Prosthetic Devices as Goal-Seeking Agents

Patrick M. Pilarski*
Div. of Physical Medicine and Rehabilitation
Dept. of Medicine, University of Alberta
Edmonton, Alberta, Canada, T6G 2E1
* Corresponding author: pilarski@ualberta.ca

Richard S. Sutton
Dept. of Computing Science
University of Alberta
Edmonton, Alberta, Canada

Kory W. Mathewson
Dept. of Computing Science
University of Alberta
Edmonton, Alberta, Canada

***Abstract*—In this article we develop the perspective that assistive devices, and specifically artificial arms and hands, may be beneficially viewed as goal-seeking agents. We further suggest that taking this perspective enables more powerful interactions between human users and next generation prosthetic devices, especially when the sensorimotor space of the prosthetic technology greatly exceeds the conventional myoelectric control and communication channels available to a prosthetic user. As a principal contribution, we propose a schema for thinking about the capacity of a human-machine collaboration as a function of both the human and machine's degrees of agency. Using this schema, we present a brief analysis of three examples from the literature where agency or goal-seeking behaviour by a prosthesis has enabled a progression of fruitful, task-directed interactions between a prosthetic assistant and a human director. While preliminary, the agent-based viewpoint developed in this article extends current thinking on how best to support the natural, functional use of increasingly complex prosthetic enhancements.**

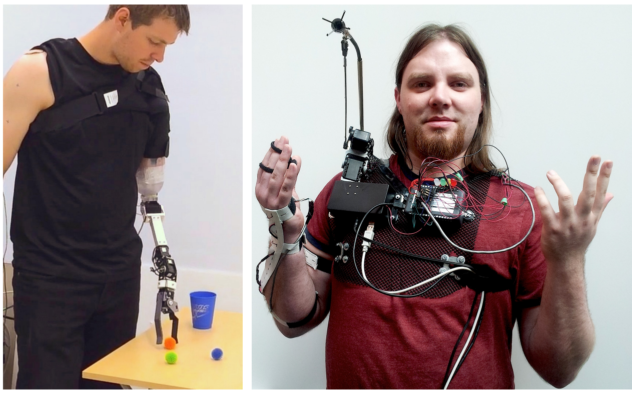

Fig. 1. Examples of human-prosthesis interaction. *Left:* a subject with an amputation using the University of Alberta Bento Arm [6] with conventional myoelectric control to complete a manipulation task. *Right:* control of a supernumerary limb by a non-amputee subject [7].

I. The Prosthetic Future

Upper-limb prosthetic devices have evolved over the last several hundred years from crude iron hands to exquisitely designed bionic body parts [1], [2]. However, despite great improvements in quality of life for those with lost limbs, the state-of-the-art has yet to create a satisfactory substitute for the nearly 1 in 200 Americans living with amputations [1]–[3]. Significant advances have been made. Extensions to hardware, software, and interfaces have paved the way for increasingly more adaptive and functional prosthetic technologies. Of note, the actuation capabilities of both commercial and experimental powered upper-limb prostheses far surpass the ability of users to manipulate all available degrees of control [4]. Advances in software, shared control between the human and the prosthesis, and machine learning in the device itself are now needed to fully bridge the gap between a user and their prosthesis [2], [5].

Prostheses are interesting in part because of the intimate way control is shared between a human and their device (Fig. 1); from a technical standpoint, the prosthetic setting is both challenging and appealing due to the dynamic, non-stationary nature of human environments [8]. Prosthetic devices must therefore maintain and update a representation of their environment, sharing some subset of their perception of the world with their human user. Modern technology enables increasingly powerful shared representations. Muscular, neural and osseo-integration allow for direct connections between the human and the device [1], [2], [9]. Onboard cameras have been shown to facilitate real-time visual object tracking and object recognition for grasp pre-shaping [10]. Microphones and speakers facilitate natural-language interactions with devices, as seen in related domains [11], [12]. Sensory feedback and surgical practice have further evolved to restore sensation to prosthesis users [13], [14].

Future prosthetic devices will receive an unprecedented density of data about the user, their needs, and their environment. This stream of data will need to be skillfully leveraged to enable the coordination of vast numbers of complementary actuators and functions. As such, and in addition to advances in communication streams between the device and the human, prosthetic limbs will soon need to actively build and improve their representation of the world around them. Prosthetic limbs will be required to structure a vast amount of data to better make decisions in support of their users' needs and goals.

