

Machines that Learn and Teach Seamlessly

Gary Stein, *Member, IEEE*, Avelino J. Gonzalez, *Fellow, IEEE*, and Clayton Barham

Abstract—This paper describes an investigation into creating agents that can learn how to perform a task by observing an expert, then seamlessly turn around and teach the same task to a less proficient person. These agents are taught through observation of expert performance and thereafter refined through unsupervised practice of the task, all on a simulated environment. A less proficient human is subsequently taught by the now-trained agent through a third approach—coaching, executed through a haptic device. This approach addresses tasks that involve complex psychomotor skills. A machine-learning algorithm called PIGEON is used to teach the agents. A prototype is built and then tested on a task involving the manipulation of a crane to move large container boxes in a simulated shipyard. Two evaluations were performed—a proficiency test and a learning rate test. These tests were designed to determine whether this approach improves the human learning more than self-experimentation by the human. While the test results do not conclusively show that our approach provides improvement over self-learning, some positive aspects of the results suggest great potential for this approach.

Index Terms—Machine learning, intelligent tutoring systems, haptic feedback, teaching agents, learning agents, augmented feedback, psychomotor skill learning

1 INTRODUCTION

THE immense potential of teaching by and with computers has been recognized for many years—almost since their inception. In the earliest computer-aided instruction (CAI), computers were initially seen as simple providers of content—an early and much more primitive version of modern eBooks. Then, intelligent tutoring systems (ITS) came along in the early 1970s to inject intelligence in the form of domain expertise, as well as a pedagogic element to computer-aided instruction. Their goal was, and remains, to use computers to autonomously and effectively teach humans. That is, to act as a surrogate one-to-one tutor and interactively provide direction and feedback to a student to improve his/her performance in carrying out a task. We seek to extend the concept of teaching with computers by using machine learning to first “train the trainer,” so to speak. Before fully discussing our objectives, a brief look at how computers have been used for teaching humans is warranted.

1.1 Teaching with Computers—A Historical Perspective

Much research has been conducted in ITS over the last nearly 45 years, beginning with the SCHOLAR system in 1970 [1], which included a set of possible right and wrong answers that could potentially be given by a student in response to a specific question. Other ITS systems of the

time focused on getting correct answers from an expert system and comparing these to the students’ answers. Furthermore, it would provide that answer to the student if he/she was deemed to need help. In the 1980s, Woolf, [2], [3] one of the pioneers in ITS, focused on planning and student modeling. Student modeling is the process of building a model of the student’s “organization of knowledge in the subject domain,” so that the ITS can attempt to understand what the student thinks he/she knows about the topic being taught [2]. This allows the ITS to adapt its teaching to the rate at which the student is learning, as well as detecting and resolving any misconceptions and confusions that the student may have. If the student model suggests that the student holds an erroneous belief about the subject, additional questions about that belief can be asked to verify its existence, and clarify its nature [3]. Woolf [3] provides examples of this with sample discourses between a student and an ITS, regarding the subjects of rain and temperature in Oregon. Student modeling has become, and remains to this day, a heavily researched element of ITS [4]. In one of her works with Beck [5], Woolf further expanded student modeling by incorporating a machine-learning (ML) agent into an ITS called AnimalWatch. This system uses “information about the student, the current topic, and the problem, and the student’s efforts to solve the problem” [5] to attempt to predict the answer given by a student to specific math questions, as well as the time taken to provide that answer. ITS have largely concentrated their efforts in teaching academic subjects, such as math, science, history, and so on, through carefully and extensively authored systems that provide extensive feedback to the trainee about his/her performance. ITS research has flourished in the last 30 years, as can be seen in the 2009 special edition of IEEE TLT on ITS [6].

In parallel, researchers in the 1980s and 1990s began to make use the emerging power of computer simulation and graphics to place trainees in virtual environments that are nowadays often barely distinguishable from the real world.

• G. Stein is with the Intelligent Systems Laboratory, Division of Computer Science, University of Central Florida, 4000 Central Florida Blvd., HEC 346, Orlando FL 32816-2362, and Primal Innovations, Inc., Sanford, FL. E-mail: gary@primalinnovations.com.

• A.J. Gonzalez and C. Barham are with the Intelligent Systems Laboratory, Division of Computer Science, University of Central Florida, 4000 Central Florida Blvd., HEC 346, Orlando FL 32816-2362. E-mail: avelino.gonzalez@ucf.edu.

Manuscript received 9 Oct. 2012; revised 31 Jan. 2013; accepted 25 Aug. 2013; published online 11 Sept. 2013.

For information on obtaining reprints of this article, please send e-mail to: lt@computer.org, and reference IEEECS Log Number TLT-2012-10-0133. Digital Object Identifier no. 10.1109/TLT.2013.32.

Simulators have been used effectively to teach pilots, astronauts, and truck drivers how to operate the complex machines they control. Simulators have been mostly used to teach tactical skills, particularly by the US military as well as those of other nations. While they are also used to teach psychomotor skills, they originally contained no pedagogy—just presentation of a virtual environment on which the trainee could practice.

Integration of intelligent tutoring concepts into training simulators began in the 1980s. Interrante and Biegel developed an intelligent simulation training system (ISTS) [7] in an attempt to introduce feedback, assessment, and pedagogy into simulation-based training—a kind of ITS for simulation systems, as its name suggests. The ISTS contained a graphic simulation component and an intelligent tutor [7]. The tutoring system serves to instruct the student in the correct manner of manipulating the simulation, and to evaluate the student's performance at the end. Other instances of intelligent tutors linked to training simulation for military flight training can be found in [8] and [9].

Our motivation comes from the general desire to further integrate (some of) the teaching ability found in ITS into simulation-based training systems. This would increase the effectiveness of simulations-based training, beyond merely serving as immersive environments where the trainee(s) can exercise their skills. Furthermore, we seek to provide this interaction through a *haptic* interface (one that uses our sense of touch—see the definition in Section 2.2), rather than through words, voice or graphics. Furthermore, and equally importantly, we also seek to reduce the time and effort required to author such systems by developing a method that uses machine learning to seamlessly create agents that learn and teach autonomously, with little human intervention. We refer to these agents as *Learning and Teaching Agents*, or *LATA*.

We focus our work on effectively teaching psychomotor skills to a human, rather than the academic subjects normally addressed by ITS, or the tactical skills covered by most military simulators. The Encarta Dictionary [10] defines psychomotor skills as “relating to bodily movement triggered by mental activity, especially voluntary muscle action.” In our context, this implies the manipulation of some control elements (e.g., steering wheel, joystick, brakes, buttons, dials) to effectively manage the actions of a device. We focus on providing feedback in real time—a form of *coaching*—and through haptic interaction.

This paper reviews the state of the art in Section 2. Then, in Section 3, we set forth the objectives of the research in the context of prior work described in Section 2. We propose and describe our approach in Section 4. Section 5 describes the infrastructure used to test the prototype system. Section 6 describes how the LATA agents were taught the psychomotor skills, while 7 describes how they were used to teach humans the same skills. Relevant test results are included and discussed in these sections. Lastly, Section 8 concludes and summarizes the paper.

2 STATE OF THE ART

Our work draws from several research areas in engineering and computer science as well as in education. These

relevant areas enjoy rich and extensive bodies of research literature. A comprehensive coverage of each area is beyond the scope of this paper. In this section, we provide a review of the relevant literature in these areas to place our work in the context of the state of the art. Given its central aspect within our work, we begin by discussing feedback to the student.

2.1 Feedback and Coaching

Coaching is a very traditional method of teaching physical skills to unskilled humans. Webster's dictionary [11] defines it as “to instruct (a person) in a subject, or prepare (a person) for an examination, by private tutoring.” The term coaching, if not the concept, originated in sports [12], where the terms “coach” and “coaching” seem to be ubiquitous, for example, football coach, tennis coach, first-base coach. When coaching a trainee, a coach seeks to exhort the trainee into performing the skill correctly, teaching him/her by demonstration of correct action and by continuously providing feedback. Human-based coaching is generally oral, such as when providing spoken directions on how to perform a task properly. However, it can also be physical, such as when a coach holds a child's hands as the child holds a baseball bat, moving both through the correct motion to help the child learn to swing at a baseball resting on a tee. This process is generally one-to-one, as suggested in the dictionary's definition above.

In recent years, the term “coach” has evolved into one with many related, but somewhat different meanings. It is commonly used alongside other seemingly unrelated nouns, such as *life coach*, *health coach*, *workout coach*, *executive coach*, and several others. Whitmore [12] defines coaching as “unlocking people's potential to maximize their own performance,” implying directed self-learning. However, we assert that coaching is still best recognized as the general concept of instructing a trainee through demonstration and feedback on a one-to-one basis, with the goal of improvement in the trainee's performance.

