03.18

Computer

E-COACHING FOR HEALTH

IEEE

IEEE computer society

vol. 51 no. 3

www.computer.org/computer

Computer

MARCH 2018

FEATURES

MARCH 2018
CONTENTS

ABOUT THIS ISSUE
E-COACHING FOR HEALTH

As intelligent systems, e-coaches exhibit abilities similar to their human counterparts: sensing, learning, and actuating.

See **www.computer.org/computer-multimedia** for multimedia content related the features in this issue.

Circulation: *Computer* (ISSN 0018-9162) is published monthly by the IEEE Computer Society. IEEE Headquarters, Three Park Avenue, 17th Floor, New York, NY 10016-5997; IEEE Computer Society Publications Office, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720; voice +1 714 821 8380; fax +1 714 821 4010; IEEE Computer Society Headquarters, 2001 L Street NW, Suite 700, Washington, DC 20036. IEEE Computer Society membership includes a subscription to *Computer* magazine.

Postmaster: Send undelivered copies and address changes to *Computer*, IEEE Membership Processing Dept., 445 Hoes Lane, Piscataway, NJ 08855. Periodicals Postage Paid at New York, New York, and at additional mailing offices. Canadian GST #125634188. Canada Post Corporation (Canadian distribution) publications mail agreement number 40013885. Return undeliverable Canadian addresses to PO Box 122, Niagara Falls, ON L2E 6S8 Canada. Printed in USA.

FEATURES CONTINUED

COMPUTING PRACTICES

COLUMNS

Departments

Membership News

Computer

IEEE COMPUTER SOCIETY http://computer.org // +1 714 821 8380
COMPUTER http://computer.org/computer // computer@computer.org

ELSEWHERE IN THE CS

EDITOR LORI CAMERON
l.cameron@computer.org

Computer Highlights Society Magazines

The IEEE Computer Society's lineup of 13 peer-reviewed technical magazines covers cutting-edge topics in computing, including scientific applications, Internet computing, machine intelligence, pervasive computing, security and privacy, digital graphics, cloud computing, and computer history. Here, we highlight recent issues of other Computer Society magazines.

computing in SCIENCE & ENGINEERING

CoFlaVis: A Visualization System for Pulverized Coal Flames

One problem concerning researchers of various domains, such as chemistry and fluid dynamics, is the optimization of coal combustion processes to increase coal-related systems' efficiency, safety, and cleanliness. Researchers examine the combustion process by using complex simulations that normally yield highly complex data comprising many characteristics. Scientists employ such datasets to validate their hypotheses or to present new hypotheses. And the analysis is mostly restricted to time-consuming workflows capable of handling only a portion of the data. To support the experts, various suppliers have developed interactive visualization and analysis tools to manage and understand multivariate data. This article from *CiSE*'s November/December 2017 issue demonstrates how one of these tools can improve data exploration of pulverized coal combustion.

IEEE Annals of the History of Computing

The "IBM Family": American Welfare Capitalism, Labor, and Gender in Postwar Germany

This article from *Annals*' October–December 2017 issue examines IBM's corporate labor and gender relations from a transatlantic perspective. It argues that IBM chair and CEO Thomas Watson Sr. shaped labor relations in his company's West German subsidiary. In the US, Watson acted as a business progressive, expanding operations, opening professional careers to young women, and implementing welfare capitalist measures. When IBM took tighter control of its foreign operations after World War II, Watson sought to implement welfare capitalist measures in the companies' overseas subsidiaries. With his wife by his side, he presented himself as the caring paterfamilias. German IBM employees embraced Watson's rhetoric but continued to join national unions and formed a workers' council, thwarting the major welfare capitalist goal of averting labor organization. Against such local labor practices, Watson undergirded a loyal workforce even in critical situations, an overlooked factor contributing to the company's success.

IEEE CLOUD COMPUTING

Intelligent Resource Management in Blockchain-Based Cloud Datacenters

Today, more and more companies migrate business from their own servers to the cloud. With the influx of computational requests, datacenters consume tremendous amounts of energy every day, attracting great attention in energy-efficiency circles. In this article from *Cloud Computing*'s November/December 2017 issue, researchers investigate the energy-aware resource-management problem in cloud datacenters, to which green energy with unpredictable capacity is connected. By proposing a robust blockchain-based decentralized resource-management framework, researchers save the energy that the request scheduler otherwise consumes. Moreover, they propose a reinforcement learning method embedded in a smart contract to further minimize energy costs. Because the reinforcement learning method is informed by historical knowledge, it doesn't rely on request arrival and energy

0018-9162/18/$33.00 © 2018 IEEE

supply. Experimental results on Google cluster traces and real-world electricity prices show that the approach reduces datacenter costs significantly.

IEEE Computer Graphics and Applications

Visual Communication and Cognition in Everyday Decision Making

Visual communication's role has quickly changed. And new materials and technology have helped create new commercial-art and graphic-design approaches. From ancient writing systems to Johannes Gutenberg's development of metal movable type and the start of the printing revolution, materials and technology have created opportunities for visual communication to reach more people and share more diverse messaging faster than ever before. Learn more in this article from *CG&A*'s November/December 2017 issue.

IEEE Intelligent Systems

Robust Tracking of Soccer Robots Using Random Finite Sets

Maintaining a good estimation of other participating robots' positions is crucial in soccer robotics, as in most multirobot applications. Classical approaches use a vector representation of the robots' positions and Bayesian filters to propagate them over time. However, these approaches suffer from the data-association problem. To tackle this challenge, this article from *Intelligent Systems*' November/December 2017 issue presents a new methodology for the robust tracking of robots based on the Random Finite Sets framework, which doesn't require any explicit data association. Moreover, the proposed methodology can integrate information shared by teammate robots, including their positions and their estimations of the other robots' positions. The proposed method reduces errors in the estimates of robots' positions by about 35 percent.

IEEE Internet Computing

Nowcasting of Earthquake Consequences Using Big Social Data

Messages posted to social media in the aftermath of a natural disaster have value beyond announcing the event itself. Mining such digital traces helps provide a timely estimate of the disaster's consequences on the population and on infrastructures. Yet, to date, researchers have paid little attention to such automatic assessments of disaster-related damage. In this article from *Internet Computing*'s November/December 2017 issue, the authors explore the process of feeding predictive models with tweets conveying on-the-ground social sensors' observations of earthquakes, and nowcasting the temblors' intensity.

IEEE micro

Ultra-Low-Power Processors

Society's increasing use of connected sensing and wearable computing has created robust demand for ultra-low-power (ULP) edge-computing devices and associated system-on-chip architectures. In fact, ULP processing's ubiquity has already made such embedded devices the processor component with the highest production volume, with even more market dominance expected in the near future. The Internet of Everything calls for an embedded processor in every object, necessitating billions of chips. At the same time, the explosion of data that these devices generate, in conjunction with the traditional model of using cloud-based services to process the information, will place tremendous demands on the limited wireless spectrum and energy-hungry wireless networks. Smart ULP edge devices are the only way to meet these demands. Learn more in this article from *Micro*'s November/December 2017 issue.

IEEE MultiMedia

Deep Learning Triggers a New Era in Industrial Robotics

Deep learning's pattern-recognition capabilities have pushed the limits in various fields, including industrial robotics. Deep learning alone won't solve all of this field's problems but will certainly improve robotics systems' perception capabilities, given its power to robustly recognize complex real-world patterns. The author of this article from *MultiMedia*'s October–December 2017 issue examines deep-learning applications for robotics.

IEEE pervasive COMPUTING

What Will We Wear After Smartphones?

Wearable computing research has been going on for 20 years. This survey from *Pervasive Computing*'s October–December 2017 issue looks back at how the field developed and explores where it's headed. According to the authors, wearable computing is entering its most exciting phase yet, as it transitions from demonstrations to the creation of sustained markets and industries, which in turn should drive future research and innovation.

IEEE SECURITY & PRIVACY

The Future of Digital Forensics: Challenges and the Road Ahead

Today's huge volumes of data, heterogeneous information and communication technologies, and borderless cyberinfrastructures create new challenges for security experts and law-enforcement agencies investigating cybercrimes. This article from *S&P*'s November/December 2017 issue explores digital forensics' future, with an emphasis on the challenges it faces and the advancements needed to effectively protect modern societies and pursue cybercriminals.

IEEE Software

Safe, Secure Executions at the Network Edge: Coordinating Cloud, Edge, and Fog Computing

System design in which cyberphysical applications are securely coordinated from the cloud could simplify the development process. However, this would push private data to the cloud, causing users to lose the control they have when they execute applications locally. Meanwhile, computing at the network edge is still lacking support for straightforward multidevice development, which is essential for a wide range of dynamic cyberphysical services. This article from *Software*'s January/February 2018 issue proposes a novel programming model. It also proposes a secure-connectivity framework for leveraging safe coordinated device proximity as an additional degree of freedom between the remote cloud and the safety-critical network edge, especially under uncertain environment constraints.

IT Professional

The Economics of "Fake News"

False information has economic, political, and social consequences. The authors of this article from *IT Pro*'s November/December 2017 issue analyze the real and perceived costs and benefits to people that create false information and those that provide them with platform support. The authors particularly consider digital advertising ecosystems that support fake-news creation. They also discuss the context of false-information consumption and suggest that fake-news creators, consumers, and arbiters can reinforce each other and form a vicious circle. The article proposes mechanisms to break the circle and alter this activity's cost-benefit structure. C

Low-Power Content Addressable Memory

Daniel J. Sorin, Duke University

This installment of Computer's series highlighting the work published in IEEE Computer Society journals comes from IEEE Computer Architecture Letters.

Content addressable memory (CAM) plays an important role in computer architecture, but architects have been unable to use it as widely as they would like because of its major disadvantage when compared to RAM—conventional NOR CMOS CAM is highly, and even prohibitively, power hungry.

In "Resistive Address Decoder" (*IEEE Computer Architecture Letters*, vol. 16, no. 2, 2017, pp. 141–144), authors Leonid Yavits, Uri Weiser, and Ran Ginosar propose a new CAM design—the resistive address decoder—that makes CAM power usage comparable to that of RAM and thus opens the door to much greater use of CAM.

Consider a typical NAND address decoder, where the address of a memory row is hardwired by connecting either address bit line or inverse bit line to the gate of an NMOS transistor. In a row where the address matches the hardwired pattern, all NMOS transistors are turned on and the row is selected. The authors add two resistive elements, such as memristors or spin-torque transfer magnetic RAM, to each NMOS transistor in a voltage-dividing manner (see Figure 1). Memristors are two-terminal devices whose resistance is changed by the electrical current. The resistance of the memristor typically has two states: R_ON (low resistive state) and R_OFF (high resistive state). By programming the left memristor of a bit cell to R_ON and the right one to R_OFF, the NMOS transistor gate is effectively connected to the address bit line. Conversely, by programming the left memristor to R_OFF while programming the right one to R_ON, the NMOS transistor gate is effectively connected to the inverse address bit line. The additional transistor in a bit cell is used to program the memristors. Such an approach achieves a two-fold effect: the address pattern in each row becomes programmable rather than hardwired, and the memristor pair forms an XOR gate, allowing comparison of the input address bit to the bit "programmed" into the resistive elements. These effects enable content addressability, effectively turning an address decoder into CAM.

0018-9162/18/$33.00 © 2018 IEEE

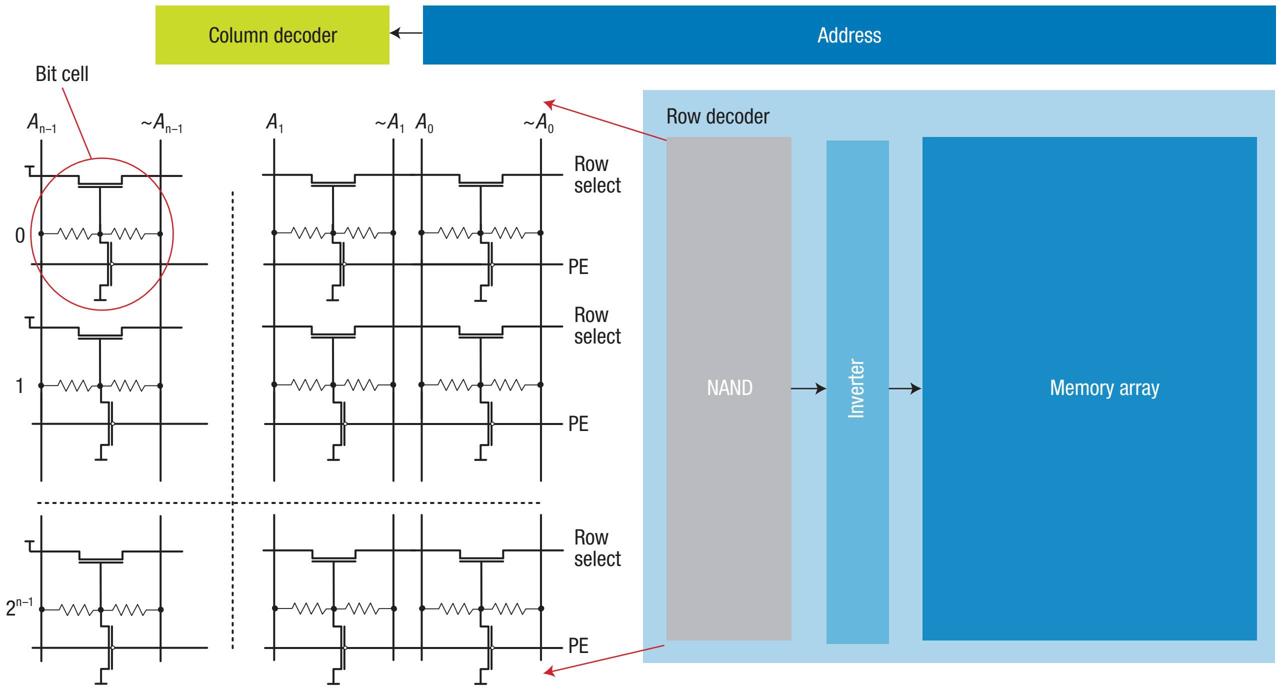


Figure 1. Two resistive elements are added to each NMOS transistor in a voltage-dividing manner.

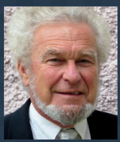
The authors' experimental results are quite promising. They show that the programmable resistive address decoder's energy and timing overheads are nonexistent during memory read and negligible during memory write. The area overhead is also minimal.

There are many possible applications of low-power CAM. Within contemporary processor cores, various storage structures are currently less than fully associative, including caches and translation look-aside buffers. If their hardwired address decoders are replaced with the programmable resistive address decoders, these structures become fully associative at the timing and energy cost of direct mapped ones.

