curves show the joint activity signal $r_j$ being predicted by the learning system. Arrows indicate the dominant predictions at the termination of each movement; dominant predictions accurately anticipate the future joint actuation. Shown for elbow (blue), shoulder (green), wrist (red), and hand (black) activity.

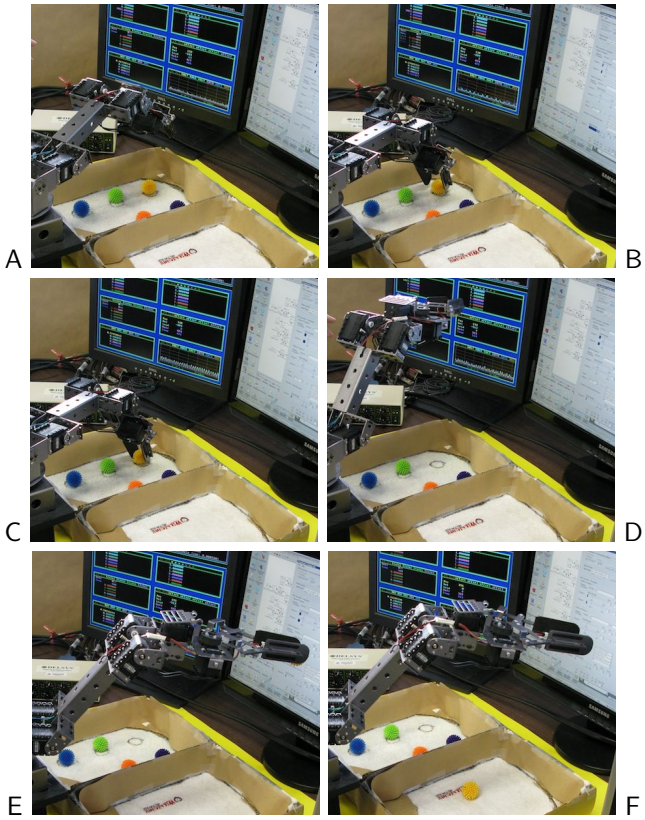


Fig. 6. **Example of a user-directed control sequence** corresponding to the joint activity progression presented in Fig. 5. However, in this sequence, the subject chooses to use the elbow actuator (A→B) instead of the wrist actuator (Fig. 5, step 1800) when approaching the yellow ball.

4, green bars). In terms of its potential for decreasing DoF transition times, the dynamic switching order produced a cumulative transition cost decrease of 1.49min as compared to the cost of the best computed switching order—a potential savings of 7.2% with respect to the total experiment time (20.7min), or 14.3% of the total transition time (10.4min). When compared to the expert-designed approach, dynamic ordering led to a potential decrease of 16.0% (1.67min) in total transition time. This advantage is expected to increase as the number of DoF in the switching order increases.

Figure 5 presents an example of the learning system's predictions for a subset of the testing data after five learning iterations. Solid lines indicate the normalized magnitude of the predictions $p_j$ for the future movement of each joint, and filled curves show the joint activity signal $r_j$ being predicted by the learning system. As shown in Fig. 5, predictions for the next joint to be actuated increase sharply at, or just prior to, the termination of the previous movement. Arrows show the dominant prediction in the dynamic switching list at the termination of each user movement. For reference, images from a similar sequence of user actions is shown in Fig. 6.

Of additional note in Fig. 5, predictions for subsequent future DoF can also be seen to rise in magnitude prior to actuation. This promises to help optimize user switching in the case of shifted intent or task changes. For example, a user would normally switch to elbow control after gripping a ball (Fig. 5, step 2000, and Fig. 6C→D); however, if the middle barrier between compartments was suddenly removed, the system's increased prediction for shoulder activity would allow them to skip to shoulder control with only one additional switching action. The learning system predicts correctly that shoulder activity will follow elbow activity in this context.

Performance did not rely on the user performing a perfectly repetitive task; the learning system was able to utilize regularities in the task environment while remaining robust to novel deviations. For example, we consider the case of a dropped (fumbled) ball, or a user's choice of an alternate approach to achieve their task—e.g., the difference between using the wrist joint (Fig. 5, step 1800) or elbow joint (Fig. 6A→B) prior to gripper actuation. These cases were common in the testing data; experience with these situations and new information from the environment (e.g., grip force) allowed the system to re-order its predictions, and thus the suggested switching order, to help match the user's new goals.

Generalization and learned adaptation was enabled in part by the use of function approximation to create the feature vector $\mathbf{x}$. By expanding the initial set of 32 observations into a well structured approximation space, the system was able to represent rich interconnections and non-linear relationships between observed signals. This approximation could then be used to learn good state-related predictions. These results hold with our recent work on function approximation for simultaneous multi-joint myoelectric control [14].