There is no reasonable doubt that availability of feedback to a student is a very important part of learning. Bilodeau and Bilodeau [13] found that for motor skill learning, nearly any kind of feedback improves the performance of a student. They state that student feedback is “the strongest, most important variable controlling performance and learning” [13]. Todorov et al. [14] classify feedback into *knowledge of results* and *augmented feedback*. In knowledge of results, students can learn to improve their performance by simply knowing the optimal results of the task or skill being learned, and comparing their performance against it. Such feedback is often naturally available from the environment. Augmented feedback, on the other hand, requires “extra, artificially generated information” [14] to be provided. Kulhavy and Stock [15] refer to these same types of feedback as *verification* and *elaboration*, where the former only informs the student of the correctness of their answer while the latter provides direction on how to improve performance.

Todorov et al. [14] used a virtual environment to provide augmented feedback about the performance of a game in the real world that requires psychomotor skills (ping-pong). Their experiments measured arm motion in

the real world and transferred those to a simulation that was used to provide the feedback. The feedback provided was visual. They found that feedback augmented by this simulation did improve the learning of psychomotor skill in some situations.

Computer-based coaching emerged as a discipline in the 1970s, in conjunction with the progress made in ITS, as these can serve as a form of computer-based coaches. ITS can provide an assessment of the student after completing an exercise as well as provide interventions and guidance during the exercise. ITS can also demonstrate the correct approach if the student requires it and facilitate a one-on-one learning experience. Mason and Bruning [16] in their extensive review of the literature on computer-based feedback in 2001 found that there was no best way to provide feedback, and that each application was different vis-à-vis the best type of feedback to provide. The following examples illustrate applications of ITS as computer-based coaches.

Burton and Brown [17] discuss the subject of computer coaching and point to how subtly difficult it can be to build such systems. Deesen and Tilak [18] patented a method for computer-based coaching for graphic art design in 1992, although they list several other potential applications. It appears to consist mainly of providing textual feedback. Dickson and Parsons [19] argue for computer-based coaching for athletic endeavors, although they provide no new research. Constantino-González and Suthers [20] designed and built an ITS called COLER, which served as a computer coach for collaborative problem solving, with database design using entity-relationship (ER) modeling serving as the test application [20]. COLER offers advice to help resolve conflicts and reconcile differing solutions among group members [20].

Zissos and Witten [21] designed and built Anchises, an ITS that serves as a computer coach for users of the EMACS text editor. Anchises' coaching took the form of advice offered to the user when "bad plans" were detected, or the system determined the user was ignoring a command that may be helpful [21]. In a pilot experiment involving eight users ranging from novice to expert, Anchises was found to give 88 percent of the advice that it should have, and only 10 percent of the advice that it should not have [21].

Much closer to our work, Acovelli and Gamble [22] have one of the earliest mentions of agent-based coaching. They discuss the limitations of traditional feedback in coaching and focus their work on providing timely and contextually appropriate textual feedback to the trainee.

Geng et al. [23] introduce an ITS architecture to teach psychomotor skills, in particular, how to fly an airplane. They state that "most flight simulators do not provide enough automated assessment" Their feedback is in the form of spoken instructions.

Research has been ongoing in teaching marksmanship, another important psychomotor skill in a military context. Chung et al. [24] describe their research, as well as that of several others in this domain. Most of the research they describe comes from psychology and the science of marksmanship, rather than training. Nevertheless, their own system provides feedback to the trainee by selecting prerecorded videos of how to perform the action correctly.

Belghith et al. [25] describe a simulation-based ITS for learning how to manipulate a telerobotic arm on the space station. Their feedback is largely through selected movies and generated paths.

The provision of feedback in military simulators has focused on after-action review (AAR), rather than real-time coaching. In an AAR session, the instructors meet the trainee team after the training session to provide them with feedback on their performance. This feedback typically refers to the tactics used by the trainees and is generated by the human instructors, generally happens offline and immediately after the training session. However, there have been some efforts at automating this process and providing feedback in real time during the training session. Jensen et al. [26] devised automatic AAR to detect different types of common errors in command and control simulations. They also incorporated analysis that could explain the events that led to the error. Surdu and Pooch [27] use agents to detect deviations in the ongoing operations as compared to the planned operations. Jensen et al. [28] use finite state automata to track the current state in the simulation and compare it to the expected behavior, which is hardcoded into an evaluation algorithm. Abbot [29] also presents a means to do automated AAR as a comparison between the actions of the trainee and an expert agent modeled by a machine-learning technique. Fernlund et al. [30] also used machine learning to build the agents against which to compare trainee tactical behavior. In a way, the works of Abbot and of Fernlund et al. were two early instances of agents that can learn and teach. However, the teaching and learning functions were not seamless, the teaching was done only as AAR rather than in real time, and the subject was tactical behavior rather than psychomotor skills.

In summary, while feedback is currently provided with most, if not all ITS, it is not as prevalent in simulation-based trainers. When it is provided, it is done mostly in terms of AAR for military trainers. Feedback is generally not provided in real time (with some exceptions as noted above) and when it is, it is generally provided through text, voice or graphics. Lastly, the knowledge necessary to provide the feedback is generally (likewise with some exceptions) created by hand at presumably great expense (although cost is not discussed in the papers reviewed). We next review haptics and haptic feedback.

2.2 Coaching through Haptic Feedback

Humans have learned to rely most strongly on vision and hearing for sensing the world around us. Yet, many times, and especially in stressful situations, these sensory channels can be overwhelmed with such a dense input stream that we cannot process all the inputs received. The fifth sense, touch, is often underused, to the detriment of learning systems.

Haptics is the study of the sense of touch. Srinivasan and Basdogan [31] define a haptic interface as one that "displays tactual sensory information to the user by appropriately stimulating his or her tactile and kinesthetic sensory systems." Under the most general of definitions, mice and keyboards could be considered passive haptic devices because they provide input to the computer, as the human is able to feel the interaction with the device, and by extension, the computer. Additionally, one's ability to use

his/her sense of position to move one's hand and press a button could also be considered haptic interaction. However, these are basic, one-way interactions, as the computer does not reflect back any outputs through these same devices. Such reflection traditionally takes place visually through the display and/or audibly through speakers. Computer haptic devices, on the other hand, refer to an active force-feedback system within an interface device. These devices can provide sensing cues that are presented to a human in a natural way.

Our work employs the touch and feel of a control device, such as a joystick, to act as the interface between the human teacher/learner and the computer. Haptics is an important element of our presented approach. Our focus is to use haptics as the primary means to provide verification as well as elaboration feedback to trainees in psychomotor teaching applications, where arms, hands, and/or feet are used to control a machine or device. We use haptics as the only means of providing feedback to the trainee when teaching psychomotor skills.

One obvious use of computer haptic interfaces has been to use them to engage trainees who suffer from vision and/or hearing impairment [32]. Another obvious application is training for surgery. In fact, much discussion has taken place in the medical literature about training surgeons via virtual reality [33], [34]. This has been particularly accelerated by the emergence of robotic surgical procedures, where the actual robots are used to practice, thus providing a semblance of haptic feedback [35], albeit only as verification feedback and not as elaboration. van der Meijden and Schijven [36] conducted a study about the effectiveness of haptic feedback (among other things) on training for surgery, and found that "no firm consensus exists on the importance of haptic feedback in performing minimally invasive surgery," despite of some positive results.

Haptic response to surgical robots has been a research issue since 1985. Westebring-van der Putten et al. [37] reviewed the literature for this body of research and found it to be extensive. However, their discussion of haptics in medical applications (surgery) mostly refers to haptic feedback in the actual devices, and not in the context of training. Okamura [38] indicates that the lack of such haptic feedback in surgical devices "is a limiting factor" in the success of these devices.

Other applications of haptic feedback in medical education include the work of Schaffer et al. [39] who used haptics to train cardiologists. Burdea et al. [40] use virtual reality and haptic feedback to train physicians in the diagnosis of prostate cancer. Brooks et al. [41] use haptics to explain molecular docking to medical students. Williams et al. [42] use haptics to teach osteopathic medical students about back problems. Nijholt et al. [43] introduce the concept of haptic feedback for training medical students to do subcutaneous injections on simulated patients. However, they did not seek to model expertise, but only to provide predefined haptic feedback for specific tasks. In each case, the learner was able to improve better than through trial and error.

In other areas, Lloyd and Bull [44] use a haptic interface to allow students to inspect their student models and see for

themselves how their knowledge is being modeled in an adaptive learning environment. Their system combines a graphic interface with a haptic feedback device that distinguishes between hardness and softness to indicate topics mastered by the student and those not so. They found that the students found haptic feedback preferable to a text-based one.

Williams et al. [45] used haptics to teach engineering mechanics in college-level classes. They found that students preferred this interface to the conventional ones. Okamura et al. [46] devised a "haptic paddle" to better communicate effects of dynamic systems such as viscous damping and inertia, and to be able to better build control systems.

For psychomotor skills, as they are defined in this paper (manipulation of control elements to effectively manage the actions of a device), adding haptic feedback has the potential to increase learning over purely visual learning. In a study on teaching force skills by Morris et al. [47], visuohaptic feedback was shown to provide better recall of a desired sequence of motions, and the desired force to be applied to each motion, than purely visual or haptic feedback. A study by Feygin et al. [48] on the benefits of haptic guidance in teaching psychomotor skills demonstrated that haptic feedback can benefit performance, especially when training the temporal aspects of a task. One reason why haptic feedback may be beneficial is that psychomotor skills involve haptic input (pushing buttons, moving joysticks or steering wheels, etc.), and because the human body has an excellent kinesthetic (knowledge of body position) memory [48]. Presenting the correct haptic input through haptic feedback removes the need for the brain to perform complicated sensorimotor transformations [48]. Interestingly, however, there is also literature that suggests that haptic guidance can have negative effects on learning, by either distracting the learner or by causing an overdependence on the guidance or feedback [49].