Another intriguing application of low-power CAM is a virtually addressable main memory, which is enabled by replacing the hardwired NAND address decoder of main memory with the programmable resistive address decoder. Virtually addressable memory might significantly increase the CPU performance when running virtual machines and hypervisors (where nested memory mapping is required), as well as improve memory performance and utilization (by reducing the physical memory fragmentation), and enhance the system security. C

DANIEL J. SORIN is the Addy Professor in the Department of Electrical and Computer Engineering at Duke University. Contact him at sorin@ee.duke.edu.

50 & 25 YEARS AGO



EDITOR ERICH NEUHOLD
University of Vienna
erich.neuhold@univie.ac.at

MARCH 1968

Attainment of Reliable Digital Systems through Redundancy (p. 2) "The primary thesis of this paper is that the complex tasks of modern ultrareliable computing systems can only be accomplished through the judicious use of redundancy. First, a point of view is presented that partitions the field into the various technical topics of interest. Then, reference is made to specific, selected articles that have been published, or are otherwise available to the researcher, and brief comments are made that relate their contribution to the technology. Finally, the referenced documents themselves are listed in the form of an extensive bibliography that provides a source list for anyone wishing to delve further into the details of the particular approaches to the attainment of reliable digital structures through redundancy."

Hybrid Computers (p. 18) "Both analog and digital computers have become indispensable in many scientific and engineering disciplines. Each offers some unique advantages that cannot readily be achieved in the other. Digital computers are better for generating multi-variable functions. Analog computers are better (in most cases) for work that requires direct interface with external analog for work equipment. … Fortunately, the two machines can be combined to form a hybrid system which takes advantage of the best features of both. Simulation, which implies that there is a math model of the system, device, or phenomenon and that parameters of this model relate to certain important real variables of that system, is probably the major application for hybrid systems."

Standardization of Nondiagrammatic Presentation of Digital Logic (p. 25) "To achieve standardization we need to know how equations or their equivalent are being used to represent digital logic. You can help us by answering this questionnaire and returning it to us at the address listed at the end of this form."

MARCH 1993

www.computer.org/csdl/mags/co/1993/03/index.html

Processor Reconfiguration through Instruction-Set Metamorphosis (p. 11) "Computationally intensive applications typically spend most of their execution time within a small portion of the executable code. A general-purpose machine can substantially improve its performance in many of these applications by adapting the processor's configuration and fundamental operations to these frequently accessed portions of code. … To keep the processing platforms general purpose, RAM-based field-programmable gate arrays (FPGAs) provided the mechanism for implementing the augmented instructions. … [In the] PRISM approach, a configuration compiler is specialized to accept a program as input and produce both a hardware and a software image as output."

Adapting, Correcting, and Perfecting Software Estimates (p. 20) "Continuous software estimation models are needed that can constantly evaluate costs and schedules. This article proposes a hybrid estimation model and demonstrates how to use it. … According to Boehm, 'Good software cost estimation is not an end in itself but rather a means toward more effective software lifecycle management'. Indeed, without a reliable software estimation capability, effective project planning and control is next to impossible." *[Editor's note: The article discussed various models for cost and time estimation of software projects. Unfortunately, the frequent cost and time overruns in today's software projects demonstrate that such models are still lacking.]*

Concurrent Runtime Monitoring of Formally Specified Programs (p. 32) "Anna (Annotated Ada) is an Ada language extension to include facilities for formally specifying the intended behavior of Ada programs. … Anna is based on first-order logic, and its syntax is a straightforward extension of the Ada syntax. Anna constructs appear as formal comments within the Ada source text (within the Ada comment framework).” *[Editor’s note: Ada didn’t become as successful as envisioned and the formal extension technique hasn’t transferred to other languages. In my mind, this would help in today’s climate of power-programming, in which we seem to rely more on customer-based debugging than on formal verification.]*

Constrained-Latency Storage Access (p. 44) “Multimedia applications that require real-time access to large amounts of storage are forcing designers to take new approaches to the structure and function of storage systems. ... We use the term constrained-latency storage access (CLSA) for applications that have strict deadlines for the completion of some secondary storage accesses.” *[Editor’s note: Many other techniques have been developed over time, some of them realized in streaming data applications.]*

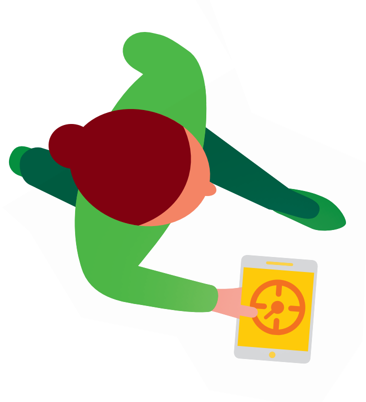
Object-Oriented Intelligent Computer-Integrated Design, Process Planning, and Inspection (p. 54) “An integrated manufacturing environment includes a computer-aided design [CAD] system, a computer-aided manufacturing [CAM] system, and a vision system. The CAD system is used to design the part, the CAM system to automatically generate a process plan and detailed instructions for machining the part, and the vision system to inspect the finished parts and monitor the execution of operations. An automated, flexible, and intelligent computer-integrated manufacturing (CIM) system might not be as inconceivable as once believed.” *[Editor’s note: This research used mostly constraint-based programming techniques and thus is only somewhat “intelligent”; for example, no machine learning approaches were applied. Large applications for intelligent CAD/CAM systems haven’t yet been developed, but the challenges resulting from 3D printing might provide new impetus.]*

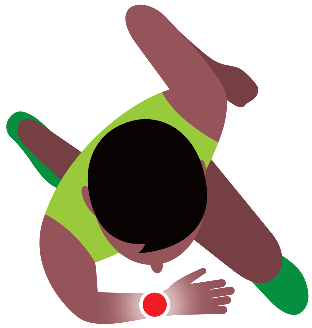
Layoffs in the Computer Industry (p. 66) “The current slump in the computer industry has long-term implications that will affect future employment. Many jobs lost during the downturn are likely to have gone forever. ... The computer manufacturing industry is in turmoil that goes deeper than the general economic recession, and this turmoil is reflected in severe job cuts.” *[Editor’s note: Since 1993, we’ve gone through a number of “restructurings,” for example, the 2001 Dot-com bubble. In other instances, companies once considered safe laid off workers by the thousands, most recently in 2016 and 2017.]* C

E-Coaching for Health

Oresti Banos, University of Twente
Christopher Nugent, Ulster University

 0018-9162/18/$33.00 © 2018 IEEE





E-coaching is an emerging computing area in which intelligent systems are used to encourage progress toward specific health-related goals by providing tailored training and guidance. Progress in this field could well be a key enabler of increasing health span and well-being for our increasingly care-demanding society.

Over the past few decades, our society has experienced radical changes incomparable with any other time in human history. This information technology revolution continues to transform the way in which we perceive and interact with the world, including how we work, entertain, and relate to one another.

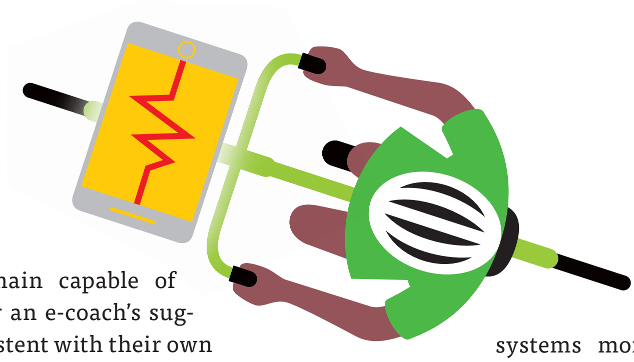
Although there have been many positive developments, major negative consequences of the widespread use of computing technology are that people are less physically active, isolate themselves more, tend to have increased levels of stress, and their diet is less nutritious. These lifestyle changes are already viewed as a global epidemiological threat, compromising healthy longevity not only for citizens of developed countries but also those living in developing nations, where obesity and diabetes are becoming the new malaria and HIV. It is therefore imperative for us to find effective new means for raising public awareness and changing behavior in order to improve our ability to remain active, healthy, and resilient. One area that offers this potential is e-coaching. E-coaching, which is the use of computing technology to help shift an individual's behavior to improve well-being, promises to play a leading role in helping us realize a healthier and more mindful society.

Changing behavior is far from straightforward. Over the years, a multitude of theoretical frameworks have been explored to help explain the reasons why humans break or follow established patterns of action. Examples of these are the Social Cognitive Theory, Goal-Setting Theory, or the Transtheoretical Model, all of which offer a degree of potential in influencing the individual to adopt new habits, set attainable goals, or positively manage expectations. Nevertheless, transforming these theories into practice is difficult, especially in an era in which human attention span and ability to be self-reflective are diminished. Few of us are self-aware enough to recognize the subtle ways in which we can positively and negatively affect our situation, nor do we possess enough self-determination to take action. This in turn means that a vast majority of us need help.

Really, we must accept that changing behavior to improve our health is a lifelong and resource-demanding activity. Most of us need to be frequently, if not continuously, observed, listened to, questioned, understood, reasoned with, taught, and/or advised—that is, coached—in several aspects of our lifestyle. Although it is fair to say that we encounter (human) coaches every day, including physical therapists, friends, clinicians, and so on, they are often only present at infrequent and often random moments. Most individuals would benefit from opportunistic coaching—to occur at the time and place when it is needed—instead of when possible.

Recent advancements in ubiquitous computing, artificial intelligence, and human–computer interaction have made it possible to develop the first generation of autonomous coaching systems, which are expected to deal with the inherently complex nature of coaching humans. As intelligent systems, e-coaches exhibit abilities similar to those of their human counterparts: sensing (observing, listening, questioning), learning (acquiring knowledge, understanding), and actuating (reasoning, clarifying, advising). An e-coach can observe, reason about, learn from, and predict a user's behavior to provide tailored and effective guidance to reach an individual's goal.

Sensing options are currently unparalleled due to the number of smart, ubiquitous sensing systems developed and deployed on a global scale. Instrumented devices such as smartphones or wearables—ingrained throughout our daily activities—enable unobtrusive observation and detection of a wide variety of behaviors as we go about our physical and virtual interactions with the world.



These observational capabilities are largely augmented by the widespread adoption of the Internet of Things; these new sensing ecosystems mean nothing we do goes unnoticed.

The vast amount of data generated by such sensing infrastructures can be digested by powerful machine-learning and data-mining algorithms, which map the raw data (for example, our GPS coordinates at specific times) into predictive trajectories of behavior (for example, when you walk out of your apartment in the morning, your smartphone displays the next departing train to your usual destination and approximate time to walk to the station). The processed data is combined with computerized behavior-change frameworks and domain knowledge to dynamically generate tailored recommendations and guidelines through advanced reasoning. Acting upon the user's detected situation and goals requires a system that contains the proper content and applications that can build collaborative relationships and fluent communication channels between e-coaches and the user. This can, for example, be realized through soft embodiments, such as virtual or augmented reality, or hard embodiments, such as robots or other interfaces.

E-coaching comes with some important challenges and dilemmas still to be addressed. Privacy, respect for autonomy, sufficient expertise, and policies and laws that support both integrity and responsibility of such systems are among the most relevant. For example, e-coaches should not eradicate a user's ability to make decisions. They should encourage changes to user behavior in accordance with that individual's own stated goals or vision for his or her lifestyle. Likewise, users should remain capable of assessing whether an e-coach's suggestions are consistent with their own values and priorities in life. A myriad of sometimes sensitive data is normally necessary for e-coaches to fully understand a user's situation and context to provide tailored recommendations, thus proper data management, anonymization, and encryption are essential to stave off any potential personal threat.

E-coaching is still in its infancy. The enormous technical, ethical, and societal challenges will require significant effort in order for these tools to be deployed successfully and to their full potential. Nevertheless, given our health and clinical care challenges, now is the time to push for these effective new technologies to both improve our lives and extend our health span. Shall we give it a try?

IN THIS ISSUE

We received a wide range of impressive articles to our open call for papers. So many that we will be publishing a second issue on this topic later in 2018.

In this issue, we present four articles addressing relevant theoretical and practical aspects of e-coaching systems for health and well-being applications.

In "Architecting E-Coaching Systems: A First Step for Dealing with Their Intrinsic Design Complexity," Sergio F. Ochoa and Francisco J. Gutierrez propose a new theoretical framework to assist in the design of an e-coaching system. The authors decouple the structural elements of the system from the coaching strategies to facilitate the scalability and adaptability to new domains or unforeseen aspects in the design phase. This framework could also help to prototype e-coaching systems more rapidly and also guide the design of the digital coaches as a collaborative activity in which software engineers, domain specialists and end users participate in the co-creation of the system.

In "Emotion and Motivation in Cognitive Assistive Technologies for Dementia," Julie M. Robillard and Jesse Hoey discuss seminal advancements in the integration of emotions in computing, including their use in building e-coaching technologies that are more effective and acceptable to the user. The authors analyze the limited success that assistive technologies have had in practice so far and link this to a lack of social and emotional alignment. Through the Bayesian Affective Control Theory, the authors introduce a quantitative socio-psychological framework that integrates salient sources of information such as prior expectations and observed emotions and actions to model emotional behavior in computerized assistive systems. This framework helps provide the grounds for moving beyond current emotion-appraisal–based e-coaching systems and toward systems more aligned with the values and needs of users.

In "Recommender System Lets Coaches Identify and Help Athletes Who Begin Losing Motivation," Paolo Pilloni, Luca Piras, Salvatore Carta, Gianni Fenu, Fabrizio Mulas, and Ludovico Boratto describe a method for automatically anticipating if and when an athlete will stop training due to a decline in motivation. The authors present a dropout-predictive model developed on the historical workout data collected from users' smartphone sensors and self-reports. Once a risk of

dropout is detected, the system notifies the personal coach with a report on the exercise statistics and possible recommendations. Coaches can then communicate with these individuals through a web-based interface to encourage these individuals to keep training through more personalized routines and effective strategies to discourage training abandonment.

In "Monitoring Eating Behaviors for a Nutritionist E-Assistant Using Crowdsourcing," Mario O. Parra, Jesus Favela, Luis A. Castro, and Arturo Morales present an approach designed to automatically infer and counsel people regarding eating behaviors, a fairly uncharted and challenging area in the e-coaching domain. The authors combine semi-autonomous crowdsourced food intake assessment with a virtual conversational assistant to monitor and coach individuals undergoing treatment for weight reduction. Users are asked to regularly photograph their meals, which are then evaluated for nutritional content, calories, and wholesomeness by a group of Internet users. This information is used to inform the digital assistant about the user's diet so that he or she can then respond with tailored recommendations and motivational tactics. The authors report on the latency, cognitive load, and accuracy of various crowdsourced food-intake assessment methods and also provide some preliminary findings on users' perceptions of the nutritionist e-assistant.

These four articles provide an exceptional overview on the innovative work taking place in the e-coaching area. Being chiefly multidisciplinary by definition, e-coaching will continue to revolutionize the way we interact and communicate with the world, augmenting our self-awareness and ultimately yielding an unparalleled ecosystem to support our future health and wellness. C

ABOUT THE AUTHORS

ORESTI BANOS is an assistant professor of creative technology and computational behavior modeling at University of Twente. His research interests include human-aware computing, behavior and context modeling, intelligent coaching systems, and smart pervasive sensing. Banos received a PhD in computer science from the University of Granada. Contact him at o.banoslegran@utwente.nl.