The principal contribution of this work is therefore to suggest that a prosthetic device should be an *agent*—i.e., that it should be an autonomous system that both has and seeks goals. In more general terms, we propose that the parts of a larger information processing system (e.g., both sides of a tightly coupled human-machine interface) are well thought of as each being full information-processing systems with goals. We further suggest that for maximum benefit all parts of an interface should model the other parts as being goal-seeking systems. In the remainder of this manuscript we will develop the intuition behind an agent-based viewpoint.

II. Goal-Seeking Communication and Control

There are multiple means by which the human and an agent—e.g., an assistive robot like a prosthesis—can beneficially interact to achieve the human's objectives [11], [15], [16]. In much of the existing literature, one or more feedback channels are used as a means by which a non-expert can train, teach, and interact with a system without explicitly programming it. This shaping allows for the human to learn how the robot accepts and interprets feedback, and for the robot to learn what the human's goals are for their shared interaction [17]. A selection of representative examples are briefly surveyed below. Pilarski et al. introduced the use of actor-critic reinforcement learning for myoelectric limb control and showed that a user could train a virtual robotic appendage with a single, scalar reward signal provided by the user [18]. Knox and Stone explored a wide variety of strategies for incorporating feedback with environmental reward. They found that Action Biasing and Control Sharing, both using feedback as policy modifiers rather than changing the reward function, produced the best results [19]. Griffith et al. built on the work of Knox and Stone with Advise, a framework to maximize the information gained from human feedback by associating policy labels [20]. Advise outperformed other modern methods in robustness to noise. They also explored how other parameters, such as feedback consistency, affected the performance of a learning agent. Loftin et al. have further expanded the space of human interaction through detailed investigation of human teaching strategies and developed systems which model the human feedback. Their systems learn in substantially fewer episodes and with less feedback than other approaches [21].

III. Agency

We now turn to the specific case of a human interacting directly with a prosthetic device, and define a schema for thinking about the levels of agency that each side of the human-machine interface may obtain. In what follows, we will refer to the human as the *director*, and their prosthesis or other assistive device as the *assistant*. For the purposes of our present discussion, we define agency as the ability of an autonomous system to have and seek to achieve goals. This definition is similar to the Belmont Report (1979), wherein a system assumes agency if it is "capable of deliberation about personal goals and of acting under the direction of such deliberation" [22]. Hallmarks of agency include the ability to take actions, have sensation, persist over time, and improve with respect to a goal; these hallmarks give rise to an agent's ability to predict, control, and model its environment (including other agents).

Agency is not easily identified as present or absent in a non-human system. As one contribution of the present work, we therefore attempt to identify one viable schema for thinking about agency in a prosthetic setting. Figure 2 presents this schema, where each level includes and extends the capabilities of the preceding levels as follows:

1) Mechanism: The system acts in a fixed or predetermined way in response to the state or stimulus. For example, the standard case of a body-powered prosthesis, or a conventional myoelectric controller that processes EMG signals via a fixed linear proportional mapping to create control commands for the prosthetic actuators [23]. Even mechanistic systems may be considered to have a degree of agency—goals may be present in the form of error measures or simple homeostatic control feedback loops.

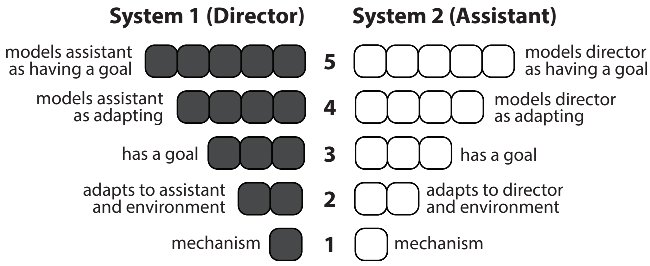

Fig. 2. Levels of goal-seeking agency during human-prosthesis interaction.

2) Adapts over time: A changing mechanism. In addition to acting mechanistically, the system has the capacity to adapt or change in response to the situation and signals perceived from the other agent. One clear way of thinking of this case is a system that gradually acquires knowledge about its situation in the form of changing parameters, thresholds, or forecasts about how to act and what will happen in the future (i.e., prediction and control learning). Adaptation can occur over scheduled periods of time, as in the supervised learning of a pattern recognition classifier, or during ongoing experience [2], [24].

3) Has a goal: The system has defined goals or objectives, with the intent to maximize or optimize some measure of its own situation. One way that goals may be defined is in the form of scalar signals of reward (success), as in the computational and biological reinforcement learning literature [25]. This level of agency is the common case for the director—the human user of a commercially available myoelectric or mechanical prosthesis.