In summary, the works reviewed make use of haptics for some applications, although few are for teaching psychomotor skills. Surgery training is one exception, although most of the work involves improving the haptics of the robotic surgical systems, rather than being used in psychomotor skill training.

2.3 Learning Agents

There is a rich extensive body of research literature in machine learning (ML) compiled over the last 50 years or so, ever since Samuel's seminal checkers program in 1963 [50]. Some impressive results have been accomplished over the years—too many to fully do justice here. Much of the ML literature involves learning how to classify a set of elements found in a database by using examples of such classifications. Many techniques exist to do this. These include neural networks with its many different varieties, support vector machines, genetic programming, inductive learning, neuroevolution, as well as many others. Learning can be said to be either *supervised*, where the truth-value of each example is provided along with the examples; *unsupervised*, where the truth-value is not provided; and *reinforcement learning*, a sort of trial and error learning where successful behavior is reinforced and unsuccessful behavior is punished in the process of the agent experiencing the

environment. A comprehensive review of these methods and approaches is beyond the scope of this paper.

Of particular interest to our work is *observational learning* (OL), where agents can be built by merely observing a human (or another agent) performing a task. Observational learning is similar in some ways to supervised learning, but there are two important differences. In observational learning, the learning examples are time based, continuous, and not easily separable throughout the exercise. Second, there is no explicit linkage between cause and effect. This important aspect of learning must be extracted from the behavior observed by the learning entity. A form of supervised learning that is more related to OL is that of *sequential learning* [51].

Much research has been done in building agents through OL. Work in this area began in 1979 with Bauer's learning to program from examples of human-written program source code [52]. More recently, Sammut et al. [53] built a pilot agent by observing someone fly an airplane. Sidani and González [54] developed a system that learns how to behave in a signaled traffic intersection by observing a human do the same in a simulation. Pomerleau [55] created a driving agent that learned from observing a human drive a real automobile. Fernlund et al. [56] used contextual reasoning and genetic programming to build agents of automobile driving through observation of humans in a simulation.

Observational learning has also seen significant advances and applications in robotics, where it is normally called learning from demonstration. See Ontañón et al. [57], and Bentivegna and Atkeson [58] for examples of work in this area, among many others.

In our work, we utilized a previously developed machine-learning algorithm called *PIGEON* [59] that was shown in [60] to successfully teach an agent how to perform a task or carry out a mission. It did so through a combination of unobtrusive observation of a human performer (an *actor*) plus exploratory self-learning. *PIGEON* was able to produce a robust agent capable of performing physical tasks proficiently and in a human-like manner in three different domains (see [59] for details). For the scope of our research, we limited the learning tasks to psychomotor skills through a haptic computer interface.

2.4 Teaching and Learning

As anyone who has ever taught others knows, learning and teaching are inseparable—one must learn something well before teaching it to others. In fact, it can be said that the true test of one's understanding of a subject is his/her ability to teach it effectively to others. Once an agent has learned something, it can be copied inexpensively and easily disseminated to teach that something to others in remote locations, such as a battlefield or a remote hospital. Such an agent could also provide an infrastructure for later evaluating a trainee. Furthermore, it could easily be enhanced when necessary.

Nechyba and Xu [61] introduced the idea of integrating teaching and learning. They used pretrained neural networks to provide feedback, but used visual cues rather than a haptic interface. Sano et al. [62] use haptics to provide coaching for human-to-human knowledge transfer using a

machine interface. Their system, however, coached a human trainee with a human trainer, not an agent.

In summary, little work has been reported in the literature that uses learning agents and teaching agents in a seamless fashion.

3 OBJECTIVES OF THIS RESEARCH

Our work addresses two problems simultaneously: 1) the incorporation of augmented, automated feedback into simulation-based training for psychomotor skills; 2) facilitating the authoring of such computer coaches, often termed a difficult and time-consuming task. We address the first through real-time haptic feedback, and the latter through a combination of learning and teaching by the same agent. The seamless combination of these two functions in an agent has to the best of our knowledge, not been reported in the literature.

Our work seeks to teach (coach) psychomotor skills to humans through computer agents that possess the appropriate skill to be transferred. These psychomotor skills are naturally haptic. We refer to these as *teaching agents* and assume that they are expert at the skill. The teaching system compares the trainee actions to those of the teaching agent, and provides the necessary feedback to the human through the haptic interface.

Teaching humans, however, is only part of our research. We also address the (earlier) step of building these teaching agents in ways that avoid, or at least reduce, the long and expensive efforts required to do this by hand. In effect, we seek to teach what will later become the teaching agents mentioned above. We can call them *learning agents* while they learn, but they are the very same agents that will later be used to teach humans. We can reduce the time and effort taken to teach the target skills to these learning agents by using machine learning. We use a simulation of the physical world to test our system.

We have two specific objectives in this research: 1) a technical objective to determine whether such a system can be built and work effectively from a computational standpoint; and 2) an educational objective to determine whether this concept can be effectively used to train humans in learning complex psychomotor skills. Note that we do not attempt to quantitatively show that we can build teaching agents less expensively than the traditional manual way. We found no data in the literature about the time and cost consumed in building these systems. Rather, we sought to show that creating these agents nearly automatically is feasible, and this can be construed to require less effort than doing so manually.

The result of our work is the LATA system, which as mentioned above, stands for *Learning And Teaching Agent*. In effect, LATA is an intermediate agent that learns from proficient humans in a central location and then teaches the same skill to less proficient humans.

4 THE LATA SYSTEM

The LATA system is the result of our research embodied in a prototype system. The first part of this research (learning) has been conceptually shown by us elsewhere, albeit in a

different application (see [59]) and is only briefly repeated here for the sake of completeness. The teaching part is the focus of this paper.

A prototype was built and a series of experiments were designed and executed to prove or disprove our hypotheses that such a system can improve learning in humans beyond self-experimentation on the environment. We selected the manipulation of a crane used to move large containers (boxes) around a large flat surface such as a shipyard. This task requires the psychomotor skills to move the gripper smoothly and grab the appropriate box in the right place, and moving it to its destination.

This research had two phases: 1) to teach the LATA agent how to manipulate the crane, and 2) to use the LATA agent to coach another human in manipulating the same crane. The second phase is the central focus of this paper and is described in detail. Nevertheless, the first phase is also discussed for the sake of completeness, albeit briefly. We begin with a discussion of PIGEON—our machine-learning algorithm.

4.1 The Learning Algorithm—PIGEON, NEAT, and PSO

We briefly discuss the machine-learning algorithm of choice for our work—PIGEON, which was developed by us and fully described elsewhere. See [59] and [60] for these details.

PIGEON stands for *Particle swarm Intelligence and Genetic programming for the Evolution and Optimization of Neural networks*. It combines neuroevolution in the form of the NEAT system [63] with Particle Swarm Optimization (PSO) [64] to perform quick learning. NEAT was developed by Stanley and Miikkulainen [63] and stands for *Neuro-Evolution of Augmenting Topologies*. It is a hybrid algorithm consisting of artificial neural networks and genetic algorithms (GA). The output of NEAT is a trained neural network. In PIGEON, NEAT “complexifies” the architecture of a population of neural networks in an evolutionary manner, and then uses PSO to optimize the weights of the neural networks. PSO is a social optimization algorithm that trains a population of individuals similarly to GAs. PSO numerical optimization capability has been shown to be able to quickly train neural networks (set the weights) when directly applied to them. PIGEON combines NEAT and PSO in an alternating manner, with cycles of NEAT and PSO interlaced. It was applied to three different domains of increasing difficulty and complexity [60]: *Chaser* (purely reactive psychomotor skills), *Sheep* (reactive psychomotor skills with planning), and *Car* (more complex reactive psychomotor skills with planning and tactical decision making). PIGEON succeeded in creating proficient and human-like agents that performed nearly as well as their human instructors—sometimes better.

PIGEON uses NEAT to provide the structure for the agent—the neural network—and allows NEAT to enhance that structure by “complexifying” the neural network structure of the emerging agents to meet the requirements of the fitness function. NEAT is essentially a GA, so it will create a population of neural network agents, rather than just one. These agent neural networks are all competing to become the best performing one. The fitness function used by NEAT is based on the *Similarity Factor* or on the

Performance Rating, depending on the phase of agent training. During observational learning, the Similarity Factor fitness function used was the similarity of the agents’ behavior to that of the human actor. In the experiential learning phase, rewards depended on the performance rating of the agent’s behavior against some absolute standard. The agent neural networks being evolved by PIGEON seek to maximize their fitness and/or rewards. The output of PIGEON is a neural network-based agent that is able to exercise a skill in the performance of a task. The interlaced enabling and disabling of NEAT and PSO in an alternating manner is the essence of PIGEON.