CHRISTOPHER NUGENT is a full professor of biomedical engineering at Ulster University. His research addresses the themes of the development and evaluation of technologies to support ambient assisted living, including mobile-based reminding solutions, activity recognition, and behavior modeling. Nugent received a PhD in biomedical engineering from Ulster University. He is *Computer*'s Digital Health area editor. Contact him at cd.nugent@ulster.ac.uk.

Architecting E-Coaching Systems: A First Step for Dealing with Their Intrinsic Design Complexity

Sergio F. Ochoa and Francisco J. Gutierrez, University of Chile

E-coaching systems are a highly valuable asset for promoting healthier lifestyles. However, the design of these systems is intrinsically complex. This article proposes a loosely coupled architecture to support the modeling of e-coaching solutions, reducing design complexity and enhancing the flexibility required for the e-coaching process.

E-coaching is a plan conducted by a system to lead a person from an initial stage to a target goal—for example, to make sure that a sedentary person walks for 30 minutes every day. The plan typically includes a set of milestones (interim goals) and several potential paths (coaching actions) that can be used to reach milestones. However, the suitability of each path depends on a particular coaching context (for example, personal characteristics and behavioral attitudes), the coaching goal to be reached, and the user's current progress. Therefore, the paths in the coaching process must be defined dynamically (considering the set of available alternatives) and be based on the current context and past experiences of the user.

The widespread use of smartphones and commercial wearable devices—such as Fitbit, Nike+ FuelBand, and Jawbone UP—has created several opportunities to develop e-coaching systems for health prevention.[1] Bart Kamphorst defines an e-coaching system as "a set of computerized components that constitutes an artificial entity that can observe, reason about, learn from, and predict a user's behaviors, in context and over time, and that engages proactively in an ongoing collaborative conversation with the user in order to aid planning and promote effective goal striving through the use of persuasive techniques."[2] Although these systems are typically autonomous, they can also be used to support blended coaching strategies in scenarios where humans

0018-9162/18/$33.00 © 2018 IEEE

(for example, health specialists) play an active role. Typically, the complexity of these systems increases with their required level of autonomy.

Designing e-coaching systems, particularly those addressing healthcare scenarios, is intrinsically complex. It requires knowledge from several domains (psychology, health, persuasive computing, human–computer interaction, data science, and software engineering[3,4]), and the design must be context-aware. In this sense, the system design must consider not only the sensing and representation of the e-coaching context, but also contextualized reasoning and learning based on the impact of past coaching actions.[2,5] Moreover, making these systems effective requires dealing with the complexity of identifying suitable persuasion mechanisms, using artificial intelligence techniques, and implementing system personalization (using self-adaptive mechanisms based on user models and context).[5]

In addition, there are almost no structural design guidelines (such as reference architectures or patterns) to help deal with such complexity. Therefore, software designers and researchers must conceive and model these systems using their intuition instead of reusing designs that have proven to be useful in the past.

To reduce the complexity of designing these systems, we present a loosely coupled architecture for e-coaching systems (LAES), which addresses the main design concerns of these systems and decouples the system structure from its behavior. This strategy allows for addressing the design of these systems in an incremental way to obtain a flexible solution that can self-adapt its behavior during the coaching process. The proposed architecture also provides a guide for conceiving e-coaching systems and exploring the particular design concerns (persuasion, self-adaptation, personalization, reasoning, and diagnosing). This represents a first step toward a personal informatics theory that organizes, understands, and contextualizes the knowledge in this study domain, helping to address the study and modeling of these systems in a more affordable way.

E-COACHING SYSTEM DESIGN

Kamphorst[2] proposes a set of minimal features that e-coaching systems must include: social ability, credibility, context-awareness, personalization (user tailoring), learning of user behavior (including behavior change), proactiveness, and guidance (coaching planning). These requirements can be used as a checklist to determine the potential coaching capabilities under development or in already implemented systems or designs. However, the relations among these requirements are not explicit, so it is not clear how to use them for constructing a system model.

It is important to note that the design of these systems involves two major concerns: the system structure and its e-coaching mechanisms. The system structure determines the capabilities potentially required by the system to conduct a coaching process, such as context awareness, decision making, and self-adaptation. This structural design must be independent of any coaching task and represents only the underlying infrastructure of the intelligent system. Typically, this design is quite conceptual and generic; therefore, it can be reused to conceive or adjust other e-coaching solutions.

The second design concern (e-coaching mechanisms) is focused on determining how to properly use the available resources to conduct the e-coaching process in an effective way. This design stage includes the definition of the particular e-coaching process and work context to be considered in the application domain. This also includes characterizing the target person (such as gender, age, attitude, coaching goal, progress, and past experiences). Suitable results of this second design stage are more difficult to reuse because they are context-dependent. Therefore, they are recommended to be used only in the same or similar application contexts. Both design concerns are mandatory and equally challenging for designers. However, addressing the latter makes sense only if the former has been properly identified. Software engineering has shown that reusing this type of solution (design and architectural patterns for a particular application domain) reduces the design complexity of systems.

Although related literature reports that many e-coaching systems support healthcare scenarios,[1,3,6] in most cases the reusability of these designs tends to be low or null for several reasons. For instance, many proposals present their system designs as the fusion of structural and specific coaching components.[7–9] Given their specialization, they are potentially reusable only when designing systems that follow the same coaching purposes and are therefore used in the same application context. In many other cases, the designs are presented as articulated collections of software and hardware components, representing the system infrastructure. These designs are usually addressed in certain types

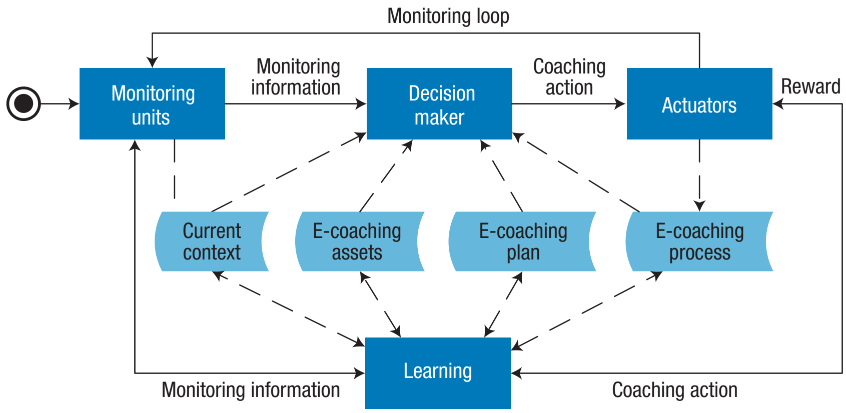


FIGURE 1. Conceptual model of an e-coaching system.

of systems, such as socially assistive robotics[10] and Internet of Things–based health-supporting systems.[11] The reusability of these designs is also limited, not only because specialized and general structures should be reused together, but also because of the high complexity of these supporting technologies, which makes these proposals obsolete in the short term.

The literature also reports some conceptual models that can be used to support the structural design of e-coaching systems and thus help deal with their inherent design complexity. Some of the most well-known models are Computerized Behavior Intervention,[12] PHG (personalized health guide),[13] the transtheoretical model of behavior,[14] and several flavors of cognitive architectures.[15] These models are highly useful and reusable, but they represent the system structure in terms of psychological and cognitive concepts. They do not propose a conceptual representation from a software engineering point of view. Therefore, developers—although being able to reuse these models—still have to bridge the gap between what the models propose and what needs to be designed for the operative structure of the e-coaching system.

A CONCEPTUAL MODEL

The structural design identifies the main components of the system, as well as the way in which the components interact to provide contextualized e-coaching services (see Figure 1). This structure organizes the system in terms of functional components and supporting data, and it recognizes the four major activities of the e-coaching processes: monitoring, decision making, persuasion/awareness provision, and diagnosing/learning.[2,4]

The first component monitors particular context variables and events, and accordingly informs both the e-coaching decision maker and the learning unit.[3] It also includes the notification of relevant changes in the e-coaching context. The decision maker uses this input to determine whether an e-coaching action is required. If it does, that component determines when and how to intervene in the user space, as well as what action is required to be triggered to persuade the user to do something or to make the user aware of a situation.[4] The actuator conducts the coaching actions, which can be persuasive or reflective.[3] Persuasive actions motivate behavior change using explicit actions, whereas reflective actions provide awareness information that can be accessed on demand by the user.[16]

The impact of each coaching action should be evaluated to determine not only its usefulness in the future, but also to assess the need to trigger a new action if the former action was not effective. A learning unit is in charge of this activity, which also delivers intrinsic or extrinsic rewards to keep the user engaged with the e-coaching process.[16] As shown in Figure 1, the first three components run the monitoring loop, and the last component improves the knowledge available for performing more effective coaching actions in the future.

The e-coaching process should also count on supporting information to properly perform the activity. As depicted in Figure 1, this information includes the current context, the e-coaching assets, the e-coaching plan, and the e-coaching process. The current context characterizes the ongoing coaching situation, which includes the user model, the status of the context variables being monitored, and the sequence of events sensed by the monitoring units.[3,5] The e-coaching assets include the contextualized and prioritized rules and actions that can be used to instantiate and adjust the coaching plan. The plan establishes the interim and final goals, but the details about how to reach these goals should be built on demand by the learning unit according to previous experiences and using the available e-coaching assets. It is important to note that the findings

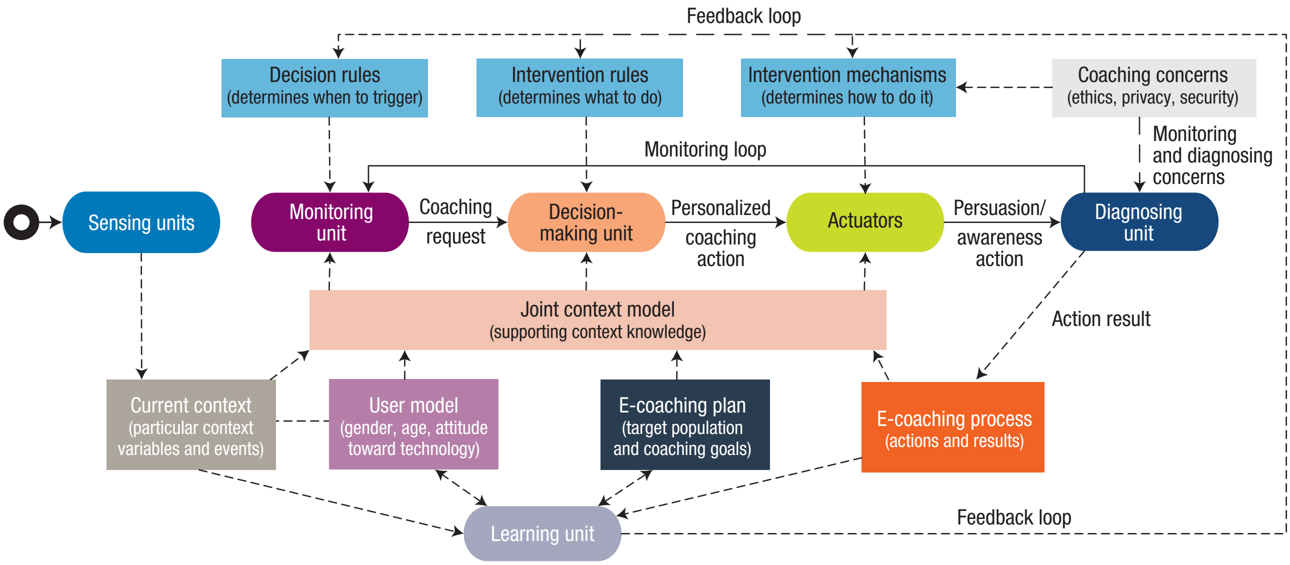


FIGURE 2. Loosely coupled architecture of an e-coaching system.

reported in e-coaching studies can be represented as contextualized assets and thus become reusable for developers and researchers.

Finally, the e-coaching process establishes the sequence of actions and rewards delivered to the user, trying to move the user's behavior from the current status to a target status. Given that these actions temporarily affect the user behavior, the feedback obtained from each action is recorded in the e-coaching process and then used to adjust the coaching plan, the context status, and the priority of the available assets.

In this conceptual model, the behavior of the e-coaching system is mainly specified through data representations. Therefore, it can be tailored, personalized, and extended on the fly (for example, through unattended remote updates) without changing the system source code. This design strategy makes these systems more flexible, adaptable, and better prepared to address medium-to-long coaching processes. According to Thomas Fritz and his colleagues,[3] most systems reported in the literature only evaluate their effectiveness in the short term, probably because their behavior adaptation requires changing source code. In this sense, LAES makes it easier for the system to self-adapt and can even improve the system behavior according to the current e-coaching context and process evolution.

LAES: LOOSELY-COUPLED ARCHITECTURE FOR E-COACHING SYSTEMS

Figure 2 shows the LAES architecture, which was derived from the model shown in Figure 1 but includes many more design details. For instance, the component initially identified as "e-coaching assets" was divided into four pieces of knowledge: the decision rules that determine when a coaching action is potentially required; the intervention rules that establish what action to take according to the coaching context; the intervention mechanisms that indicate how the coaching action must be conducted; and the coaching concerns that establish the ethics and privacy and security issues to be considered during the e-coaching process. This definition of e-coaching assets is aligned with the steps identified by Heleen Rutjes and her colleagues[4] to conduct an e-coaching action. These pieces of knowledge are contextualized because the suitability of their components is typically expressed through predicates that involve context variables and values assumed by them. Therefore, the findings of the studies reported in the literature can be specified as rules, intervention mechanisms (to persuade or provide awareness), and constraints to the e-coaching actions.

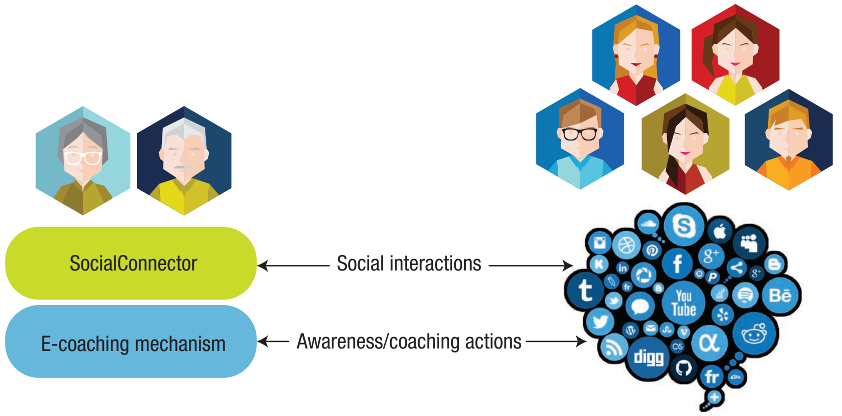


FIGURE 3. E-coaching scenario of the SocialConnector system.