4) Models the other agent as adapting: The agent views the other agent as changing and building up expectations (for example, predictions) during ongoing interaction, and in response to the signals it generates. This alters the way the first agent presents signals to the second agent. An example of this level of interaction is an amputee (the director) training a pattern recognition controller (the assistant), knowing that assistant is adapting to the signals the director generates.

5) Models the other agent as having a goal: The agent views the other agent as not only changing in response to received signals, but also as having its own objectives. This may be viewed as the agent having at least a preliminary "theory of mind," further altering the way one agent presents signals to the other agent. Viewing another agent as an adaptive, goal-seeking system enables more advanced forms of direction, collaboration, and instruction.

Figure 3 uses this schema to depict three example combinations for the director and assistant that can be readily identified from the contemporary literature. The ultimate case is when both systems accurately think of the other as having a goal and making predictions to achieve that goal.

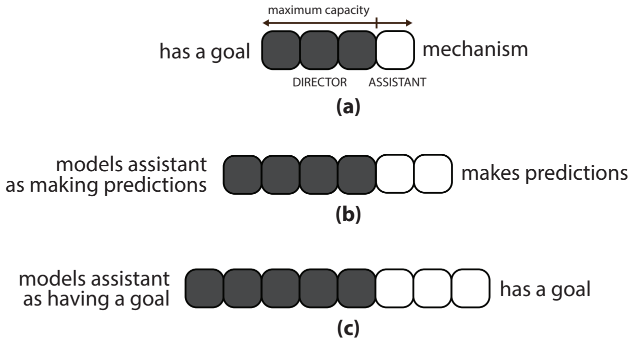

Fig. 3. Combined levels of agency and the resulting capacity. a) The standard setting for prosthetic control, where the director utilizes a fixed mechanism; b) the case demonstrated by Edwards et al. [26], and in commercially deployed pattern recognition, wherein the human interacts with a prosthetic system known to acquire, maintain, and use predictions (knowledge) in control; c) the case where the assistant has knowledge-supported goals and the director views the assistant as having knowledge-supported goals—for example, reward-based training of a myoelectric controller by Pilarski et al. [18]. (n.b., boxes need not be equal in scale.)

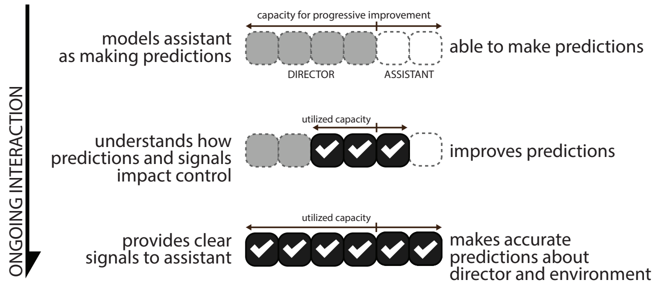

Fig. 4. An example progression wherein the combined ability of the director and assistant approaches its full capacity via ongoing human-machine interaction. Over time and through better modelling of each others' behaviour and the environment, the systems build up *communicative capital* (shown as ticked boxes) that can be used to achieve their goals. One natural type of communicative capital is knowledge in the form of predictions.

In the examples that follow, we assume that a human director has a set of goals that relate to their task, needs, and environmental setting. Defining the goal of the assistant is not always straightforward, but one possible and immediate goal for the assistant is gaining the approval of the director. Approval may be communicated to the assistant via any of the normal communication channels between the two agents, or through a privileged channel dedicated to reward. Like an assistant in the corporate sense, a goal-seeking prosthetic assistant would strive to maximize approval, but would also have its own goals that may be overridden by directions from the user. For example, the assistant may have the goal to protect itself at the onset (to prevent its motors from overheating during use or its battery from running dead), but in order to secure the user's approval, have the capacity to align its behaviour over time to the user's goals as they become clear to the agent. How the goals of the director and assistant can come to align in a general sense is an interesting problem for future discussion.

IV. Improvement through Interaction

So far we have considered the setting in which a communication channel is opened up between the director and the assistant. Communication will improve the lot of both systems. If the receiver's reward is not improved by the information received, then it will ignore the information. If the sender's reward is not improved by what the receiver does with the information, then the sender will not send it. The sender can send any of many possible things. It follows that the sender should vary its communication to find the information to send that results in both systems improving with respect to their goals. One example of a simple progression is shown in Fig. 4.