4.2 General Inputs and Outputs to PIGEON

The inputs to PIGEON include environmental variables determined a priori by a human designer. These variables are deemed important criteria with which to make decisions, either micro decisions such as when making use of the psychomotor skill, or higher-level decisions such as when making explicit tactical decisions, for example, turn left. These inputs include the current internal state of the agent and its external environment. The output of the neural network-based agents is a command to the actuators that control the platform of interest. Such commands manifest themselves in the actions of the platform in the simulation.

Once PIGEON has finished its work and a LATA agent has been trained, the inputs used by the LATA agent in the teaching process are the same it used in learning. Its outputs are also the same commands used to control the platform. Inputs and outputs are highly domain dependent, so they can be better explained in the context of the evaluation environment in Section 4.

4.3 Feedback through LATA Agents

Feedback is provided by the LATA system through counterforce exerted by the haptic device(s) that the trainee uses to perform a task. This counterforce indicates the correct manipulation of these devices. This counterforce can be felt by the student in real time as he/she performs the task, and can be interpreted as corrective feedback. In effect, the corrective feedback reflects how the teaching agent would perform that skill at that very moment in time, and this is felt by the trainee as the counterforce on the haptic device (a joystick in our case).

However, the Chaser, Sheep, and Car domains on which Stein did his experiments [59] were not found suitable for our work because they were either too simple or irrelevant from a practical point of view (i.e., Chaser and Sheep), or already mastered by most people (i.e., Car). A fourth domain called Crane was used. Crane provided a complex task with which most humans are not familiar and would have to learn. We describe the Crane simulation as well as what was to be learned and taught.

5 TEST INFRASTRUCTURE—THE CRANE DOMAIN AND ITS SIMULATION

The crane manipulation task was chosen because it involves a psychomotor skill that requires the coordination of both hands, as well as multistage planning and tactical decision making. Additionally, a task had to be selected

that would not already be commonly known to human test subjects. Volunteer participants would probably already know how to drive a car; however, operating a crane would likely be a new experience. Therefore, tests that include only learning from a baseline of a novice should conceivably show the appropriate amount of learning attributable to the LATA system.

5.1 Task to Be Learned

The task presented to the human actor was to seize, lift, and move several boxes (one at a time) from a given location to a designated collecting area in a fixed amount of time. The human test subjects had to use spatial reasoning to move in three dimensions, as well as psychomotor skills to maneuver the gripper above the target box, pick up the box, and move it safely around obstacles (other boxes). Ten boxes were randomly distributed on the “playing field” such that there is minimum spacing between individual boxes. Each trial was limited to 60 seconds during which the subject must pick up as many boxes as possible and move them to the *drop off zone*. With the time allotted, it is possible, although difficult, to pick up all 10 boxes. Competent humans were able to pick up and move nearly eight boxes on average. The cutoff point to be considered competent at the task was arbitrarily set at picking up and moving seven boxes.

5.2 The Crane Simulation and Its Physics

A simulation of the crane and its environment was built as part of this research. The physics model of the simulation calculates the motion based on first-order physics. The inputs from the user were directional vectors for motion proportional to the actions and not direct positions. This model approximated the movement of a real crane, with limitations on speed of translation and gripping. The simulation provided collision detection, friction, and gravity for interaction between the crane and boxes, boxes with other boxes, and boxes with the ground. The physics model provided a simplified version of a crane easily understood by users.

5.3 Visual Appearance of Simulation

The visuals of the simulation are a three-dimensional environment rendered using a simplified block model. The important aspects of the simulation were for the test subjects to easily identify the target boxes to pick up as well as the drop off location. As seen in Fig. 1, the test subject is presented with two views. The top view shows a third person perspective of the crane (represented by the “C”-shape) that follows the crane through the environment. Therefore, the crane appears to be fixed in position at the center of the screen, with the world moving around it. A second view was added in the bottom of the screen, where a top-down perspective gives an overview of the entire map. Using this view, the user can place himself/herself in the world, as well as strategize about the order in which to pick up the boxes. A projection of the crane on the floor was provided to help in the alignment process.

The border of the “playing field” is delineated by dark lines. The drop-off zone is likewise delineated by a lined square on the lower right side of the “playing field.” The crane is the facedown, black C-shaped clamp in the upper

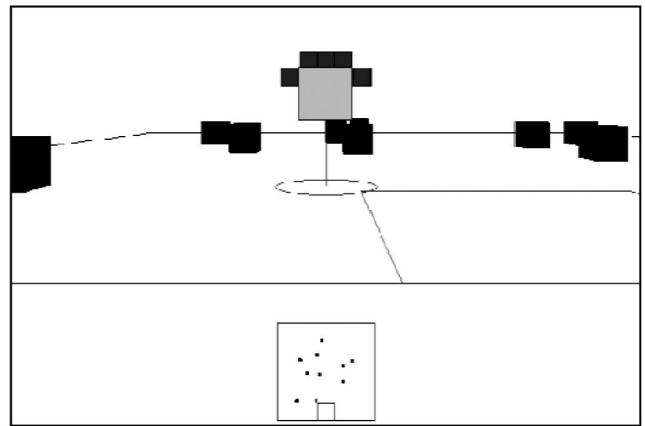


Fig. 1. Visual interface of the crane simulation.

middle of the screen; the boxes to pick up are represented by black cubes located on the surface; the box being picked up turns gray as it is grasped by the crane. Boxes already moved to the drop-off zone remain solid black (none shown in Fig. 1). Additionally, the visual projection of the crane mentioned above appears as a circle on the “floor” below the crane.

5.4 Interaction

The interaction with the simulation is strictly through the joystick interface. Typically, joysticks are considered only as input devices because they translate the angles of the joysticks into the computer for processing. However, force-feedback joysticks are also able to translate output from the computer into a force that can be felt by the user.

Two Logitech Flight System G940 force-feedback joysticks were utilized for their high torque capabilities. A custom Linux kernel driver was written to communicate with these devices. Each joystick is limited to two-axis movement (X-Y) and each axis could be independently given force values. These joysticks were not just force resistance but active force feedback to allow the agents’ full actions to be felt by the trainee while being coached.

This simulation provided the testbed for training the LATA agents through observation and through experience (practice). It was also the same environment where just-trained LATA agents were used to coach the human subjects in operating the crane. The crane operation task is moderately difficult for humans because multiple variables must be tracked. The general population in the US does not have experience operating this kind of heavy machinery, and it therefore provided for a large group of eligible test subjects. In preparation for the tests, a human participant taught himself to an expert level in operation of the crane in this simulated environment to serve as the actor in observational learning.

5.5 Specific Inputs and Outputs

Six inputs and four outputs compose the internal structure of the neural network agent for the crane simulation. The six environmental inputs to the networks are *angle* and *distance* to the target location from the crane position, the *height* above the floor, the *gripper position*, and the *angle* and *distance* to the closest noncaptured box. All angles are in

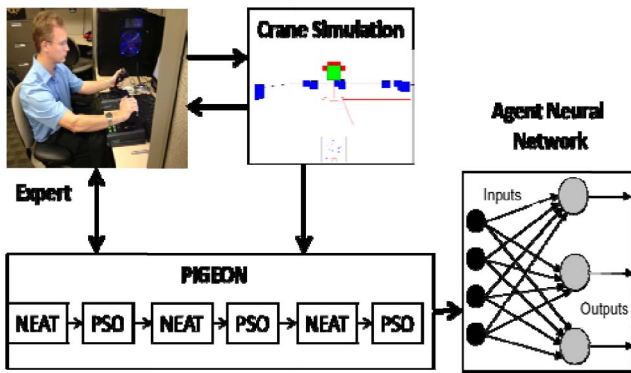


Fig. 2. Agent learning process (observational learning only).

radians from $-\Pi$ to Π . The distances are the raw values, and the playing surface is limited to a 3×3 area.

The four network outputs represent the four controllable actions: X-Y-Z translation and gripper operation. These are the same as those used in training. The left joystick controlled the X-Y translation of the crane while the right joystick controlled the Z translation and the opening/closing of the gripper. These joysticks were then mounted on a desk. Utilizing only these actions, the human could fully manipulate the crane, pick up boxes, and drop them off.

6 PHASE I—TEACHING THE LATA AGENT

Phase I is about teaching the LATA agents the task of manipulating the crane. The (observational) learning function by the LATA agent is depicted in Fig. 2. The expert (or *actor*) executes the skill in the simulator and his/her actions are observed by PIGEON and learned by the ultimate agent—the neural network output of PIGEON. In this section, we briefly discuss Phase I, which is itself composed of two parts—observational learning and experiential learning.

6.1 Observational Learning

Observational learning collects data unobtrusively from an actor performing the task in a real or simulated environment. These data include the actions of the actor and environmental variables for use in learning. The LATA agent is trained to perform the task using only these data. Observational learning does not require information verbalized by an expert. The manual process traditionally used to build computer coaches required a large effort, and experts could not always explain why they performed certain actions—“they just know” what to do. Experts can also be noncommunicative, overcommunicative and even hostile, thereby making the knowledge acquisition process difficult. Since observational learning (as we define it here) does not involve direct communication with the expert, it alleviates this problem.