The previous monitoring units component is now decoupled into sensing and monitoring units. Sensing units are in charge of keeping the value of the context variables updated, and monitoring units identify context changes and determine whether they eventually require a coaching action. This structure adheres to a regular self-monitoring process, like the one described by Michael Klein and his colleagues.[12] Therefore, it represents an input for the decision-making process.

The previous decision maker and actuators were not specified in detail because their design will depend on the particular coaching activity they are supporting. Therefore, their design will not be reusable in another context. The previous learning unit was decoupled into two loosely coupled components: the diagnosing unit and the learning unit. The diagnosing unit infers the potential impact of any coaching action and records such information in the coaching process component to support learning from previous experiences. The learning unit consumes information from the joint context and produces knowledge that refines, extends, or updates the coaching assets. Many researchers propose the use of machine-learning techniques to generate this new knowledge, which is quite favorable in systems adhering to the LAES architecture, given that they contextualize and centralize the coaching information of the users.[4] The learning unit also uses such knowledge to adjust the user model and instantiate the e-coaching plan.

Figure 2 also shows the use of a joint context model, which is the aggregation of the current context, user model, and the e-coaching plan and process. Although the e-coaching literature recognizes the relevance of considering the coaching context and user modeling, there are almost no reports indicating the structure of these components. For instance, Klein and his colleagues suggest several context variables that can be used to support e-coaching for therapy adherence.[12] Although useful, such a proposal is specific for that application domain and is therefore not easy to reuse in other coaching purposes. In that respect, the structure of context presented in this model also represents a contribution to the state of the art, recognizing that each particular e-coaching system must derive specific context models from this structure during the design process.

Deriving an e-coaching system

We used the LAES architecture to model an e-coaching mechanism that is embedded into the SocialConnector application[17] to promote the social health of older adults. This e-coaching system encourages adult children and grandchildren to interact with older family members using social media (see Figure 3). This helps the elderly improve their emotional status, and helps the rest of the family feel that they are complying with their implicit filial obligations.

The main design goal of this system was to encourage people to respect the social interaction rhythm agreed upon between the elderly and each family member.[19] Figure 4 shows the design of this e-coaching system, where the color of components and assets maps to the components defined in the LAES architecture (see Figure 2).

To derive an e-coaching system from the LAES architecture, the designer should determine what components of the architecture are required to support the new coaching process and how to instantiate them to reach the coaching goals. For instance, in this case, the user model component (see Figure 2) was decomposed into four components given the nature of the participants in this scenario: family, member, relationship, and user preference (see Figure 4). Similarly, the role of the sensing unit (see Figure 2) is played by the interaction sensor (see Figure 4), which senses and records both the social interactions involving

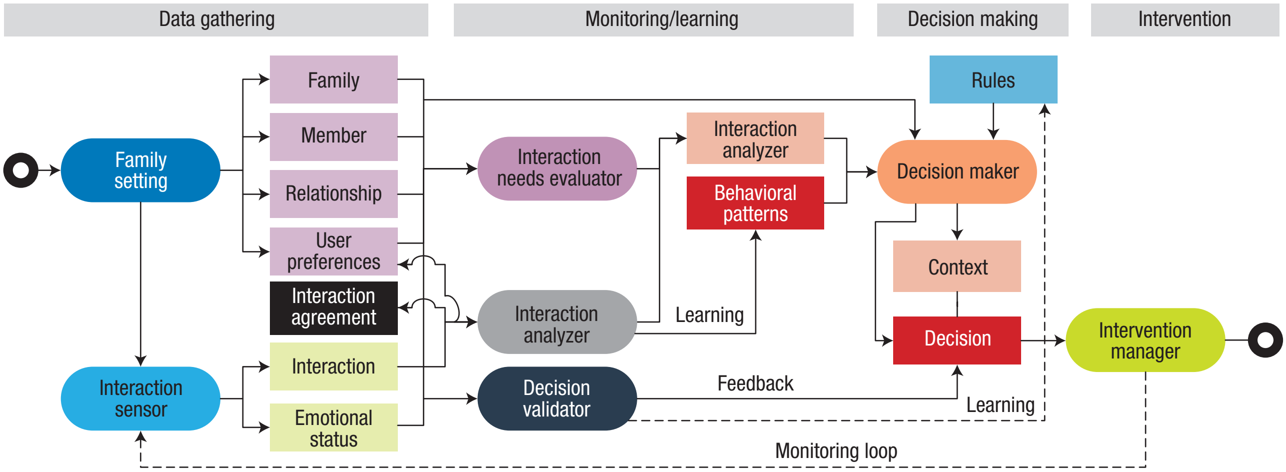


FIGURE 4. Structure of the SocialConnector e-coaching system.[19]

older adults and the emotional status of the elderly (in other words, it records the current context).

The most complex instantiation is always the definition of the rules and mechanisms to conduct the coaching process according to the current context (see Figure 2). This complexity is given by the uncertainty about the suitability of each potential alternative to be used in the process. In this sense, findings of coaching studies can be used to instantiate these assets. Similar to other models, the LAES architecture does not ensure the proper instantiation of these components. However, it represents such an instantiation in the form of data abstractions that allows human designers or coaches to add and adjust the e-coaching rules remotely, quickly, and with low effort. Given that these update processes are not visible to the end users, designers can tune the e-coaching strategy whenever they want, as many times as required. This capability is the result of the decoupling of components proposed by the LAES architecture.

This e-coaching system was used during nine weeks by nine family networks (64 people), involving two changes of the coaching strategy (at the end of weeks three and six). The obtained results show that the system was useful for reaching the coaching goal, and that its structure allowed users to tailor the coaching strategy on the fly and remotely, without requiring a physical access to the device used by the participants.[18] This shows in practice the self-adaptation capabilities that can be expected from systems derived from the LAES architecture.

The minimal requirements for e-coaching systems

Considering the list of minimal requirements stated by Kamphorst[2] for e-coaching systems, we conducted an inspection of LAES using an approach inspired by the Architecture Tradeoff Analysis Method (ATAM).[20] The inspection lasted for almost three hours and involved three external software architects with more than 10 years of experience designing collaborative systems. After reviewing the core of these requirements, we presented the LAES architecture and the reviewers evaluated the capabilities of the architecture to address the stated requirements, asking for design justifications when required. Following are the evaluation results.

- Social ability. The social ability of an e-coaching system allows it to engage with users in an ongoing interaction. The monitoring unit (see Figure 2) helps reach such a goal, as it regulates the interaction with the user considering three criteria: taking action when the user requires support, triggering coaching actions when the process requires it, and keeping the social link with users without overwhelming them.
- Credibility. The coaching actions triggered by the system must make sense to the users and be perceived as being trustworthy. Otherwise, they

will be ignored, jeopardizing the e-coaching process. In LAES, the use of contextualized assets and a joint context model both support e-coaching contextual decision making. Therefore, if the e-coaching actions were properly defined, it is expected that users will perceive the system as credible.

- Context-awareness. The context-awareness capability of the system is required not only for recommending tailored coaching actions (what to do), but also to determine when and how these actions should be performed. In this respect, all functional components of the architecture are context-aware and use contextual data representations to determine their behavior in real time, thus making the system behavior self-adaptable.
- Personalization (user tailoring). The user model component is utilized to personalize the e-coaching activities. However, its representation was not defined in the architecture given that the relevance of the contextual attributes (such as gender, age, or attitude toward technology) depends on each particular application domain.
- Learning of user behavior (including behavior change). It is expected that user behavior will change over time; therefore, the system must identify these changes and learn from them. The learning unit uses the context information and the coaching plan and process as input to identify the behavior changes. The generated knowledge (the learning) is represented through decision and intervention rules, and in the usage priority of the intervention mechanisms.
- Proactiveness. This requirement is addressed mainly by the monitoring unit, which is in charge of keeping a permanent social link with the user and taking action when required. This component uses the record of past experiences and the coaching plan to allow the system to react properly by itself, considering the context situation.
- Guidance (coaching planning). Given that the user behavior usually changes over time, it only makes sense to define detailed coaching actions in the plan for the short term. The coaching plan records the interim and final goals, and the learning unit periodically defines coaching actions required to reach the next goal, according to the plan and using the available assets.

The reviewers agreed that LAES addresses and supports the functional and quality requirements defined by Kamphorst.[2] They also acknowledged that the proposed functional decoupling allows e-coaching systems to self-adapt their behavior just by changing the supporting data. This represents a low-cost adaptation capability that is highly demanded by most e-coaching systems.

The proposed LAES architecture, which decouples the e-coaching system structure from its behavior, allows for the incremental design of e-coaching systems and strategies. It can also be used to guide the design of an e-coaching system as a collaborative activity in which software architects, coaching specialists, and end users participate, leading to more effective systems. Given that the structure of LAES-based systems is decoupled from its behavior, which is also based on data (mainly context-aware rules), it is expected that these systems have a high self-adaptation capability involving low effort. Moreover, such supporting information can be extended or tuned on demand using the results of past experiences or findings of studies reported in the literature, increasing the coaching capabilities of the system. This capability is not present in the models discussed in this article, so our work represents a contribution to the state of the art in this study domain. ■

ACKNOWLEDGMENTS

This work has been partially supported by the Fondecyt Project (Chile), Grant #1150252. The work of Francisco J. Gutierrez has been supported by the PhD scholarship program of Conicyt Chile (CONICYT-PCHA/Doctorado Nacional/2013-21130075).

REFERENCES

1. J. Rogers and A. Maini, *Coaching for Health: Why It Works and How to Do It*, Open University Press, 2016.
2. B.A. Kamphorst, "E-Coaching Systems: What They Are, and What They Aren't," *Personal and Ubiquitous Computing*, vol. 21, no. 4, 2017, pp. 625–632.
3. T. Fritz et al., "Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness," *Proc. SIGCHI Conf. Human Factors in Computing Systems* (CHI 14), 2014, pp. 487–496.
4. H. Rutjes, M.C. Willemsen, and W.A. IJsselsteijn, "Understanding

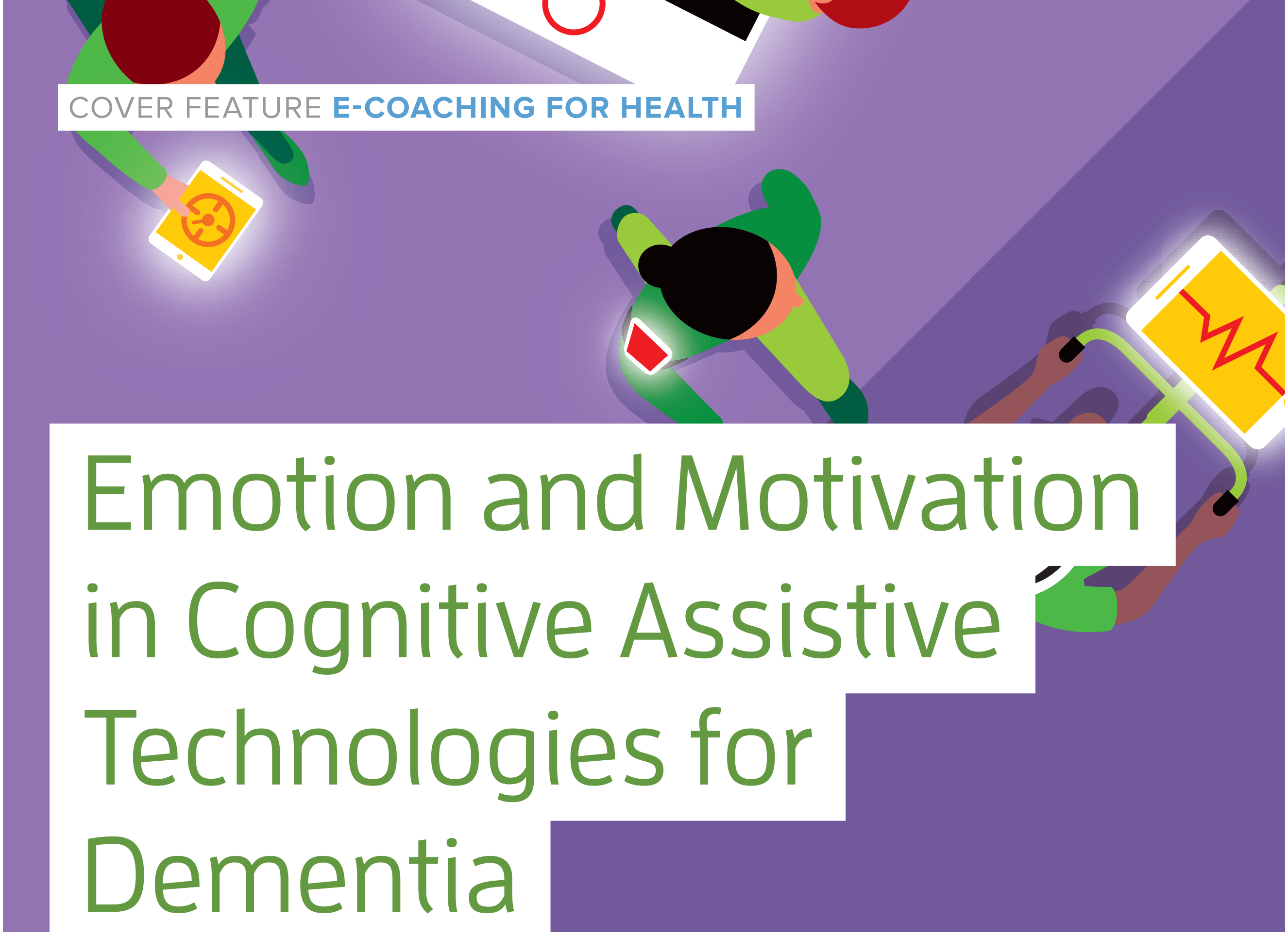
ABOUT THE AUTHORS

SERGIO F. OCHOA is an associate professor in the Department of Computer Science at the University of Chile. His research interests include computer-supported collaborative work, software engineering, and mobile and ubiquitous computing. Ochoa received a PhD in computer science from Pontificia Universidad Católica de Chile. He is a member of IEEE, ACM, the Chilean Computer Science Society, and the LACIR (Latin American and Caribbean Collaborative ITC Research Initiative) steering committee. Contact him at sochoa@dcc.uchile.cl.

FRANCISCO J. GUTIERREZ is an adjunct professor in the Department of Computer Science at the University of Chile. His research interests include the design and evaluation of social computing systems, persuasive technology for promoting wellbeing, and empirical studies in software engineering. Gutierrez received a PhD in computer science from the University of Chile. He is a member of ACM, SIGCHI, SIGCSE, and the Chilean Computer Science Society. Contact him at frgutier@dcc.uchile.cl.