We suggest that the most general and powerful way for this mutual improvement to happen is when the two sides are both goal-seeking systems and the discovered interaction is good for both (in either the short or long terms). This is in line with viewpoints from of the field of interactive machine learning (surveyed in Section II), wherein a non-expert director interactively trains a reinforcement learning agent to perform according to the director's goals in a new or previously unspecified environment. In these forms of interactive learning, the director must be aware of the mechanisms that shape the agent's choices, so as to be able to use that knowledge to stage better interactions. However, progressive improvement is possible even in existing myoelectric control systems where the assistant lacks explicit goals (Fig. 2, Levels 1 or 2). We present three representative cases of incremental improvement below, corresponding to the cases shown in Fig. 3 (a–c).

A. Mechanisms: Conventional Myoelectric Control

Conventional myoelectric control interactions between someone with an amputation and their prosthesis is a standard case of a goal-seeking agent (the director) interacting with a mechanistic assistant (Fig. 3, a). In this setting, the only agent capable of progressive change is the director, who learns to better use the mechanism to achieve their goals. The capacity of the complete system is therefore a function of the director's ability to learn, improve, and adapt, with a fixed (but potentially significant) contribution from the nature of the assistant. A comparable analogy is an elite athlete adapting to their sporting equipment. We know from clinical experience that training is a large part of successful myoelectric control by people with amputations.

B. Learning: Predictively Enhanced Myoelectric Control

There are multiple examples of prosthetic director-assistant interactions where the director views the assistant as adapting (specifically, as making predictions or control forecasts) and the assistant acquires knowledge about the director to better execute the director's intention (the progression depicted in Fig. 4) [2], [24], [26]. A first example is commercial pattern recognition, wherein the director is able to engage a training phase to inform the assistant about the right motions to perform in response to their myoelectric commands. The director becomes more skilled at providing clear training commands, in part because of their knowledge that the assistant is learning from their demonstrations. A second example is *adaptive switching* [26], wherein the assistant learns and makes ongoing predictions about how a user will switch between the many functions of a prosthetic device. In adaptive switching, the director improves their ability to quickly execute tasks based on the assistant's switching suggestions; at the same time, the assistant improves its suggestions based on ongoing observations about the director's actions and preferences [26].

C. Goals: Reward-Based Myoelectric Control

Goal-seeking behaviour in prosthetic assistants is less prevalent. However, previous work by Pilarski et al. demonstrated how both predefined and also human-delivered reward could be delivered to a goal-seeking assistant to gradually improve the control capabilities of a myoelectric control interface [18], [27]. By using a goal-seeking reinforcement learning agent to control the joints of a prosthesis, informed by implicitly or explicitly acquired predictions about future movement, the director-assistant team was found to be able to progressively achieve greater levels of simultaneous multi-joint myoelectric control. In these studies by Pilarski et al., the approval and disapproval was delivered by the director to the assistant with full knowledge of the assistant's learning capacity. Extensions of these initial studies to more complex settings are in progress, and should inform whether or not this form of goal-seeking by the assistant holds merit for daily-life myoelectric control by people with amputations.

V. Conclusion

In this work we developed a schema for thinking about the levels of agency that are present during human-prosthesis interaction, and suggested how increasing the agency of a prosthesis (the assistant) may improve the capabilities of its human user (the director). Using this schema, we examined three examples of myoelectric control with varying degrees of agency on the part of the prosthesis. While this work is preliminary, and further studies are needed to examine the impact of agency on prosthetic control capacity, we believe that a goal-seeking viewpoint on assistive technology contributes unique and complementary ideas to the future development of a highly functional prosthetic devices.

References

[1] K. J. Zuo and J. L. Olson, "The evolution of functional hand replacement: From iron prostheses to hand transplantation," *Can. J. Plast. Surg.*, vol. 22, no. 1, pp. 44–51, 2014.

[2] C. Castellini, P. Artemiadis, M. Wininger, et al., "Proceedings of the first workshop on peripheral machine interfaces: going beyond traditional surface electromyography," *Frontiers in Neurorobotics*, vol. 8, no. 22, Aug. 2014.

[3] K. Ziegler-Graham, E. J. MacKenzie, P. L. Ephraim, et al., "Estimating the prevalence of limb loss in the united states: 2005 to 2050," *Arch. Phys. Med. Rehabil.*, vol. 89, no. 3, pp. 422–429, 2008.

[4] B. Waryck, "Comparison of two myoelectric multi-articulating prosthetic hands," *Proc. of the MEC'11: Myoelectric Controls Symposium*, Fredericton, NB, August 14–19, 2011.

[5] L. Resnik, S. L. Klinger, et al., "Controlling a multi-degree of freedom upper limb prosthesis using foot controls: user experience," *Disabil. Rehabil. Assist. Technol.*, vol. 9, no. 4, pp. 318–329, 2014.