The human expert (now called the actor) was observed operating the crane, moving the designated boxes. All (preselected) environment variables seen by the actor as well as the actions executed by the actor were recorded. These data were stored to disk and used by PIGEON offline in the observational learning process. The Similarity Factor metric used these data to grade how similarly the agent

Ensure: Action values limited to between 0.0 and 1.0

1: TotalFitness = 0

2: for $t = 0 : Dt : \text{Time}$ do

3: $\text{Error} = \sum_{i \in \text{Actions}} (\text{HumanActions}_i - \text{AgentActions}_i)^2$

4: $\text{Fitness}_t = Dt * (1.0 - \sqrt{\text{Error}} / \sqrt{|\text{Actions}|})$

5: TotalFitness = TotalFitness + Fitness_t

6: end for

7: SimilarityFactor = TotalFitness / Time * 100

Fig. 3. Algorithm 1: Calculation of Similarity Factor.

performs to the actor, given the same task and environment. The similarity serves as the fitness function.

The algorithm shown in Fig. 3 calculates the Similarity Factor between the actor and the agent; its output is between 0.0 and 1.0. It utilizes the sum-squared error calculated over the action differences between the actor and the agent. Since the actions are normalized from 0.0 to 1.0, the squared error can be extended to an unlimited number of inputs because it is averaged over the number of inputs. The Similarity Factor is higher as the error decreases. The PIGEON algorithm seeks to maximize fitness (minimize error). This value is then scaled to between 0 and 100 percent for easier understanding by humans.

The Similarity Factor, however, does not reflect how well the agent accomplishes the task—just how similarly to the human it behaves. The proficiency of an agent, or of the actor for that matter, is judged by the Performance Rating, an objective score that grades how proficiently the task is completed. The primary purpose of the Performance Rating is to provide an unbiased number that can be used to compare the performance between individuals, whether human or computer agents. The Performance Rating is much more domain specific than the Similarity Factor. For the crane simulation, the Performance Rating is based on the number of boxes picked up (10 points awarded per box picked up and moved successfully) as well as partial credit for how close any other box may have been brought to the designated area (the Manhattan-distance between the box's location and the edge of the designated area). These partial scores can break ties when two agents collect the same number of boxes. This partial credit can never be more than picking up and fully delivering one box. A formal definition of the Performance Rating calculation for the crane domain is given in Fig. 4.

When the agent is trained through observational learning, the agent is kept on “rails” in that it is physically positioned to the same location as the human trainer, such that the environmental inputs are the same at every time-step. When the agent is acting on its own, however, it is making all the decisions. Those decisions can move the agent into unforeseen environmental conditions that can cause the agent to perform poorly with respect to its Performance Rating, in spite of possibly having scored a high Similarity Factor.

Because observational learning is the first training method used in our learning process, PIGEON starts the observational learning process with a random initial seed in the neural network. The experiential learning process (practice), however, begins with the best agent trained


```

Ensure: Playing Field (3x3)
Ensure: Destination Square (0.5x0.5)
1: Captured = 0
2: BoxScore = 0.0
3: for i < Boxes do
4:   if Boxi is in Designated Area then
5:     Captured = Captured + 1
6:     Distancei = 0.0
7:   else
8:     Distancei = Manhattan(Boxi)
9:     BoxScore = BoxScore + (3.75-Distancei) / 3.75
10:  end if
11: end for
12: if Captured = |Boxes| then
13:   BoxNormal = BoxScore / (|Boxes|-Captured)
14: else
15:   BoxNormal = 0.0
16: end if
17: TotalFitness = 100 * (Captured + BoxNormal) /
   |Boxes|
    
```

Fig. 4. Algorithm 2: Performance Rating calculation in crane.

TABLE 1
Observational Learning—Resulting Similarity Factor Compared to Actor per 1,000 and 10,000 Generations

Entity	Avg. - 1,000	Best - 1,000	Avg. - 10,000	Best - 10,000
LATA	83.8	85.8	84.9	87.1

TABLE 2
Observational Learning—Resulting Performance Ratings

Entity	Average	Best	Standard Dev.
LATA agent	64.6	73.3	1.5
Human Expert	77.8	95.4	10.1

through observation, and seeks to improve the Performance Rating of that agent.

In our observational learning, the PIGEON algorithm was run on observed data for the human actor for 1,000 generations. The average Similarity Factor was 83.8 while the highest was 85.8 (out of a possible maximum of 100). For 10,000 generations, the average improved slightly to 84.9 while the best agent improved to 87.1. See Table 1. Therefore, the agent was able to learn to match the actions of the actor reasonably well when presented with the same environmental input values that were seen by the actor.

Table 2 displays the Performance Rating of the best performing agent (100 is the maximum score attainable). By any definition, these agents can be considered competent performers, although not quite as good as the human actor.

6.2 Experiential Learning

The LATA agents trained through observation were subsequently subjected to experiential learning (practice) to improve their proficiency. PIGEON was used in this learning process as well. This resulted in a more competent agent, not so much on the average but on the best run by the best performing agent. Table 3 indicates this.

TABLE 3
Experiential Learning—Resulting Performance Ratings

Entity	Average	Best	Standard Dev
LATA	64.9	96.1	20.4
Human Expert	77.8	95.4	10.1

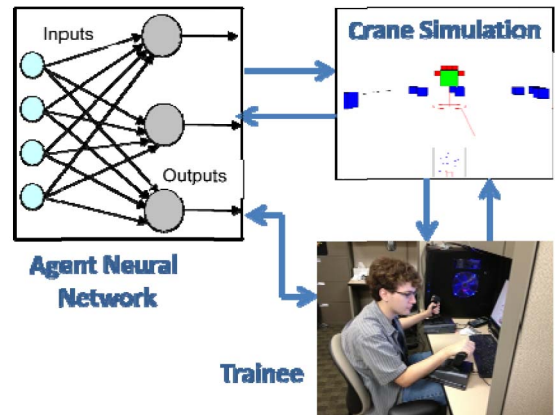


Fig. 5. LATA agent in the process of teaching a trainee.

Using these agents trained to competence by the above methods, we now turn to see how they were used to train humans, and whether their use made any difference in human training

7 PHASE II—TRANSFERRING LEARNING VIA HAPTIC FEEDBACK

We now describe Phase II—determination of whether skills can be transferred successfully from a LATA agent to a less proficient human via haptic-based coaching.

7.1 Teaching Humans via Coaching

Feedback is provided “haptically” in real time while the human is operating the crane, and not through a score computed at the end of the run, as in verification feedback. The LATA agent acting as a coach reacts in real time to the needs of the task and the trainee’s actions when controlling the crane, and provides counterforce corrections to the trainee based on what he/she does at every moment in real time. Fig. 5 depicts the arrangement of the trainee and the LATA.

The underlying concept of our teaching LATA agent is to provide real-time corrections. An analogy is the student-driver car that teenagers typically use in high school driver education, where the student and the teacher have identical setups except that the teacher’s steering wheel and brakes always override the actions of the student’s. If the student did something wrong, the teacher could provide corrective input to show proper performance (and sometimes avoid an accident!). In our application, the teaching LATA agent is the teacher and the “identical setup” is the very same haptic device now controlled simultaneously by the trainee and by the LATA agent. If the trainee performs an action deemed improper, the LATA agent provides feedback by physically forcing the joystick toward the desired (and correct)

position, thereby actively and forcefully countering the actions of the trainee. The trainee will notice this and accept it as constructive feedback. For all experiments described, the joystick is the main interface device for bidirectional communication—the human trainee uses it to control the crane and the LATA agent uses it to indicate to the trainee how well it is performing the task.

7.2 Force Feedback

The LATA coach compares the position at which the human has set the actual joysticks to the positions of “virtual joysticks” computed by the LATA agent. Between these two positions, a force vector can be drawn from one point to the other. If we treat the joystick like a virtual spring, the farther these two points are from one another, the larger the force that would be given. Alternatively, if the points were close to one another, the force would be minimal. This computed force vector is then broken up into its components for each axis and a command is given to an electric motor connected to each axis of the joystick independently. This motor in the joystick is strong enough to provide many pounds of force at the handle.

7.3 Experimental Plan and Experiments

A set of experiments was designed to evaluate the merits of the LATA system acting as a coach on novice human trainees. The first experiment, called the *Proficiency Test*, evaluated the overall performance level achieved by the human test subjects. The second experiment, called the *Learning Rate Test*, determined how quickly the test subjects could learn a task to a certain level of competence with the help of the LATA coach. Both experiments gauged how much knowledge was gained using the LATA coaching system when compared against self-learning. In both cases, a control group of test subjects who learned the skills strictly through unsupervised practice served as a benchmark for comparison.