Effective Coaching on Healthy Lifestyle by Combining Theory and Data-driven Approaches," *Proc. Personalization in Persuasive Technology Workshop*, 2016; ceur-ws.org/Vol-1582/13Rutjes.pdf.

5. A. Van Wissen, "Agent-Based Support for Behavior Change: Models and Applications in Health and Safety Domains," PhD dissertation, VU Univ. Amsterdam, 2014.
6. M. Kaptein, "Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking," *ACM Trans. Interactive Intelligent Systems*, vol. 2, no. 2, 2012, article no. 10.
7. R.J. Beun, "Talk and Tools: The Best of Both Worlds in Mobile User Interfaces for E-Coaching," *Personal and Ubiquitous Computing*, vol. 21, no. 4, 2017, pp. 661–674.
8. S. Obdržálek et al., "Architecture of an Automated Coaching System for Elderly Population," *Studies in Health Technology and Informatics*, vol. 184, 2013, pp. 309–311.
9. R. op den Akker, "A Personal Context-Aware Multi-Device Coaching Service that Supports a Healthy Lifestyle," *Proc. 25th BCS Conf. Human-Computer Interaction* (BCS-HCI 2011), 2011, pp. 443–448.
10. J. Fan et al., "A Robotic Coach Architecture for Elder Care (ROCARE) Based on Multi-User Engagement Models," *IEEE Trans. Neural Systems and Rehabilitation Eng.*, vol. 25, no. 8, 2016, pp. 1153–1163.
11. A. Amato and A. Coronato, "An IoT-Aware Architecture for Smart Healthcare Coaching Systems," *Proc. IEEE 31st Int'l Conf. Advanced Information Networking and Applications* (AINA 2017), 2017, pp. 1027–1034.
12. M. Klein, N. Mogles, and A. van Wissen, "An Intelligent Coaching System for Therapy Adherence," *IEEE Pervasive Computing*, vol. 12, no. 3, 2013, pp. 22–30.
13. A. Honka et al., "Rethinking Health: ICT-Enabled Services to Empower People to Manage Their Health," *IEEE Rev. Biomedical Eng.*, vol. 4, 2011, pp. 119–139.
14. J.O. Prochaska, "Stages of Change and Decisional Balance for 12 Problem Behaviors," *Health Psychology*, vol. 13, no. 1, 1994, pp. 39–46.
15. P. Langley, "Progress and Challenges in Research on Cognitive Architectures," *Proc. 31st AAAI Conf. Artificial Intelligence* (AAAI 17), 2017, pp. 4870–4876.
16. S. Munson and S. Consolvo, "Exploring Goal-Setting, Rewards, Self-Monitoring, and Sharing to Motivate Physical Activity," *Proc. 6th Int'l Conf. Pervasive Computing Technologies for Healthcare* (PervasiveHealth 12), 2012; doi: 10.4108/icst.pervasivehealth.2012.248691.
17. D. Muñoz et al., "A Social Cloud-Based Tool to Deal with Time and Media Mismatch of Intergenerational Family Communication," *J. Future Generation Computer Systems*, vol. 53, issue C, 2015, pp. 140–151.
18. F.J. Gutierrez, S.F. Ochoa, and J. Vassileva, "Mediating Intergenerational Family Communication with Computer-Supported Domestic Technology," *Proc. 23rd Int'l Conf. Collaboration and Technology* (CRIWG 17), 2017, pp. 132–147.
19. F.J. Gutierrez, S.F. Ochoa, and J. Vassileva, "Mediating Asymmetries in Family Communication: Supporting the eInclusion of Older Adults," *Proc. Int'l Conf. Universal Access in Human-Computer Interaction* (UAHCI 15), 2015, pp. 438–448.
20. R. Kazman, M. Klein, and P. Clemens, "ATAM: Method for Architecture of Evaluation," tech. report CMU/SEI-2000-TR-004, ESC-TR-2000-004, Carnegie Mellon Software Engineering Institute, 2000.



Emotion and Motivation in Cognitive Assistive Technologies for Dementia

Julie M. Robillard, University of British Columbia and BC Children's and Women's Hospital and Health Centres

Jesse Hoey, University of Waterloo

The adoption and effectiveness of cognitive assistive technologies hinge on harnessing the dynamics of human emotion. The authors discuss seminal advances in the integration of emotions in assistive technologies for dementia and propose Bayesian Affect Control Theory (BayesACT), a quantitative social-psychological theory, to model behavior and emotion in such systems.

Technology that persuades people to act must understand and leverage the shared structure and dynamics of human emotion. Intelligent assistive systems for dementia have had limited success in practice, and we hypothesize that integrating emotional reasoning will lead to greater effectiveness, acceptance, and usability. Toward this end, we propose Bayesian Affect Control Theory (BayesACT), a quantitative social-psychological theory about how people perceive others and act socially, to model behavior and emotion in assistive systems and, in turn, promote the alignment of these technologies with users' values and needs. Integrating BayesACT in assistive technologies offers many potential benefits but also presents several challenges.

ASSISTIVE TECHNOLOGY FOR DEMENTIA

Dementia is characterized by the progressive deterioration of cognitive and functional capabilities, leading to a loss of the capacity to perform activities of daily living

 0018-9162/18/$33.00

such as bathing and medication taking. Intelligent assistive technologies have been proposed as a possible solution to support people with dementia in performing these activities independently.[1]

COACH (Cognitive Orthosis for Assisting aCtivities in the Home) is an example of one such technology.[2] It uses a probabilistic and decision-theoretic model of a given daily activity such as hand washing, tooth brushing, or cooking. The model maps from various sensor inputs, including cameras, to a set of prerecorded audio and video cues that are played when the person stops making progress in the task. When tested in a long-term care facility, COACH was found to reduce the need for human assistance, in some cases by as much as 100 percent.[2] However, for some older adults, COACH failed to provide appropriate assistance, leading to confusion or agitation. This result might be due to an emotional misalignment of COACH with the specific needs or personality of the individual user.[3]

Although significant effort has been made to design prompts based on the methods and styles of human caregivers, such as task-focused strategies or paraphrased repetition,[4] a simple one-size-fits-all style of prompting could be limiting. For example, some individuals might respond to a more servile approach, while others might prefer a more imperative style. These responses can be predicted to some extent by models for technology adoption[5] and are influenced by a range of factors such as personal background, sense of self and identity, and emotional responses to prompts, whether given by human or machine.

Recent interviews with older adults with dementia and their caregivers show that emotional processing remains considerably more intact than cognitive processing in dementia, particularly in social situations.[3] These results support the proposition that, while beliefs grounded in cognitive memories fade as people lose their ability to remember people and events, affective aspects could persist longer, even without situational context. Therefore, explicit models of affective meanings offer an attractive mechanism for developing more personalized assistive technologies for dementia.

While Damasio's work is not without criticism, emotional reasoning has emerged as a necessary guide for cognitive deliberation. To some extent, these considerations have been left out of the quest to build artificial intelligence (AI), which to date has largely focused on mapping the perception of stimuli to action by invoking rational utility. Rational AI agents must have preferences that obey certain axioms, and must be able to assess probabilities of outcomes. However, rational decision making engenders numerous paradoxes and inconsistencies,[7] which researchers often attribute to the limited capacity of human intelligence to hold axiomatic preferences or estimate correct probabilities. From this perspective, paradoxes and inconsistencies exist only because of a shortcoming of human intelligence.

Cognitive scientists have attempted to build theories of intelligence that account for deviations from rationality—for example, by following Herbert A. Simon's view of emotions as interrupting mechanisms to cognitive processing so that the system can attend to urgent needs in real time,[8] by modeling the impact of emotion on decision making through the lens of behavioral economics,[9,10] or by focusing on the function of emotion as "bridging the gaps of rationality."[11]

EFFECTIVELY INTEGRATING AFFECTIVE RESPONSES INTO ASSISTIVE TECHNOLOGY REQUIRES A CLEAR, OPERATIONAL DEFINITION OF EMOTION.

INTEGRATING EMOTIONS IN COMPUTING

Effectively integrating affective responses into assistive technology requires a clear, operational definition of *emotion*. Emotions are increasingly recognized as critical components of decision making and human social action.[6] Antonio Damasio's hypothesis that emotions guide behavior contrasts with the Platonic "high-reason" view of intelligence, in which rationality is primarily used to make decisions. Damasio argues that learned neural markers focus attention on likely successful actions, and act as a neural bias allowing people to work with fewer alternatives. These somatic markers are interpreted internally as "cultural prescriptions" for behaviors that are rational relative to social conventions.

Rationality is at the forefront of most of these theories: goals and plans are analyzed and interpreted, and emotions are subsequently generated, preparing the agent for action.

In many of these theories, emotions are viewed as discrete labels describing categorical primitive feelings (for example, "happy" and "sad"), or as a single dimension of valence (good vs. bad). However, this view is at odds with the hundreds of different subtle emotional states humans can describe. Theorists usually propose some mixing mechanism to account for this variety, but it is unclear how this process occurs. This view is challenged by Lisa Feldman Barrett, who argues that there is little empirical evidence to support the idea of discrete emotions like "anger" and "sadness" as fundamental entities in humans.[12] Although some emotion theories do not take the emotions categories to be natural kinds, most still use the primitive emotion categories as endpoints of the analysis. James A. Russell and Barrett propose that a 2D "core affect" might be the fundamental affective entity that is first elicited by stimuli and then categorized using appraisal processes to arrive at interpretations of emotions.[12,13] This core affect forms the basis of "dimensional" theories of emotion. Dimensional emotion has a long history dating back to the 1800s, but modern approaches usually take the view that at least three dimensions are needed (valence, arousal, and potency), possibly four if uncertainty is included.[14] Although popularized in the affective computing literature, the 3D structure of emotions was originally explored by Charles E. Osgood in the 1960s.[15]

The field of affective computing[16] has yet to fully embrace the deep-seated connection between emotion and action, perhaps because, until recently, there has not been a precise, computationally implementable definition of how emotions guide action.

ACCOUNTS OF EMOTION IN AFFECTIVE COMPUTING OFTEN CONSIDER THE APPRAISAL PROCESS—WHAT "MAKES" AN EMOTION.

Accounts of emotion in affective computing often consider the appraisal process—what "makes" an emotion. For example, the classic decision tree in the Ortony, Clore, and Collins (OCC) appraisal model has cognitive appraisal decision nodes and emotions as leaves.[17] Seminal work by Clark Elliott used an OCC model augmented with "love," "hate," and "jealousy" to make predictions about human emotional ratings of semantically ambiguous storylines.[18] OCC models were also integrated with Bayesian networks and probabilistic models for tutoring applications, with a focus on understanding student emotions, but leaving intervention to future work.[19] Klaus Scherer breaks emotion down into five components: appraisal, activation, expression, motivation, and feeling.[20] A behavioral component handles relevance to the organism and prepares possible reactions to stimuli. Emotion is proposed as a facilitator of learning and as a mechanism to signal and predict forthcoming action, but this relationship is not fully elucidated. Other approaches have attempted to reverse-engineer emotion through reinforcement learning (RL) models that interpret the antecedents of emotion as aspects of the learning and decision-making process, but relegate the function of emotion to characteristics of the RL problem.[21] Many of these approaches borrow from behavioral economics and cognitive science to characterize the consequences of emotion in decision making and integrate this knowledge as "coping rules" or "affect heuristics" to influence AI agents' behavior.

In affective computing research, coping is usually modeled as a separate mechanism that maps emotions to a set of action filters that can guide or change decisions made cognitively. Jonathan Gratch and Stacy Marsella[22] proposed a five-stage coping process wherein beliefs, desires, plans, and intentions are first formulated, and upon which appraisal frames are computed. Appraisals are then mapped to multiple emotions using an OCC model, and these emotions are aggregated using an overall emotional state, or "mood." Coping strategies next apply a set of rules to handle the emotions either inwardly, by modifying elements of the model such as probabilities and utilities, or outwardly, by modifying plans or intentions. Christine Lisetti and Piotr Gmytrasiewicz defined specific coping mechanisms as "action tendencies," highlighting their importance in guiding actions.[23]

The BayesACT model described in this article is based on the social-psychological Affect Control Theory (ACT), which defines emotions precisely as vectors in a 3D sentiment space used for sharing and interpreting cultural expectations. Sentiments, also 3D vectors, are inextricably attached to cognitive symbolic interpretations of identities and actions such as words and gestures. The model is grounded in a very tight connection between perceptions and actions through an emotional channel defining cultural expectations, compelling action systems to be guided by a fast cultural heuristic. BayesACT is consistent with research demonstrating affect and uncertainty as heuristics for decision making,[10] but goes a step further by providing a more complete and dynamic model of emotional responses in social situations. BayesACT fundamentally changes the affective computing paradigm by explicitly putting the emotional and volitional "horse" before the cognitive and deliberative "cart."

LIMITATIONS OF CONVENTIONAL AFFECTIVE COMPUTING

Appraisal models have largely been the focus in emotional AI and affective computing. Appraisals usually start by defining a set of variables, each of which models some basic entity such as novelty, control, and uncertainty. However, assessing these variables has some limitations.

In AI, novelty can be used in two contexts. First, in a "decision theoretic" mechanism, novelty acts as a positive reward for exploration: newly discovered elements have the potential to be beneficial, and optimism under uncertainty is the by-product. Second, in an "expectation violation" mechanism, novelty evokes an emotion that guides future action through some coping mechanism. In assistive technology for dementia, framing novelty as a reward or dimension on which to base action choices can be problematic, as memory loss can make events, objects, or persons appear novel when they are not, making modeling difficult without a precise statement of how novelty maps to action as a function of biographical memory.

Further, control and power are normally evaluated with respect to a reward function, as agents who can get higher rewards independently of other agents are more powerful. This feature might be relevant in e-coaching technology, as users could experience feelings of dependence related to their health condition. However, power-based decision making should consider subtle and shifting forms of power and control. For example, older adults with dementia might often feel or express strong control as they are enacting an identity held at some time in the past, yielding a priori expectations and notions of control that are subsequently updated as a result of inference. Finally, uncertainty is a key consideration in developing emotionally responsive e-coaching applications. While many appraisal theories do not incorporate a formal notion of uncertainty, many affective computing and behavioral economics approaches have embraced it.

Overall, most computational models of affect build on conventional AI reasoning techniques (planning, logic, decision theory, and so on) and might fail to capture how people actually make judgments based on culturally and emotionally defined markers. Dimensional theories of emotion have been explored to bridge this gap, and valence has been proposed as a heuristic to guide action based on somatic markers.[6,10] However, these theories provide incomplete support for determining how an agent should use emotion-related information to react to a user in a culturally sensitive way to increase motivation and provide the needed assistance.

> **SENTIMENTS, ALSO 3D VECTORS, ARE INEXTRICABLY ATTACHED TO COGNITIVE SYMBOLIC INTERPRETATIONS OF IDENTITIES AND ACTIONS.**

AFFECT CONTROL THEORY

In contrast to the cognitive-rational models traditionally used in AI, ACT proposes that the main drivers of action are differences between established cultural *sentiments* and transient situational feelings or *impressions*.[24] This affective discrepancy—called *affective deflection* in ACT—generates an initial response to a situational event and interacts with cognitive processing to adjust and refine actions to meet the situation's real-time demands.