Figure 7 shows the improvement in mean absolute error gained by iterating over the training data. The normalized error on testing data after one iteration (online operation) was smaller than the typical difference between dominant predictions and the next largest prediction—accurate predictions were achieved after only a single iteration through

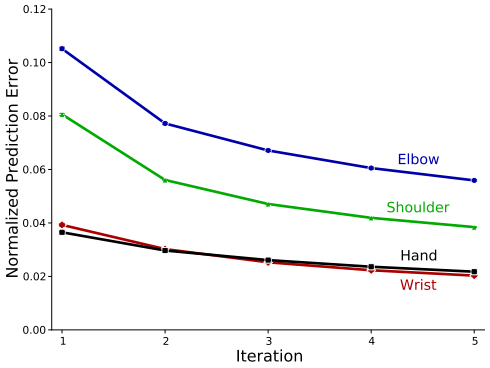


Fig. 7. **Improvement in prediction error after multiple learning iterations over the training data**, in terms of the mean absolute normalized error between learned predictions and the true, post-hoc predictions.

the training data. Further improvement occurred with added online data or additional passes through the training data.

Taken as a whole, the techniques described in this work enable a control approach that automatically adjusts switching order to offer a patient the most useful DoF for a given context and task. The results presented above suggest that this dynamic switching approach will substantially reduce the delay and cognitive load required to use a switching-based controller. While clinical proficiency tests such as the box and blocks task are not a replacement for out-of-clinic user testing, this preliminary study indicates that a dynamic switching approach may extend well to common daily tasks. However, since the patient will no longer be able to memorize a fixed switching order, an adaptive controller must also be combined with a feedback system to relay switching outcomes and options to the patient. We are currently working on tactile, vibratory, and audio feedback devices to achieve this functionality. This combined learning and feedback system should further decrease the number of switching errors and thus minimize switching related delays. Future work will describe comprehensive testing with a population of amputee patients, and the use of real-time prediction to augment other aspects of myoelectric control.

## VI. CONCLUSIONS

In this work we discussed the problem of controlling an assistive biomedical device with fewer control channels than available degrees of freedom. Specifically, we demonstrated a real-time machine learning approach to help support a user in manually switching between the joints of a multi-function myoelectric training prosthesis. The presented learning system was capable of accurately predicting and anticipating future control activity and control signals to be executed by a user, and using these predictions to generate an adaptive switching order that greatly improved upon a fixed, expert-designed switching scheme. This promises to optimize patient-prosthesis interaction, reduce delays related to manual function switching by patients, and increase patient uptake of conventional multifunction prostheses. Online learning of this kind also provides a basis for future work on intuitive, simultaneous control of multiple joints.

While we demonstrated this work in the context of myoelectric control, the described techniques are equally applicable to other domains involving direct human-machine interaction. Examples include semi-autonomous wheelchair control, assistive walking, remote microsurgery, and tele-operation of autonomous vehicles. There are a number of situations where humans must control a system using limited control channels (in terms of number or cognitive power). We suggest that real-time predictive learning is a fundamental tool for making control not only faster and more intuitive, but more powerful in terms of the functionality a user is able to achieve using an assistive robotic system.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the Alberta Innovates Centre for Machine Learning (AICML), the Natural Sciences and Engineering Research Council (NSERC), Alberta Innovates – Technology Futures (AITF), and the Glenrose Rehabilitation Hospital Foundation. Thanks as well to Joseph Modayil, Adam White, Jacqueline Hebert, K. Ming Chan, and Michael Stobbe for a number of insightful discussions.

## REFERENCES

[1] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82–94, 1993.

[2] P. Parker, K. B. Englehart, and B. Hudgins, "Myoelectric signal processing for control of powered limb prostheses," *J. Electromyography Kinesiol.*, vol. 16, no. 6, pp. 541–548, 2006.

[3] T. W. Williams, "Guest Editorial: Progress on stabilizing and controlling powered upper-limb prostheses," *The Journal of Rehabilitation Research and Development*, vol. 48, no. 6, pp. ix–xix, 2011.

[4] M. A. Oskoei and H. Hu, "Myoelectric control systems—a survey," *Biomed. Signal Process. Control*, vol. 2, no. 4, pp. 275–294, 2007.

[5] 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," *The Journal of Rehabilitation Research and Development*, vol. 48, no. 6, pp. 643–660, 2011

[6] 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," *The Journal of Rehabilitation Research and Development*, vol. 48, no. 6, pp. 719–738, 2011.

[7] Todd A. Kuiken, Guanglin Li, Blair A. Lock, Robert D. Lipschutz, Laura A. Miller, Kathy A. Stubblefield, Kevin B. Englehart, "Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms," *Journal of the American Medical Association*, vol. 301, no. 6, pp. 619–628, 2009.

[8] 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.

[9] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction*. MIT Press Cambridge, Massachusetts; London, England, 1998.

[10] R. S. Sutton, J. Modayil, M. Delp, T. Degris, P. M. Pilarski, A. White, D. Precup, "Horde: a scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction," in *Proc. of 10th Int. Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, May 2-6, 2011, Taipei, Taiwan, pp. 761–768, 2011.

[11] M. R. Dawson, F. Fahimi, J. P. Carey, "The development of a myoelectric training tool for above-elbow amputees," *The Open Biomedical Engineering Journal*, in press, 2012.

[12] V. Mathiowetz, G. Volland, N. Kashman, and K. Weber, "Adult norms for the box and block test of manual dexterity," *The American Journal of Occupational Therapy*, vol. 39, no. 6, pp. 386–391, 1985.

[13] S. Singh and R. S. Sutton, "Reinforcement learning with replacing eligibility traces," *Mach. Learn.*, vol. 22, no. 1, pp. 123–158, 1996.

[14] 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. IEEE International Conference on Rehabilitation Robotics*, pp. 134–140, 2011.