The essence of this evaluation was to determine whether the LATA system could train a set of novice human test subjects to a higher performance level (as measured by the Performance Rating), and how quickly could the LATA system train an individual to a fixed competency level. Since each of these experiments assumes novice human test subjects, there is no overlap between any control and test group; furthermore, there is no overlap between the Proficiency Test group and the Learning Rate Test group. Test subjects were volunteers drawn from a pool of students in the College of Engineering and Computer Science at our home institution. They were assigned to the test and control group based on a coin flip. All experiments took place in a closed cubicle without any external interaction. The experiments took place over the course of approximately one month, with several different test subjects performing the exercise each day. Both groups received initial verbal instruction on what each joystick did and an explanation of the crane simulation in terms of objects, colors, and goals.

7.3.1 Proficiency Test Plan

Twenty-eight volunteers were used for the Proficiency Test. They were randomly partitioned into two groups of 14. For both groups, the experiment required that each test subject perform 20 trials of the crane simulation to pick up and

deliver as many boxes as possible to the designated area in a 60-second span. The test subjects were given 1 minute of rest between trials. For the control group, all 20 trials of the crane simulation were unsupervised. That is, the test subject could gain experience by practicing with the crane simulator as he/she wished.

For the test group, the first five trial runs were unsupervised, as were for the control group. For the next 10 trial runs, however, the test group subjects would undergo coaching by the LATA coach through the joysticks. Finally, the last five runs were independently operated again. Since the first five trial runs for both groups were unsupervised, they were used to set the baseline performance of each individual (the “before” runs). Then, after 10 training runs by either unsupervised practice (control group) or haptic LATA coaching (test group), the final five trial runs were used as the measure for the skill level gained by the individual test subjects (the “after” runs). To normalize the data, the baseline performance was subtracted from the final performance to adjust for the natural abilities of the test subject.

7.3.2 Learning Rate Test Plan

Twenty human test subject volunteers were used in the Learning Rate test. They were also equally distributed between two groups of 10 individuals each. No test subject that participated in the Proficiency Test was used in the Learning Rate Test. To calculate the learning rate, the test subjects were to be trained up to the performance level of the LATA coach—retrieving seven boxes in 60 seconds. The subjects would perform trial runs as required until they were able to capture seven boxes on their own. For the sake of practicality, there was an imposed limit of 40 trial runs, but this limit was never reached. Test subjects in the control group were simply allowed to operate unsupervised until they captured seven boxes. The subjects in the test group would alternate between operating unsupervised (for one trial) then be coached by LATA on the next trial. This allowed the trainee to get a feel for the joystick as guided by the LATA coach, then try it for themselves on every other trial.

7.3.3 Qualitative Evaluation—Questionnaire

In addition to quantitative evaluations done above, qualitative metrics were also recorded. Each test subject in the Proficiency Test experiment was given an anonymous questionnaire after completing his or her experiment. The following questions were on the questionnaire (rated on a scale 1 to 5, where 5 is most positive):

- Do you feel you improved over the 20 runs?
- Would you feel you are an expert at this task?
- Do you feel the computer haptics helped in training (test group only)?

Note that this was only given to those subjects that participated in the Proficiency Test. Because all subjects in the Learning Rate test achieved the same proficiency, we felt that it was not as relevant and did not administer this questionnaire to them.

7.4 Experimental Results

All trials for the Proficiency and the Learning Rate tests took place on the same computer workstation with the same

TABLE 4
Proficiency Test Results

Proficiency Rating Range	Control Group "Before" runs	Control Group "After" runs	Test Group "Before" runs	Test Group "After" runs
0 - 15	0	0	5	1
15 - 25	6	0	5	1
25 - 35	16	2	11	7
35 - 45	20	4	23	3
45 - 55	18	11	14	8
55 - 65	8	26	9	22
65 - 75	1	21	3	17
75 - 85	1	6	0	9
85 - 95	0	0	0	2
>95	0	0	0	0

joysticks. The test subjects were all engineering or computer science undergraduate students. Although joystick movements were recorded to disk and each subject filled out a questionnaire, no personal identifying information was kept.

Because the test groups were small, the samples cannot assume a normal distribution. Therefore, the comparisons within groups and between groups were calculated with rank statistics. Rank statistics computes its values from the median and relative rank of the values. For paired tests (such as the same individual before and after), the Wilcoxon signed-rank test (*signrank*) is used when the sample size is less than 20. For unpaired tests (such as comparing two independent groups), the Wilcoxon rank-sum test (*ranksum*) is used because it is robust with respect to outliers and works with population sizes of less than 20.

7.4.1 Proficiency Test Results

The data from the Proficiency Test experiment from both groups in the "before" and "after" cases can be seen in Table 4, a tabular histogram of how many runs scored at what level of proficiency. The maximum proficiency is 100. There are 70 runs total in each column (14 subjects \times 5 runs each). From this histogram, we can see that there was clear improvement in both groups, where the average performance was able to improve approximately 20 points better between the beginning and ending trials. However, an eyeball comparison of the magnitude of improvement is not precise and requires statistical analysis.

The first determination to make was whether the two samples were from the same population by comparing the performance of both groups from the "before" baseline runs. According to the rank-sum test, the p-value of the two groups is 0.945, which is not less than the 0.05 needed to reject the Null-Hypothesis that the two groups are initially *the same*. Therefore, they are deemed to be from the same population.

The second determination was to confirm that there was improvement from "before" to "after." The signed-rank test for both groups in a paired test between the before and after is calculated to be 0.00012, which is much less than 0.05. This rejects the Null-Hypothesis that there was *no change* between the before and after performances. Thus, performance gain is verified for both groups.

Third, we quantify the performance gains by both groups and compare them. To compare the performance

gains by the control and test groups, these were tested against each other in the "after" training set. To assess improvement, the baseline scores were subtracted from the final scores then divided by the baseline to normalize. The results indicate that the median individual was able to improve 40 percent beyond his/her original performance level. However, the p-value of 0.80 of the rank-sum check between the control and test groups could not reject the Null-Hypothesis that the LATA coaching method did not improve performance gain over self-learning.

It is interesting to note, nevertheless, that from Table 3 we can see that the human expert who served as the actor in observational learning was able to attain a proficiency rating of nearly 78. The control group was able to attain such a high proficiency in only six runs, while the test group was able to do it in 11 runs, nearly double.

7.4.2 Learning Rate Test Results

The learning rate experiment examines how many trial sessions it would take to reach the specific proficiency level of picking up and delivering seven boxes in the 60-second time period. This is the same performance level achieved by the LATA agent coach. The median individual took an average of 7.5 training sessions to reach that performance level. According to the rank-sum test, the comparison between the learning rates was 0.285 and was not significantly low to a 0.05 value to reject the Null-Hypothesis that the use of the LATA haptic system *did not* improve the learning rate of the human subjects. Therefore, it is likely that the amount of time required for improvement is not different between the control and test groups.

However, because half of the time for the test group is considered coaching during which they were not in full control of the crane (the LATA agent was partially in control), one could argue that the amount of time to learn could be considered to be halved. If so, then the p-value on the rank-sum test would be 0.0137, which is now less than 0.05. This would reject the Null-Hypothesis and support the hypothesis that the test subjects learned more quickly with LATA haptic training.

7.4.3 Qualitative Evaluation Results—Questionnaire

The test subjects from the Proficiency Test experiment (only) were asked a series of questions on an anonymous questionnaire. This questionnaire was designed to capture the opinions of the test subjects about the experiment and about themselves. The first question was whether they thought they improved over the course of the 20 runs. The control group averaged 3.8 versus 4.5 for the test group. Both groups believed that they had some improvement but the rank-sum test had a p-value of 0.0167, which rejects the Null-Hypothesis that the groups improved the same. Therefore, evidence suggests that the test group believed they improved more.

When asked whether they felt they were an expert at the crane simulation at the end of the training, both groups slightly agreed with 3.4 and 3.9 for control and test groups, respectively. The p-value on the rank-sum was 0.247, which does not reject the Null-Hypothesis that the test group did not improve more than the control group. Evidence suggests, therefore, that both groups believe they became experts to the same degree.

The final question was whether the subjects in the test group (only) believed that the haptics helped in the learning process. The participants responded with an average value of 3.7, which is higher than the neural value of 3.00. There is some evidence that the students believed that the haptic feedback helped in the learning process.

7.5 Analysis

There is evidence to suggest that the test group subjects were able to use the LATA coach to gain a high proficiency level. The test group subjects were also able to learn up to the level of the LATA coach reasonably fast. However, there is mixed evidence as to whether LATA coaching was able to improve beyond a human's ability to learn a psychomotor skill task on his/her own through practice. Some evidence suggests that test group members were able to perform at the level of the human expert more frequently than were control group subjects. The quantitative evidence for the Learning Rate Test, when adjusted for the coaching runs, supports the premise that the LATA coaching indeed resulted in improved gains over the control group subjects. Furthermore, the qualitative data support this as well based on the responses from the test group subjects. However, the Proficiency Test results do not support this claim, as neither does the unadjusted Learning Rate Test data.

The reason for the mixed results could be traced to the test domain rather than a weakness with the LATA coaching concept. As seen in the Learning Rate experiment, the human test subjects could learn to a base competency level of seven boxes retrieved in approximately 7.5 trials. Since the experiment for Proficiency was 20 trials, the humans in the control group were able to learn up to that proficiency level before the first 10 trials. This indicates that the skills were significantly easier to master than we had anticipated. Further research is warranted with a task is harder to accomplish and would therefore take longer than 20 trials to master. Additionally, the experiment could have been made more difficult by requiring the achievement of nine-block level that an expert can achieve, which would have taken more time.