ACT proposes that humans learn and maintain a set of shared cultural

FIGURE 1. Example evaluation, potency, and activity [E, P, A] sentiments about the identities "nurse" and "patient" (top), and the deflections and transient impressions for "nurse comforts patient" (middle) and "nurse ignores patient" (bottom), per the USA 2002–2004 lexicon.

affective sentiments about people, objects, and behaviors, and about the dynamics of interpersonal events. Humans use an affective mapping to appraise individuals, situations, and events as sentiments in a 3D vector space of *evaluation* (E; good vs. bad), *potency* (P; strong vs. weak) and *activity* (A; active vs. inactive). These sentiments can be measured, and their cross-cultural consistency has repeatedly been demonstrated in large studies.[15] Humans use these culturally shared sentiments to make predictions about what others will do and to guide their own behavior, making them a keystone of human intelligence. The shared sentiments, and the resulting affective ecosystem of vector mappings, encode a set of social prescriptions that, if followed by all members of a group, results in an equilibrium or social order. This "affect control principle" has been shown to be a powerful predictor of human behavior. While ACT is based in sociological evidence about human interactions with other humans, people also ascribe affective meaning to media and to technological artifacts.[25]

EPA sentiments can be measured with the semantic differential, a survey technique in which respondents rate affective meanings of concepts on bipolar scales. In general, within-cultural agreement about EPA meanings of social concepts is high, and cultural-average EPA meanings from even a few dozen participants are extremely stable over extended periods of time.[24] Sociologists have gathered EPA ratings for numerous concepts across different cultures (USA 1975/1978/2002–2004/2013–2017, Canada 1980–1986/2001–2003, Ireland 1977, Japan 1989–2002, Germany 1989/2007, China 2001) by surveying thousands of people and compiling ACT lexicons or dictionaries that give average EPA ratings for words for which there is consensus. For example, as Figure 1 shows, the EPA for nurse (from the USA 2002–2004 dictionary, average female rating) is [2.9, 1.5, 0.2], meaning that nurses are seen as very good (E), a bit powerful (P), and a bit active (A). A patient is seen as [1.1, −0.8, −0.9]: comparatively less good, powerful, and active than a nurse. A complete description of these datasets can be found in David R. Heise's *Surveying Cultures*[26] or at research.franklin.uga.edu/act.

Social events cause transient impressions of identities and behaviors that deviate from their corresponding fundamental sentiments. ACT models this formation of impressions from events with a minimalist grammar of the form actor-behavior-object (A, B, O). The ACT predictions are based on an empirically derived impression formation function over the nine-dimension space of $\{E, P, A\} \times \{A, B, O\}$.[24] We denote the fundamental sentiment of the actor's evaluation as A_e, and use A'_e for the corresponding transient impression (and similarly for the other eight combinations). The impression formation function consists of linearly weighted polynomial features that combine fundamental sentiments in ways that represent known psychological consistency effects. For example, the transient impression of an actor's evaluation (A'_e) will be positive if (s)he does something positive (B_e), if (s)he does something either positive to a positive person or negative to a negative person (a *balance* effect, represented by a term $B_e \times O_e$ with a positive coefficient), and if (s)he does something weak to a positive person or powerful to a negative person (represented by a term $B_p \times O_e$ with a negative coefficient).

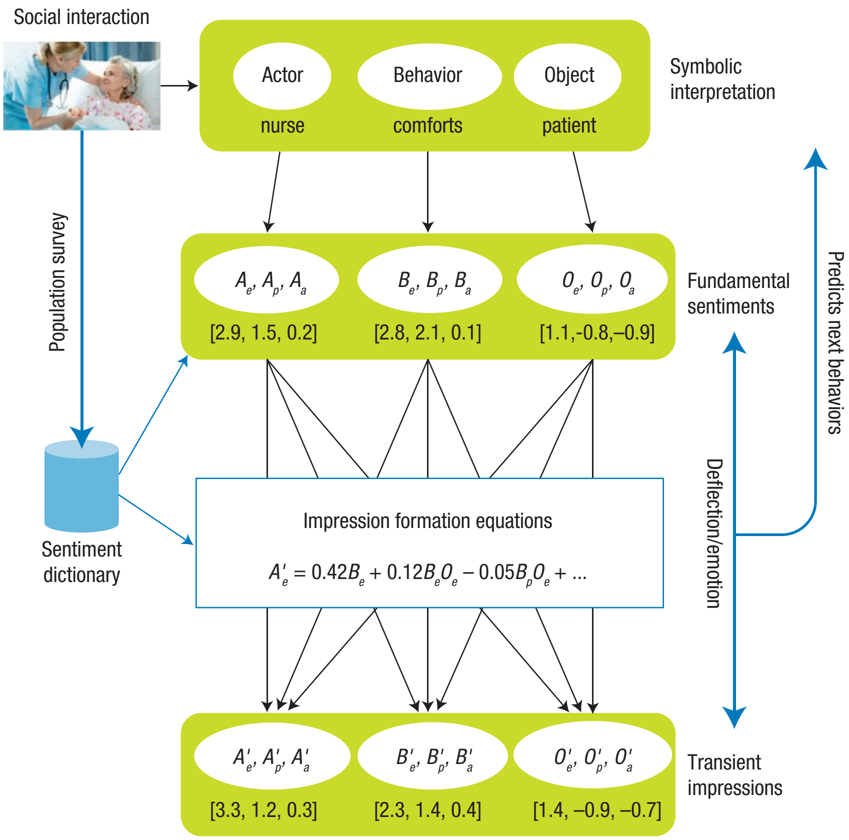


FIGURE 2. Key elements of Affect Control Theory (ACT). An actor (A) performs a behavior (B) on an object (O). A dictionary of empirically measured fundamental sentiments corresponding to the (observed/estimated) behavior and the (known or previously estimated) identities of A and O provides inputs to a set of impression formation equations (also empirically measured). These equations yield transient impressions for A, B, and O. The difference between fundamentals and transients is the deflection, which can be used as a predictor of the most appropriate behavior (for example, a response for the object). The vector difference between fundamentals and transients is the emotion, which is sent as a signal of misalignment.

These three effects are combined with weights showing relative strength as measured in the USA 1978 EPA survey as

$$A'_e = 0.42B_e + 0.12B_eO_e - 0.05B_pO_e + \ldots$$

where we only show 3 of the 20 terms in the full equation. The weighted sum of the squared Euclidean distance between fundamentals and transients is the *deflection* and is hypothesized to correspond to an aversive state of mind that humans seek to avoid or minimize. This hypothesis is known as the *affect control principle*.

As a specific example, consider a nurse (actor) who ignores (behavior, with EPA = [–1.9, −0.3, −0.9]) a patient (object). Observers agree, and ACT predicts, that this nurse appears less nice (E) and less powerful (P) than the cultural average of a nurse (transient EPA = [−0.5, 0.9, 0.3]), while the patient also seems less good and less powerful (transient EPA = [0.4, –1.4, −0.8]). The situation in which a nurse ignores a patient has a deflection of over 13 (very high), whereas if the nurse comforts (EPA = [2.8, 2.1, 0.1]) the patient, the deflection is 1 (very low). After comforting, the nurse's transients are very close to her fundamentals (transient EPA = [3.3, 1.2, 0.3]), as are the patient's (transient EPA = [1.4, −0.9, −0.7]). In general, nurses are expected to do positive and powerful things to a patient, and such actions confirm their culturally defined identity sentiments.

The affect control principle allows ACT to predict deflection-minimizing actions for agents by computing derivatives of the impression formation equations. Emotions in ACT are readouts of the 3D vector differences between fundamental and transient sentiments. They are relayed from one agent to another to promote alignment through a set of gestures, vocal tones, and facial expressions. Figure 2 shows a schematic representation of some of ACT's key elements.

A recent generalization of ACT, BayesACT integrates affective dynamics with decision-theoretic reasoning and explicitly models uncertainty as a key element.[27] BayesACT proposes that emotional processing takes place rapidly and continually, while cognitive-rational processing takes any spare cycles to compute optimal plans within the ecosystem of the affective processing unit. BayesACT views feelings or sentiments as bridging the gaps in a social order, while rationality becomes the "interrupt" mechanism that can be used to handle novelty and to repair breakdowns and disruptions.

BayesACT is thus a type of expectation violation model in which we define expectation violation in a 3D sentiment space and use deflection as a term to describe it. Deflection explicitly provides a policy based on identities, which are interpreted as a motivational force. BayesACT is also a decision-theoretic model, since it can compute plans in the (restricted by affect) game tree to optimize over utility, which might include a novelty measure or other appraisals. As deflection grows, people become less certain of themselves, and the breadth of action choices causes a cognitive overload that leads to short-term solutions. These poor short-term solutions could be disruptive for people and their social context. BayesACT resolves this problem with a single parsimonious model of human affect. When a breakdown occurs, for example, after the user gets frustrated, BayesACT can predict the resulting difficulties to make decisions and suggest correct alternatives, or focus the user on one clear identity.

INTEGRATING ACT IN ASSISTIVE TECHNOLOGY FOR DEMENTIA: THE COACH SYSTEM

To demonstrate how ACT can effectively model affective responses in assistive technology for dementia, we integrated BayesACT into the COACH system and used it to explicitly model the identities of the older adult user and the assistive agent. We used these explicit identity models to generate interventions that were tailored to a simulated user's emotional state.

The baseline function of COACH's hand-washing system relies on a Bayesian sequential model with discrete variables corresponding to the different steps of hand washing, describing the state of the tap (on/off) and hands (dirty/soapy/clean and wet/dry). The older adult user's behavior is modeled by his or her actions: turn on/off water, use soap, use towel, rinse, and null (do nothing). There are probabilistic transitions between plan steps described in a probabilistic plan-graph (for example, the user sometimes applies soap first and sometimes turns on the tap first). A binary variable describes whether the user is aware or not, and this is unobserved but inferred from the user's behavior. COACH tracks the user's hands by classifying individual body parts from a single overhead depth image on a per-frame basis. The tracker outputs the locations of the user's two hands and head, followed by a mapping to a set of predefined spatial regions (soap, tap, sink, water, towel), the index of which is used as the observation of the model's state.

The integration of BayesACT requires that COACH be able to learn information about the user's identity. Luyuan Lin and her colleagues experimented with different identities in simulations, including situational identities such as patient and assistant and biographical identities such as boss or athlete.[28] In practice, COACH uses hand coordinates obtained from the hand tracker as a signal of user emotion on the EPA dimensions. However, due to differences in the ability to automatically determine these emotional factors, BayesACT models the strength of the estimates and can factor weak or less informative signals, concordant with the idea of integrating multiple, differently salient sources of information.[28]

Deflection then plays a critical role by modeling how the interaction will deviate from the goal should the system become misaligned with the person. For example, if the deflection is high, then more unusual (for example, nonconforming) reactions by the older adult are expected. Prompting strategies will be more effective if users are aware of their situation and are well aligned with the system, such that each user shares the same emotional models of identity.

While initial tests of the COACH system were promising, more information was required about how to model identities, including about how persons with dementia perceive themselves and a virtual assistant. This knowledge gap was addressed through a set of interviews with older adults, both with and without dementia. The study involved a semi-structured qualitative interview process involving 12 older adult-care residents and 9 caregivers.[3] The interview guide covered life domains (family and origin, occupation/vocation, personal history, and relationships) and feelings related to an intelligent cognitive assistant. All interviews were transcribed and analyzed to extract a set of affective identities and coded according to the social-psychological principles of ACT. The coding was organized around three themes: biographical identities, current identities, and loss of identity or confusion. Once the main identities were extracted, E, P, and A ratings were generated for each identity using standard semantic differential scales.

Thematic analysis of the interviews showed that a set of identities can be extracted for each participant. Furthermore, the results support the proposition that, while identities grounded in denotative memories fade as abilities to remember people and events are lost, affective aspects of identity in the self-sentiment can persist longer.[3] One resident, for example, associated strongly on an emotional level with biographical identities of father (EPA = [1.9, 1.8, 0.0]) and priest (EPA = [2.2, 1.1, –1.5]), even though he was sometimes being treated as a more helpless identity, such as child (EPA = [1.5, −0.8, 2.1]). This person's affective memories did not find situational support, leading to feelings of inferiority.[3]

Future COACH development will include integration of these findings into the system. Beyond COACH, BayesACT simulations have been conducted to model a range of interactions, including to predict older adult responses to online health tools,[29] demonstrating the model's potential applicability across different settings.

BENEFITS OF BAYESACT INTEGRATION

BayesACT integrates three requirements of emotion modeling into a single reasoning framework:

- predicting how people will feel in emotion-evoking situations, building on appraisal models;[22]
- predicting how people will make decisions in emotional states, building on decision-theoretic models;[21] and
- using emotional responses to infer user personality, building on reverse appraisal models and coherence approaches.[30]

Most existing models that attempt to bridge these three requirements focus on task goals and largely ignore social goals.[10] BayesACT not only integrates social goals but also identities, which serve as heuristics for computing plans to achieve goals. Compared with appraisal theories that often consider power as an intermediate construct, BayesACT also harnesses power as a key construct of social interaction. Finally, a practical advantage of BayesACT lies in its dictionaries of cultural sentiments, which act as a large and validated domain theory.

A key barrier to the use of technology, intelligent or otherwise, in applications such as healthcare has been adoption. Despite the creation and development in recent years of assistive healthcare solutions that have been perceived to be useful in theory, and concerted efforts to bring these technologies to market, widespread adoption has been slow. Customizing technology through the integration of BayesACT to increase the likelihood of acceptance and adoption by targeted users might help to shift these technologies from the research realm to the real world. Through the use of BayesACT, technology developers can attempt to predict the affective responses of their users to specific aspects of their forthcoming applications. For example, an AI agent used as a social networking tool or personal assistant might be designed to act in a manner that the user perceives as friendly for those who prefer these types of interactions and more efficient, curt, and distant for those who do not align well with a friendly assistant. Further, BayesACT is culturally sensitive, and dictionaries of affective meanings and impression formation equation parameters have been gathered in many different countries and languages. The consideration of cultural variations in these measurements would result in different behavior from a BayesACT agent.

BayesACT simulations can also be used to optimize user experience with technology to predict the extent to which the technology is likely to achieve trust on the part of the user—a particularly important aspect of a relationship within the healthcare field. As one example, an AI technology could have the aim of aiding its user in online shopping for medication. For users to have positive experiences with such technology, it seems clear that they must, to some extent, trust that the AI agent will make accurate predictions as to what the user wants and needs, and that it will not use that information in any way that the user would disapprove of. Integration of BayesACT into such technology will allow the technology to predict the affective meanings that the user perceives and to then act in a way that is consistent with those meanings, thus promoting trust, engagement, and adoption.