[6] M. R. Dawson, C. Sherstan, J. P. Carey, et al., "Development of the Bento Arm: An improved robotic arm for myoelectric training and research," *Proc. of MEC'14: Myoelectric Controls Symposium*, 2014, pp. 60–64.

[7] A. S. R. Parker, A. L. Edwards, P. M. Pilarski, "Using Learned Predictions as Feedback to Improve Control and Communication with an Artificial Limb: Preliminary Findings," arXiv:1408.1913 [cs.AI], 7 pages, 2014.

[8] G. N. Saridis and H. E. Stephanou, "A hierarchical approach to the control of a prosthetic arm," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 7, no. 6, pp. 407–420, 1977.

[9] L. R. Hochberg, M. D. Serruya, G. M. Friehs, et al., "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, no. 7099, pp. 164–171, 2006.

[10] M. Markovic, S. Dosen, C. Cipriani, et al., "Stereovision and augmented reality for closed-loop control of grasping in hand prostheses," *J. neural engineering*, vol. 11, no. 4, pp. 046001, 2014.

[11] A. L. Thomaz and C. Breazeal, "Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance," *Proc. AAAI*, 2006, pp. 1000–1005.

[12] T. Kollar, S. Tellex, D. Roy, et al., "Toward understanding natural language directions," *Proc. 5th ACM/IEEE Int. Conf. Human-Robot Interaction (HRI)*, 2010, pp. 259–266.

[13] J. S. Schofield, K. R. Evans, J. P. Carey, et al., "Applications of sensory feedback in motorized upper extremity prosthesis: a review," *Expert Rev. Medical Devices*, vol. 11, no. 5, pp. 499–511, 2014.

[14] J. S. Hebert, K. Elzinga, K. M. Chan, et al., "Updates in targeted sensory reinnervation for upper limb amputation," *Current Surgery Reports,* vol. 2, no. 3, pp. 45-53, 2014.

[15] B. D. Argall, S. Chernova, M. Veloso, et al., "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 57, no. 5, pp. 469–483, 2009.

[16] P. M. Pilarski and R. S. Sutton, "Between instruction and reward: human-prompted switching," *AAAI 2012 Fall Symposium on Robots Learning Interactively from Human Teachers (RLIHT)*, Nov. 2–4, Arlington, VA, USA, 2012, pp. 46–52.

[17] A. L. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," *Artif. Intell.*, vol. 172, no. 6, pp. 716–737, 2008.

[18] P. M. Pilarski, M. R. Dawson, T. Degris, et al., "Online human training of a myoelectric prosthesis controller via actor-critic reinforcement learning," in *Proc. IEEE Int. Conf. on Rehab. Robotics*, Zurich, Switzerland, 2011, pp. 134–140.

[19] W. B. Knox and P. Stone, "Reinforcement learning from simultaneous human and MDP reward," *Proc. 11th Int. Conf. Autonomous Agents and Multiagent Systems*, 2012, pp. 475–482.

[20] S. Griffith, K. Subramanian, J. Scholz, et al., "Policy shaping: Integrating human feedback with reinforcement learning," in *Advances in Neural Information Processing Systems*, 2013, pp. 2625–2633.

[21] R. Loftin, J. MacGlashan, B. Peng, et al., "A strategy-aware technique for learning behaviors from discrete human feedback," *Proc. of AAAI*, 2014, pp. 937–943.

[22] Belmont Report (1979). The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research. [Online]. Available: http://hhs.gov/ohrp/humansubjects/guidance/belmont.html

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

[24] P. M. Pilarski, M. R. Dawson, T. Degris, et al., "Adaptive artificial limbs: A real-time approach to prediction and anticipation" *IEEE Robot. Autom. Mag.*, vol. 20, no. 1, pp. 53–64, 2013.

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

[26] A. L. Edwards, M. R. Dawson, J. S. Hebert, et al., "Adaptive Switching in Practice: Improving Myoelectric Prosthesis Performance through Reinforcement Learning," *Proc. of MEC'14: Myoelectric Controls Symposium*, Fredericton, NB, August 18-22, 2014, pp. 69-73.

[27] P. M. Pilarski, T. B. Dick, and R. S. Sutton, "Real-time prediction learning for the simultaneous actuation of multiple prosthetic joints," *Proc. of the 2013 IEEE International Conference on Rehabilitation Robotics (ICORR)*, Seattle, USA, June 24–26, 2013, pp. 1–8.