8 CONCLUSIONS

Our presented LATA system builds a computer agent trained to the level of the expert using observational learning and practice; then, it uses that same agent and haptic interface to teach less proficient human novices. The goals of this research were to prove technical feasibility of such a system, and show that it has educational value. The concept was limited to learning complex psychomotor skill tasks such as would be found in operating a crane in a simulated shipyard, moving boxes around.

In the context of other research, while there is significant work reported in the provision of feedback to the student, haptics (both for teaching as well as for actual use), learning agents, teaching agents, and psychomotor skill learning, no other work combined all the aspects of providing feedback in real time though a haptic interface for a system that employs agents to teach psychomotor skills that were developed though machine learning.

From a technical standpoint, we were successful in building this system in a practical manner. While not yet ready for real-world applications, the system required

modest amounts of human intervention to build the learning agent. We found that the PIGEON learning algorithm was able to teach a LATA agent to a high level of performance through automatic observation of the human's actions in the operation of a crane, and then subsequent practice. The resulting LATA agent was able to perform at close to (but not quite) the level of the original human expert from whom it learned.

The learning agent—now considered a teaching agent—was then evaluated for its ability to teach other humans using force-feedback joysticks as the interface. Through a study of both gains in proficiency and learning rate, humans test subjects were shown to be able to learn up to the ability of the LATA agent. However, it could not be statistically shown that this caused the human to improve more than learning on his/her own. Nevertheless, there is some evidence to suggest that the humans could learn more quickly with LATA, although that could not be entirely confirmed based on the experiments performed. According to the test subjects, there is evidence that they believed that they improved to a higher level with the haptic training and that they became experts themselves.

In summary, we found only weak evidence to suggest that the teaching agent has educational value above that of self-learning, at least as related to the selected crane manipulation task. This inability to show conclusively that LATA was able to train a human better than a self-learning process on the Crane simulation was disappointing. However, on the positive side, the LATA agents were able to be created nearly automatically (with little human assistance), learn the skill nearly to the level of an expert, and then be used to train a novice human to that same level. The benefits of wide distribution of qualified trainers through an agent that can be created (nearly) automatically without explicit programming are significant. Therefore, we strongly believe that further research is warranted based on our results.

In conclusion, our work was technically successful and did perform the learning and teaching functions. Furthermore, the educational value of our approach showed some potential as well as some positive feedback. However, its value in improving learning beyond current means remains to be proven conclusively.

We believe that further research on this promising avenue of teaching and learning is warranted. One should begin with subjecting the current prototype to a more suitable set of tests (i.e., more challenging) than the Crane simulation. That would help establish more clearly the utility of this approach. Flying an airplane (simulator) could be one such domain for further testing. Surgery could be another.

ACKNOWLEDGMENTS

This research was supported by the US National Science Foundation under grant number IIS-0948820.

REFERENCES

- [1] J.R. Carbonell, "AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction," *IEEE Trans. Man-Machine Systems*, vol. 11, no. 4, pp. 190-202, Dec. 1970.
- [2] B.P. Woolf, "Context-Dependent Planning in a Machine Tutor," doctoral dissertation, Dept. of Computer and Information Science, Univ. of Massachusetts, Amherst, 1984.

- [3] B.P. Woolf and D.D. McDonald, "Context-Dependent Transitions in Tutoring Discourse," *Proc. Nat'l Conf. Artificial Intelligence*, pp. 355-361, 1984.
- [4] H.K. Holden, R.A. Sottolare, B.S. Goldberg, and K.W. Brawner, "Effective Learner Modeling for Computer-Based Tutoring of Cognitive and Affective Tasks," *Proc. Interservice/Industry Training, Simulation, and Education Conf. (IITSEC '12)*, Nov. 2012.
- [5] J. Beck and B.P. Woolf, "High Level Student Modeling with Machine Learning," *Proc. Int'l Conf. Intelligent Tutoring Systems*, 2000.
- [6] C. Ramos, C. Frasson, and S. Ramachandran, "Introduction to the Special Issue on Real World Applications of Intelligent Tutoring Systems," *IEEE Trans. Learning Technologies*, vol. 2, no. 2, pp. 62-63, Apr.-June 2009.
- [7] L.D. Interrante and J.E. Biegel, "Reasoning in Time and Space: Issues in Interfacing a Graphic Computer Simulation with an Expert System for an Intelligent Simulation Training System," *Proc. Winter Simulation Conf.*, pp. 1106-1109, Jan. 1989.
- [8] S.S. Mulgand, M. Asdigha, G.L. Zacharias, K. Krishnakumar, and J.A. Dohme, "An Intelligent Tutoring System for Simulator-Based Helicopter Flight Training," *Proc. Flight Simulation Technologies Conf.*, 1995.
- [9] E. Remolina, S. Ramachandran, D. Fu, R. Stottler, and W.R. Howse, "Intelligent Simulation-Based Tutor for Flight Training," *Proc. Interservice/Industry Training, Simulation, and Education Conf. (IITSEC '04)*, 2004.
- [10] *The Encarta Dictionary of North America*. Microsoft Corporation, 1999.
- [11] *Webster's New World Dictionary*. World Publishing, 1957.
- [12] J. Whitmore, *Coaching for Performance*, fourth ed. Nicholas Brealy, 2009.
- [13] E.A. Bilodeau and I.M. Bilodeau, "Motor-Skills Learning," *Ann. Rev. of Psychology*, vol. 12, pp. 243-280, Feb. 1961.
- [14] E. Todorov, R. Shadmehr, and E. Bizzi, "Augmented Feedback Presented in a Virtual Environment Accelerates Learning of a Difficult Motor Task," *J. Motor Behavior*, vol. 29, no. 2, pp. 147-158, June 1997.
- [15] R.W. Kulhavy and W.A. Stock, "Feedback in Written Instruction: The Place of Response Certitude," *Educational Psychology Rev.*, vol. 1, no. 4, pp. 279-308, 1989.
- [16] B.J. Mason and R. Bruning, "Providing Feedback in Computer-Based Instruction: What the Research Tells Us," <http://dwb.unl.edu/Edit/MB/MasonBruning.html>, 2001.
- [17] R.R. Burton and J.S. Brown, "An Investigation of Computer Coaching for Informal Learning Activities," *Int'l J. Man-Machine Studies*, vol. 11, no. 1, pp. 5-24, Jan. 1979.
- [18] K.C. Deesen and A.G. Tilak, "Computer Assisted Coaching Method," US patent 5,100,329, Washington, D.C.: Patent and Trademark Office, 1992.
- [19] S. Dickson and D. Parsons, "The Odd Couple—Skill Learning and Computer Software," *British J. Educational Technology*, vol. 29, no. 4, pp. 371-374, 1998.
- [20] M.A. Constantino-González and D.D. Suthers, "Coaching Collaboration in a Computer-Mediated Learning Environment," *Proc. Conf. Computer Support for Collaborative Learning: Foundations for a CSCLE Community*, pp. 583-584, 2002.
- [21] A.Y. Zissos and I.H. Witten, "User Modelling for a Computer Coach: A Case Study," *Int'l J. Man-Machine Studies*, vol. 23, no. 6, pp. 729-750, Dec. 1985.
- [22] M. Acovelli and M. Gamble, "A Coaching Agent for Learners Using Multimedia Simulations," *Educational Technology*, vol. 37, no. 2, pp. 44-48, Mar./Apr. 1997.
- [23] X. Geng, Z. Zhang, Y. Jiang, and Y. Yang, "An Intelligent Tutoring Architecture for Scenario-Based Flight Training," *Proc. Second Int'l Workshop Education Technology and Computer Science*, pp. 291-294, Mar. 2010.
- [24] G.K.W.K. Chung, S.O. Nagashima, G.C. Delacruz, J.J. Lee, R. Wainess, and E.L. Baker, "Review of Rifle Marksmanship Training Research," Report #783, Nat'l Center for Research on Evaluation, Standards, and Student Testing, Univ. of California, Los Angeles, Jan. 2011.
- [25] K. Belghith, R. Nkambou, F. Kabanza, and L. Hartman, "An Intelligent Simulator for Telerobotics Training," *IEEE Trans. Learning Technologies*, vol. 5, no. 1, pp. 11-19, Jan.-Mar. 2012.
- [26] R. Jensen, M. Nolan, and D.Y. Chen, "Automatic Causal Explanation Analysis for Combined Arms Training AAR," *Proc. Interservice/Industry Training Simulation and Education Conf.*, Nov. 2005.
- [27] J.R. Surdu and U.W. Pooch, "Rational Agents, Simulation, and Military Operations," *Proc. AI, Simulation, and Planning Conf. (AIS '00)*, pp. 327-332, Mar. 2000.
- [28] R. Jensen, C. Tasoluk, M. Sanders, and H. Marshall, "FCS Intelligent Structured Training—Experimental Result and Future Applications," *Proc. Interservice/Industry Training Simulation and Education Conf.*, Nov. 2005.
- [29] R.G. Abbott, "Automated Expert Modeling for Automated Student Evaluation," *Proc. Eighth Int'l Conf. Intelligent Tutoring Systems (ITS '06)*, pp. 26-30, June 2006.
- [30] H. Fernlund, A.J. González, J. Ekblad, and A. Rodriguez, "Trainee Evaluation through after Action Review by Comparison," *J. Defense Modeling and Simulation*, vol. 6, pp. 135-150, July 2009.
- [31] M. Srinivasan and C. Basdogan, "Haptics in Virtual Environments: Taxonomy, Research Status, and Challenges," *Computers & Graphics*, vol. 21, no. 4, pp. 393-404, 1997.
- [32] A.M. Howard, C.H. Park, and S. Remy, "Using Haptic and Auditory Interaction Tools to Engage Students with Visual Impairments in Robot Programming Activities," *IEEE Trans. Learning Technologies*, vol. 5, no. 1, pp. 87-95, Jan.-Mar. 2012.
- [33] K. Tanoue, S. Ieiri, K. Konishi, T. Yasunaga, and K. Okazaki, "Effectiveness of Endoscopic Surgery Training for Medical Students Using a Virtual Reality Simulator Versus a Box Trainer: A Randomized Controlled Trial," *Surgical Endoscopy*, vol. 22, no. 4, pp. 985-990, Apr. 2008.
- [34] T.P. Grantcharov, V.B. Kristiansen, J. Bendix, L. Bardram, J. Rosenberg, and P. Funch-Jensen, "Randomized Clinical Trial of Virtual Reality Simulation for Laparoscopic Skills Training," *British J. Surgery*, vol. 91, pp. 146-150, 2004.
- [35] W.R. Chitwood, L.W. Nifong, W.H.H. Chapman, J.E. Felger, B.M. Bailey, T. Ballint, K.G. Mendelson, V.B. Kim, J.A. Young, and R.A. Albrecht, "Robotic Surgical Training in an Academic Institution," *Annals of Surgery*, vol. 234, no. 4, pp. 475-486, Oct. 2001.
- [36] O.A.J. van der Meijden and M.P. Schijven, "The Value of Haptic Feedback in Conventional and Robot-Assisted Minimal Invasive Surgery and Virtual Reality Training: A Current Review," *Surgical Endoscopy*, vol. 23, no. 6, pp. 1180-1190, June 2009.
- [37] E.P. Westebring-van der Putten, R.H. Goossens, J.J. Jakimowicz, and J. Dankelman, "Haptics in Minimally Invasive Surgery—A Review," *Minimally Invasive Therapy & Allied Technologies*, vol. 17, no. 1, pp. 3-16, 2008.
- [38] A.M. Okamura, "Haptic Feedback in Robot-Assisted Minimally Invasive Surgery," *Current Opinion in Urology*, vol. 19, no. 1, pp. 102-107, Jan. 2009.
- [39] D. Shaffer, D. Meglan, M. Ferrell, and S. Dawson, "Virtual Rounds: Simulation-Based Education in Procedural Medicine," *Proc. SPIE Battlefield Biomedical Technologies Conf.*, vol. 3712, pp. 99-108, 1999.
- [40] G. Burdea, G. Patounakis, and V. Popescu, "Virtual Reality-Based Training for the Diagnosis of Prostate Cancer," *IEEE Trans. Biomedical Eng.*, vol. 46, no. 10, pp. 1253-60, Oct. 1999.
- [41] F.P. Brooks Jr., O.-Y. Ming, J.J. Batter, and P.J. Kilpatrick, "Project GROPE: Haptic Displays for Scientific Visualization," *ACM SIGGRAPH Computer Graphics*, vol. 24, no. 4, pp. 177-185, 1990.
- [42] R.L. Williams II, M. Srivastava, R.R. Conatser Jr., and J.N. Howell, "The Virtual Haptic Back Project," *Proc. Image Soc. Conf.*, July 2003.
- [43] A. Nijholt, S. Kole, and J. Zwiers, "Multimodal Interaction in a Haptic Environment," *Proc. First Joint Eurohaptics Conf. and Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 467-470, 2005.
- [44] T. Lloyd and S. Bull, "A Haptic Learner Model," *Int'l J. Continuing Eng. Education and Lifelong Learning*, vol. 16, nos. 1/2, pp. 137-149, 2006.
- [45] R.L. Williams, X. He, T. Franklin, and S. Wang, "Haptics-Augmented Undergraduate Engineering Education," *Proc. Int'l Conf. Eng. Education*, 2004.
- [46] A.M. Okamura, C. Richard, and M.R. Cutkosky, "Feeling Is Believing: Using a Force-Feedback Joystick to Teach Dynamic Systems," *J. Eng. Education*, pp. 345-349, July 2002.
- [47] D. Morris, H. Tan, F. Barbalgi, T. Chang, and K. Salisbury, "Haptic Feedback Enhances Force Skill Learning," *Proc. Second Joint EuroHaptics Conf. and Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2007.
- [48] D. Feygin, M. Keehner, and F. Tendick, "Haptic Guidance: Experimental Evaluation of a Haptic Training Method for a Perceptual Motor Skill," *Proc. 10th Symp. Haptic Interfaces for Virtual Environments & Teleoperator Systems*, 2002.