Overall, BayesACT provides an explicit, computational model of identity that allows for the close integration of emotion and action for a wide range of technology applications.

CHALLENGES OF BAYESACT INTEGRATION

Along with promising potential benefits are significant challenges in integrating BayesACT into assistive technology.

A high-level disadvantage is that ACT models are broad; most evidence pertains to how such models predict average group behavior, with less exploration of how they apply to specific individuals in specific situations.

Further, databases of terms and affective meanings take significant time to compile and, given the rapid progress of technology, it might be difficult for database generation to keep pace. For example, the most recent compilation of terms in Ontario, Canada, is from 2003 and has none that relate to technologies as commonplace as iPhones, Facebook, or Twitter. However, to address this issue various techniques can be used to map from known to unknown words. For example, words could be considered "close" based on their semantic similarity ("cat" is closer to "dog" than to "cab"), and this semantic similarity could be learned by observing real-world usage contexts ("the dog chased the cat" but not "the dog chased the vat"). Closeness in emotional space could then also be considered. Alternatively, both emotional and denotative meanings could be simultaneously learned.

A further limitation stems from the fact that to make accurate predictions, BayesACT requires accurate input of the affective meanings that users have toward themselves, the AI agent, and the actions that the agent performs. Incorrect or biased identity inputs could result in the agent performing actions to which the user is entirely unreceptive, which might hinder the trust in and likelihood of continued use of that technology, and, in the healthcare context, could have negative consequences on users' health. Therefore, a promising direction for future research is the development of technology that can actively and nonobtrusively capture the identity of the user of the technology, and efforts in this area are ongoing.

BayesACT integration also raises ethical issues. Technology that considers the affective meanings users hold toward certain objects, persons, and actions, and customizes itself to align with those meanings, could perpetuate objectively negative biases. For example, most existing virtual assistants have, by default, female voices (Apple's Siri, Amazon's Alexa, Microsoft's Cortana). If, for the sake of argument, this is because people associate femininity with subservience, then intentionally using female voices to accommodate the user base clearly perpetuates a negative stereotype. By gauging the affective meanings that a user holds, a BayesACT-infused agent could act in whatever way possible to accommodate the user, even if that means reinforcing particular stereotypes that the user holds.

Finally, among the potential upsides to integrating AI with BayesACT is the ability to foster trust on the part of the user. However, this potential benefit could also be problematic should technology developers and providers use this trust in ethically unacceptable ways—for example, to sway the user into making purchases or engaging in socially irresponsible acts.

LOOKING TO THE FUTURE

BayesACT models human sentiment, rooted in biographical and situational memories of identity, as a core motivational force for human action. It integrates utility and goals, allowing for appraisal dimensions to be explicitly represented if needed. Finally, it provides a clear definition of emotions as signals of incongruence or dissonance in a social interaction. Its probabilistic framework allows for affective information to be quickly used to establish a social context for an interaction, and to thereafter guide action selection in a way that preserves this same context. When interacting with a person with dementia in particular, it is necessary to maintain this social interaction on an emotional level to prevent breakdowns that most often result in a lack of motivation and a lack of action, and thus to undesired outcomes. This necessity is especially critical as older adults with dementia maintain greater and stronger memories on an emotional level than on a cognitive one.

Looking forward, the success of BayesACT integration will rest on how identity is understood. BayesACT operates using databases of identities that are rated based on affective meanings E, P, and A, but it is possible that other theories of identity can illuminate different important factors to be taken into account. Ongoing research is currently examining the predictive power of BayesACT specifically in the context of technological development for a range of different user groups, including online collaborative networks such as GitHub, and workplace settings with persons having intellectual developmental disability. Empirical evaluations of the effectiveness of the model will be followed by large-scale testing and implementation in existing and emerging assistive technologies.

This article has outlined a novel mechanism to allow for culturally shared emotional meanings to be integrated into assistive technology development. These culturally shared meanings guide and motivate human action. Most commonly used appraisal theories of emotions are limited in their consideration of these shared cultural meanings, and therefore have shortcomings that make them of limited utility in the context of applications for people with dementia. BayesACT addresses many of these limitations and clearly defines a way of ensuring that assistive technology is responsive to users on an emotional level. ◼

ABOUT THE AUTHORS

JULIE M. ROBILLARD is an assistant professor in the Division of Neurology, Department of Medicine, at the University of British Columbia and Scientist in Patient Experience at the BC Children's and Women's Hospital and Health Centres. Her research lies at the intersection of aging, health, and technology, with a current focus on the development of tools for the evaluation of the quality, ethics, and impact of technology for dementia. Robillard holds a BSc in biological sciences from Université de Montréal and a PhD in neuroscience from the University of British Columbia. She is associate director of Neuroethics Canada and sits on the Board of Directors of the Medical Device Development Centre of British Columbia. She is also vice-chair of the Ethical, Legal, Social Impacts Committee of the Canadian Consortium on Neurodegeneration in Aging, a member of the Executive Committee of the Technology and Dementia Professional Interest Area of the Alzheimer's Association International Society to Advance Alzheimer's Research and Treatment, and a Network Investigator of the AGE-WELL Network of Centres of Excellence. Contact her at jrobilla@mail.ubc.ca.

JESSE HOEY is an associate professor in the David R. Cheriton School of Computer Science at the University of Waterloo, where he leads the Computational Health Informatics Laboratory (CHIL), as well as an adjunct scientist at the Toronto Rehabilitation Institute. He works on problems in computational social science, probabilistic and decision-theoretic automated reasoning, affective computing, rehabilitation science, and ubiquitous computing. Much of his research has focused on developing systems to help persons with a cognitive disability such as Alzheimer's disease to engage in daily living activities. Hoey received an MSc in physics and a PhD in computer science from the University of British Columbia. He is a Network Investigator for the AGE-WELL Network of Centers of Excellence. Contact him at jhoey@cs.uwaterloo.ca.

ACKNOWLEDGMENTS

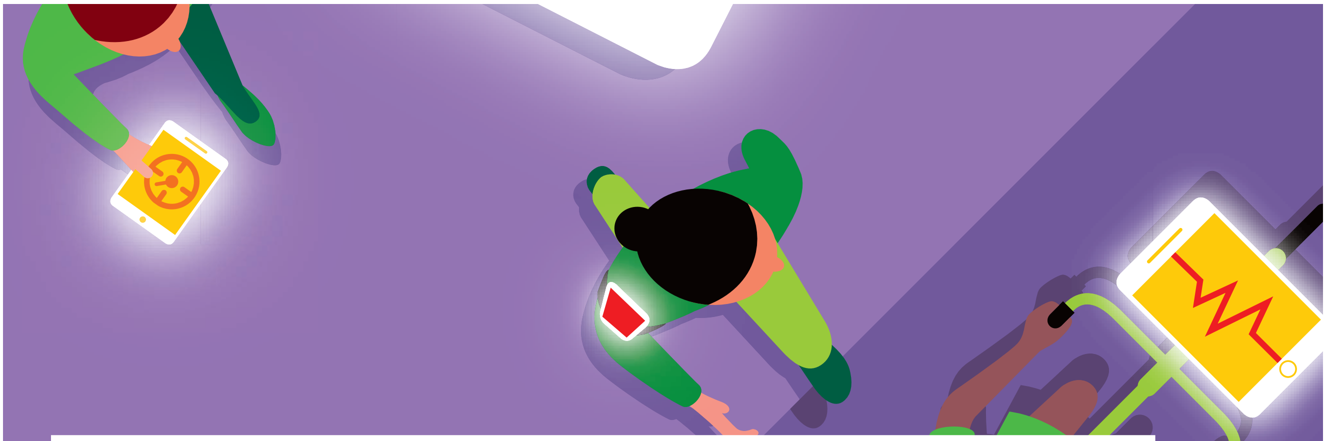
The authors gratefully acknowledge the reviewers and editor for their constructive comments, and Arlo B. Sporn, research assistant at the National Core for Neuroethics at the University of British Columbia, for his contributions to the manuscript, as well as the following funding sources: the Canadian Consortium on Neurodegeneration in Aging, AGE-WELL Inc., The American Alzheimer's Association (ETAC-14-321494), and the Canadian Natural Sciences and Engineering Research Council (NSERC).

REFERENCES

1. E.F. Lopresti, A. Mihailidis, and N. Kirsch, "Assistive Technology for Cognitive Rehabilitation: State of the Art," *Neuropsychological*

Rehabilitation, vol. 14, nos. 1–2, 2004, pp. 5–39.
2. A. Mihailidis et al., "The COACH Prompting System to Assist Older Adults with Dementia through Handwashing: An Efficacy Study," *BMC Geriatrics*, vol. 8, 2008; doi:10.1186/1471-2318-8-28.
3. A. König et al., "Qualitative Study of Affective Identities in Dementia Patients for the Design of Cognitive Assistive Technologies," *J. Rehabilitative and Assistive Technologies Eng.*, vol. 4, 2017; doi:10.1177/2055668316685038.
4. R. Wilson et al., "Quantitative Analysis of Formal Caregivers' Use of Communication Strategies while Assisting Individuals with Moderate and Severe Alzheimer's Disease During Oral Care," *J. Communication Disorders*, vol. 46, no. 3, 2013, pp. 249–263.
5. C. Nugent et al., "TAUT: Technology Adoption and Prediction Tools for Everyday Technologies," *Alzheimer's & Dementia*, vol. 12, no. 7, 2016, pp. P273–P274.
6. A. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain*, reprint ed., Penguin Books, 2005.
7. D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, 2013.
8. H.A. Simon, "Motivational and Emotional Controls of Cognition," *Psychological Rev.*, vol. 74, no. 1, 1967, pp. 29–39.
9. G. Loewenstein and J.S. Lerner, "The Role of Affect in Decision Making," R.J. Davidson, K.R. Sherer, and H.H. Goldsmith, eds., *Handbook of Affective Sciences*, Oxford Univ. Press, 2003, pp. 619–642.
10. P. Slovic et al., "The Affect Heuristic," *European J. Operational Research*, vol. 177, no. 3, 2007, pp. 1333–1352.
11. P.N. Johnson-Laird and K. Oatley, "Basic Emotions, Rationality, and Folk Theory," *Cognition and Emotion*, vol. 6, nos. 3–4, 1992, pp. 201–223.
12. L.F. Barrett, "Solving the Emotion Paradox: Categorization and the Experience of Emotion," *Personality and Social Psychology Rev.*, vol. 10, no. 1, 2006, pp. 20–46.
13. J.A. Russell and A. Mehrabian, "Evidence for a Three-Factor Theory of Emotions," *J. Research in Personality*, vol. 11, no. 3, 1977, pp. 273–294.
14. J.R. Fontaine et al., "The World of Emotions Is Not Two-Dimensional," *Psychological Science*, vol. 18, no. 12, 2007, pp. 1050–1057.
15. C.E. Osgood, G.J. Suci, and P. Tannenbaum, *The Measurement of Meaning*, Univ. of Illinois Press, 1967.
16. R.W. Picard, *Affective Computing*, reprint ed., The MIT Press, 2000.
17. A. Ortony, G.L. Clore, and A. Collins, *The Cognitive Structure of Emotions*, Cambridge Univ. Press, 1990.
18. C. Elliott, "Hunting for the Holy Grail with 'Emotionally Intelligent' Virtual Actors," *ACM SIGART Bull.*, vol. 9, no. 1, 1998, pp. 20–28.
19. C. Conati and H. Maclaren, "Empirically Building and Evaluating a Probabilistic Model of User Affect," *User Modeling and User-Adapted Interaction*, vol. 19, no. 3, 2009, pp. 267–303.
20. K.R. Scherer, A. Schorr, and T. Johnstone, eds., *Appraisal Processes in Emotion: Theory, Methods, Research*, Oxford Univ. Press, 2001.
21. T.M. Moerland, J. Broekens, and C.M. Jonker, "Emotion in Reinforcement Learning Agents and Robots: A Survey," *Machine Learning*, vol. 107, no. 2, 2018, pp. 443–480.
22. J. Gratch and S. Marsella, "A Domain-Independent Framework for Modeling Emotion," *Cognitive Systems Research*, vol. 5, no. 4, 2004, pp. 269–306.
23. C.L. Lisetti and P. Gmytrasiewicz, "Can a Rational Agent Afford to Be Affectless? A Formal Approach," *Applied Artificial Intelligence*, vol. 16, nos. 7–8, 2002, pp. 577–609.
24. D.R. Heise, *Expressive Order: Confirming Sentiments in Social Actions*, Springer, 2007.
25. B. Reeves and C. Nass, *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*, CSLI Publications, 2003.
26. D.R. Heise, *Surveying Cultures: Discovering Shared Conceptions and Sentiments*, John Wiley & Sons, 2010.
27. J. Hoey, T. Schröder, and A. Alhothali, "Affect Control Processes: Intelligent Affective Interaction Using a Partially Observable Markov Decision Process," *Artificial Intelligence*, vol. 230, 2016, pp. 134–172.
28. L. Lin et al., "Affectively Aligned Cognitive Assistance Using Bayesian Affect Control Theory," L. Pecchia et al., eds., *Ambient Assisted Living and Daily Activities*, Springer, 2014, pp. 279–287.
29. J.M. Robillard et al., "Intelligent and Affectively Aligned Evaluation of Online Health Information for Older Adults," *Proc. 2017 AAAI Joint Workshop Health Intelligence*, 2017; cs.uwaterloo.ca/~jhoey/papers/Robillard-aaai2017.pdf.
30. P. Thagard, "Why Wasn't O.J. Convicted? Emotional Coherence in Legal Inference," *Cognition and Emotion*, vol. 17, no. 3, 2003, pp. 361–383.

See **www.computer.org/computer-multimedia** for multimedia content related to this article.



Recommender System Lets Coaches Identify and Help Athletes Who Begin Losing Motivation

Paolo Pilloni, Luca Piras, Salvatore Carta, Gianni Fenu, and Fabrizio Mulas, University of Cagliari

Ludovico Boratto, Eurecat

This article presents a novel approach to monitoring athletes' behavioral changes to predict a decline in motivation. When the system detects such a decline, it refers the athlete to her coach, along with a concise explanation of the detected behavioral changes. The coach thus has all the information needed for a prompt, targeted intervention.