- [49] J. Lee and S. Choi, "Effects of Haptic Guidance and Disturbance on Motor Learning: Potential Advantage of Haptic Disturbance," *Proc. IEEE Haptics Symp.*, Mar. 2010.
- [50] A.L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," *IBM J.*, vol. 3, no. 3, pp. 211-229, 1959.
- [51] T.G. Dietterich, "Machine Learning for Sequential Data: A Review," *Proc. Joint IAPR Int'l Workshop Structural, Syntactic, and Statistical Pattern Recognition*, pp. 15-30, 2002.
- [52] M.A. Bauer, "Programming by Examples," *Artificial Intelligence*, vol. 12, no. 1, pp. 1-21, 1979.
- [53] C. Sammut, S. Hurst, D. Kedzier, and D. Michie, "Learning to Fly," *Proc. Ninth Int'l Conf. Machine Learning*, pp. 385-393, 1992.
- [54] T.A. Sidani and A.J. González, "A Framework for Learning Implicit Expert Knowledge through Observation," *Trans. Soc. for Computer Simulation*, vol. 17, no. 2, pp. 54-72, June 2000.
- [55] D. Pomerleau, "ALVINN: An Autonomous Land Vehicle in a Neural Network," *Advances in Neural Information Processing Systems 1*, D. Touretzky, ed., Morgan Kaufmann, 1989.
- [56] H.K.G. Fernlund, A.J. González, M. Georgiopoulos, and R.F. DeMara, "Learning Tactical Human Behavior through Observation of Human Performance," *IEEE Trans. Systems, Man, and Cybernetics, Part B*, vol. 36, no. 1, pp. 128-140, Feb. 2006.
- [57] S. Ontañón, K. Mishra, N. Sugandh, and A. Ram, "On-Line Case-Based Planning," *Computational Intelligence J.*, vol. 26, no. 1, pp. 84-119, 2010.
- [58] D.C. Bentevegna and C.G. Atkeson, "Learning from Observation Using Primitives," *Proc. Int'l Conf. Robotics and Automation*, pp. 1988-1993, 2001.
- [59] G. Stein, "FALCONET: Force-Feedback Approach for Learning from Coaching and Observation Using Natural and Experiential Training," doctoral dissertation, Computer Eng., Univ. of Central Florida, Aug. 2009.
- [60] G. Stein and A.J. Gonzalez, "Building High-Performing Human-Like Tactical Agents through Observation and Experience," *IEEE Trans. Systems, Man and Cybernetics, Part B*, vol. 41, no. 3, pp. 792-804, June 2011.
- [61] M. Nechyba and Y. Xu, "Human Skill Transfer: Neural Networks As Learners and Teachers," *Proc. IEEE Int'l Conf. Intelligent Robots Systems*, pp. 314-319, 1995.
- [62] A. Sano, H. Fujimoto, and K. Matsushita, "Machine Mediated Training Based on Coaching," *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics*, 1999.
- [63] K. Stanley and R. Miikkulainen, "Evolving Neural Networks through Augmenting Topologies," *Evolutionary Computation*, vol. 10, no. 2, pp. 99-127, 2002.
- [64] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," *Proc. IEEE Int'l Conf. Neural Networks*, 1995.



Gary Stein received the PhD degree in computer engineering in 2009 from the University of Central Florida. He is currently employed by Primal Innovation, Inc. in Sanford, Florida. His research focuses on machine-learning techniques to replicate or replace human operators in both simulated and robotic entities. He is a member of the IEEE.



Avelino J. Gonzalez received the PhD degree in electrical engineering in 1979 from the University of Pittsburgh. He is a professor in the Department of Electrical Engineering and Computer Science at the University of Central Florida, Orlando. His area of expertise is intelligent systems and machine learning. He is a fellow of the IEEE.



Clayton Barham is an undergraduate student in computer science at the University of Central Florida.