The spread of cheap, always-connected mobile and wearable technologies is enabling and encouraging studies and applications based on e-health persuasive technologies (eHPTs), to foster healthier and more active behaviors.[1] We are investigating the effects of eHPT on users' motivation to exercise[2] and have designed and developed the u4fit (www.u4fit.com) e-coaching ecosystem. The platform helps people connect remotely with coaches, who can then provide tailored, supervised training programs.[3]

A problem that coaches face in such virtual platforms is that the athletes they supervise might stop training without warning. In some cases, the reasons for this might be unpredictable, such as an injury. In some cases, however, athletes might exhibit behavioral changes

 0018-9162/18/$33.00

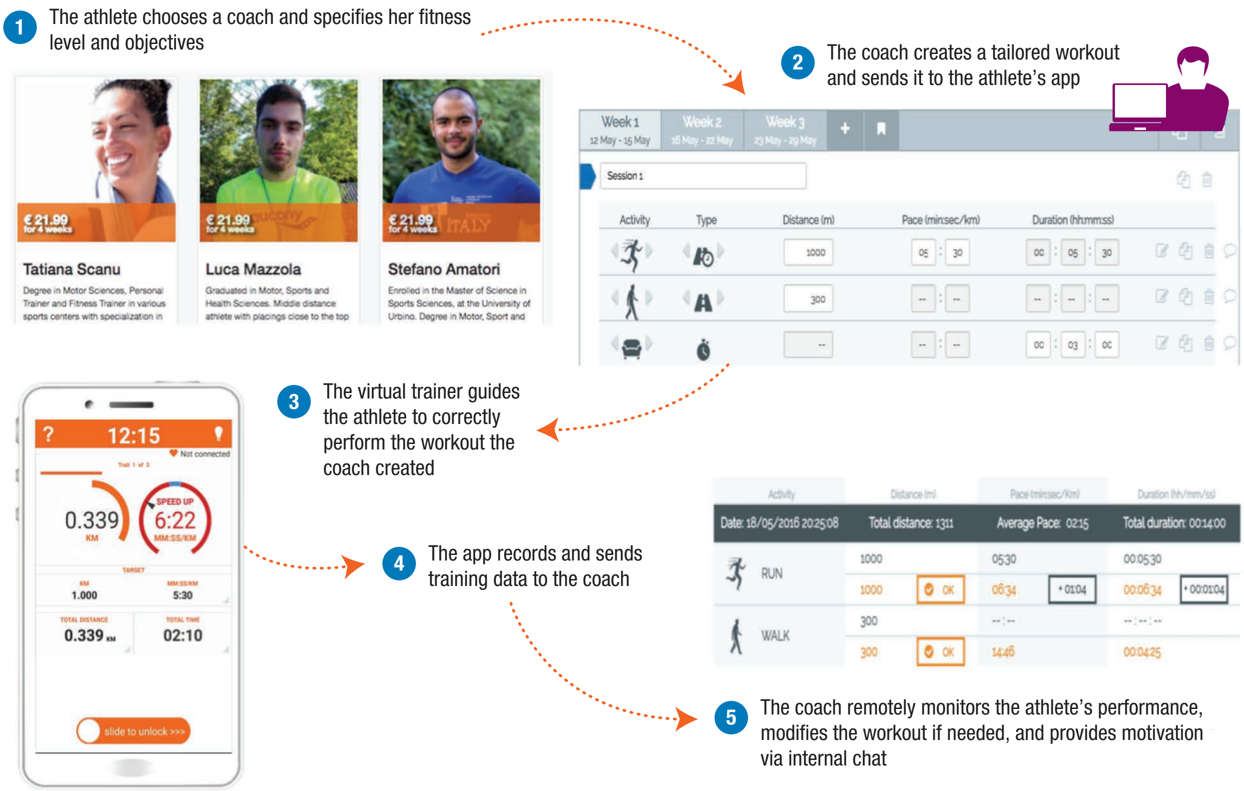


FIGURE 1. Athlete–coach interaction flow in the u4fit e-coaching ecosystem.

in the way they train that could alert coaches to a motivation loss. Predicting such behavioral changes could help a coach intervene to keep athletes motivated. Developing a methodology to accomplish this is difficult but important.

A recommender system could help a coach in this area. These systems, which are increasingly employed in health-related fields,[4] are particularly helpful when users face many options.

In an earlier study, we made a first attempt at exploiting a recommendation system for this purpose.[5] By monitoring athletes' performance, we trained a random-forest classifier to track behavioral changes and predict whether an athlete will soon stop training. The results were good. However, because the classifier employs a set of decision trees and doesn't generate an interpretable output, a coach can't use the information to determine how to get the athlete back on track. More actionable and ready-to-interpret feedback is needed.

To provide this feedback, we present an approach to spot athletes who suddenly leave their exercise routines due to a motivation decline. We based our proposal on the automatic analysis of athletes' training behavior. We trained a decision-tree classifier with athletes' past performances as recorded by our u4fit e-coaching ecosystem with the goal of predicting whether they will stop training within a specific number (x) of days. The system adapts the generated output and uses it to explain why the athlete will stop training via compact trees that make the information easily interpretable by the coach.

U4FIT: E-COACHING ECOSYSTEM

u4fit involves three main actors:

- a web-based front end that serves as users' training diaries and a virtual marketplace for them to contact and book a coach, as well as a way for coaches to manage their athletes;
- a back end in charge of handling business logic; and
- a mobile app that provides real-time guidance to help the user correctly perform the workout plan that her coach designed.

Figure 1 shows the interaction flow between an athlete and coach. Point 1 in the figure illustrates how an athlete chooses the coach that best fits her needs, fills out a questionnaire describing her current fitness level and goals, and finally books the coach. The coach can then interact with the athlete to acquire the detailed knowledge necessary to build a tailored workout plan, as point 2 shows. Once the workout is available, the athlete can start training, guided by the virtual personal trainer embedded in the mobile app, as point 3 shows.

The app tracks workout statistics, such as distances run (computed via the GPS sensor in the user's mobile phone), time spent running, and derived quantities. Data collection starts when the athlete hits the start button on the app and finishes at the end of the prescribed routine (or sooner if the athlete stops beforehand). At this point, the app seamlessly sends the data to the coach's dashboard, as seen in point 4. The coach can then evaluate the results, interact with the athlete about them, motivate her if necessary, and optimize the workout plan over time, as point 5 shows.

We have validated the effectiveness of our system and our mobile client's user experience and its impact on users' motivation to keep exercising.[6]

DATASET

The dataset we evaluated consists of a subset of the 100,000 workout results recorded by u4fit users via the mobile app. The final sample consists of 78,021 workouts performed from 22 September 2012 to 19 July 2017.

Half the users completed at least 17 workouts, and one-fourth completed at least 44. Half the users worked with the platform for at least 4.7 months, and one-fourth worked with it for at least 12.4 months.

For each training session, we considered the following statistics:

- covered distance (in meters),
- workout duration (in seconds),
- rest time (in seconds),
- average speed (in km per hour),
- calories burnt (as a function of user's weight and covered distance), and
- time elapsed since the previous workout (hours).

We excluded results whose statistics could not be considered reliable or that were likely not performed by a beginning or amateur runner. We consider a result to be reliable when

- the distance run is between 0.5 and 43 km,
- the workout lasts no more than 5 hours,
- the rest time is no more than 1 hour,
- the average running speed is no more than 16 km per hour, and
- the number of calories burnt is no more than 3,000.

CLASSIFICATION

We modeled various features and used various classifiers in our study.

Features

We used workout session statistics to model six sets of features that constitute the classifier inputs:

- ABS: This represents the absolute value of an athlete's performance during a workout—as recorded by the app, without any normalization—such as the number of meters that a runner covered. To take into account a user's level of athletic expertise, we considered, as an additional feature, the number of workouts she performed before the current one.
- INC: Each feature in this set is an athlete's change in performance from one workout to the next. We calculate this by subtracting the previous workout's results from those of the current workout and dividing the difference by the previous workout's results.
- MIN, AVG, MAX: These features represent the minimum, average, and maximum user performance.
- WL: We also wanted to consider the impact of the workouts performed by the runner in the previous seven days. We thus obtained weekly load statistics by adding the statistics of the performances of the previous week and factoring in the distance covered, the time spent running, the number of calories burnt, and the number of workouts.

Each workout is represented by 35 features: six each for the ABS, INC, MIN, AVG, and MAX sets, one represented by the number of workouts performed, and four in the WL set.

We performed a binary classification aimed at predicting if a user will stop training within the next x days. During the algorithm's training, we assigned a workout to class 1 if the user performed another workout within the next x days. Otherwise, we assigned the workout to class 0.

Classifiers

Because turning the output of the random-forest classifier into actionable knowledge is difficult, we decided to use it as our baseline and to exploit the decision tree as our actual classifier.

Decision tree is a nonparametric, supervised learning method used for classification and regression. One of its main advantages is that its results are easy to inspect, interpret, and visualize.

EXPERIMENTAL FRAMEWORK

We took a multifaceted approach to our study.

Experimental setup and strategy

Our experimental framework exploits the Python scikit-learn 0.17.1 library for comparing results from random-forest and decision-tree classifiers. We repeated each classification 10 times with a 10-fold cross validation. We ran the classifiers with the default parameters (which typically led to the best results) except for the number of trees we used in random forest (we set the n_estimators parameter to 500) and the number of decision-tree levels (we used max_depth parameter). After evaluating values of this latter parameter ranging from three to 10, we found that five levels guarantees the best performance. We took into account that using many levels would make interpreting the paths traversed in the tree—and thus, developing an intelligible explanation of athletes' behavior—difficult.

To avoid class imbalance, we tested several undersampling and oversampling strategies on the training set, with random undersampling proving to be most accurate. In addition, with $x = 1$ (that is, when we check if the athlete went running on the next day), classes are almost naturally balanced.

We performed four sets of experiments.

Classifiers comparison. We compared the classifiers by running them on all the modeled features for $x = 1$, which had previously yielded the best results.[5]

Evaluation of feature sets' importance. We evaluated each set of features' importance by measuring the DecisionTreeClassifier class' feature_importances parameter in Python scikit-learn.

Evaluation of the classifier with fewer features. We removed the least important feature sets one at a time. This let us evaluate which represent noise in the classification process and which are essential to improving accuracy.

Detection of the most frequent behavioral patterns. We detected the most frequent paths that occurred in a generated decision tree. After repeating our experiments 10 times, we regenerated the trees for each classification. We took our results from one of these trees, whose effectiveness was in line with the average results obtained with the 10-fold cross validation. We used these paths to show examples of behavioral patterns that could help generate recommendations for athletes who exhibit the same behaviors as those who appear likely to stop training soon.

Metrics

We base our metrics on four factors:

- positive instances (P), which are workouts not followed by another one within x days;
- negative instances (N);
- true positives (TP), which are instances of the positive class that the classifier correctly labeled as positive;
- true negatives (TN);
- false positives (FP), which are instances of the positive class that are incorrectly labeled by a classifier; and
- false negatives (FN).

We define accuracy as (TP+TN)/(P+N), which represents the fraction of all instances that are correctly classified. We define recall as TP/P. It measures a classifier's completeness, which is the ability to detect a large number of positive instances. We define precision as TP/(TP+FP). It measures a classifier's exactness, which is the capacity to limit the number of false detections as much as possible.

Our analysis assumes that motivating an athlete who doesn't need to be motivated (false positive) is not as bad as failing to motivate an athlete who is losing motivation (false negative). Hence, we decided to measure the F2 score, which gives more importance to recall than precision and is defined as:

$$F2 = 5\,\frac{Precision \times Recall}{4(Precision + Recall)}$$

None of the metrics presented so far takes into account the true-negative rate (defined as TN/N), which is an issue when one class is bigger than the other.[7] (A classifier that always predicts the majority class would trivially achieve accuracy.) Instead, we decided to measure informedness, which is the clearest measure of a system's predictive value.[8] Its formula is recall + (true negative rate) – 1, and its value ranges between –1 and 1.

Experimental results

Our work yielded several noteworthy results.

Classifiers comparison. Table 1 compares our classifiers' ability to predict whether an athlete will stop working out by the following day ($x = 1$). Results show that both classifiers are effective for our metrics. Indeed, for the metrics conceived in a range between 0 and 1 (all but informedness), the values are high. Informedness returns values close to 0.5, which is in line with the other metrics considering that this metric ranges between –1 and 1. Thus, we can correctly predict whether an athlete will stop training by monitoring her past performances at least 70 percent of the time. And in most cases, we can produce

TABLE 1. Classifiers' effectiveness in predicting whether an athlete will stop working out by the following day.

	Random forest	Decision tree
Accuracy	0.70	0.69
Recall	0.72	0.71
Precision	0.71	0.70
F2	0.72	0.71
Informedness	0.40	0.37

TABLE 2. Importance of each feature set in predicting whether an athlete will stop working out by the following day.

	Importance
Absolute value	0.064
Change in performance	0.03
Weekly load	0.738
Average performance	0.167
Maximum performance	0.012
Minimum performance	0.015

effective recommendations for coaches to keep their athletes training.

Even though random forest achieved the best results, decision tree represented a good tradeoff between a small loss in effectiveness and a large gain in the ability to produce helpful information for coaches.

Table 2 shows the importance of each set of features in predicting whether an athlete will stop working out by the following day. A user's weekly load clearly has more impact in the classification process.

Evaluation of the classifier with fewer features. We removed feature sets one by one to see how this affected the decision-tree classifier's effectiveness. We started with all six feature sets (column 1 in the table), then progressively removed change in performance (column 2), maximum performance (column 3), minimum performance (column 4), absolute value (column 5), and average performance (column 6). Table 3 shows that removing the three least important sets (change in performance, maximum performance, and minimum performance) doesn't decrease effectiveness, which indicates that they don't add any meaningful information to the classification process.

The remaining feature sets show alternating results. For example, weekly load alone (column 6) boosts recall and F2, which leads to a prediction that an athlete is more likely to stop working out. However, it also reduces precision and informedness, which means there is less overall accuracy and more false predictions of athletes losing motivation. Using absolute value, average performance, and weekly load (column 4) yields results that are stable although not optimal. However, the remaining metrics (accuracy, precision, and informedness) reach their highest values.

We can determine the best settings only through further tests on the u4fit platform, which we plan to do in the future.

TABLE 3. Results returned by training decision tree with different sets of features.

	1	2	3	4	5	6
Accuracy	0.69	0.69	0.68	0.69	0.69	0.68
Recall	0.71	0.71	0.71	0.7	0.72	0.78
Precision	0.7	0.7	0.7	0.7	0.69	0.66
F2	0.71	0.71	0.7	0.7	0.71	0.75
Informedness	0.37	0.37	0.37	0.37	0.37	0.34

Detection of the most frequent behavioral patterns. In generated decision trees, we see frequently recurring paths for users who are likely to stop training (class 0) and for users who are not (class 1). Figure 2 shows an example of such paths. Our system can use these short, simple patterns as the basis of explanations for coaches.

By observing the most recurrent paths, it's also possible to notice at a glance some macro behavioral differences between users of the two classes. For example, it's evident that users who work out more than an average of 3.5 times per week are more motivated, and that users who spend less time training are more likely to lose motivation.

Recurrent paths can also help produce recommendations for athletes who probably won't stop training in the next *x* days but whose behavior indicates a loss of motivation. A coach could check if the athlete's current behavior matches a significantly high percentage of one of these patterns. If this happens, the coach could intervene with some mild motivation. We plan to study the effectiveness of this approach with real coaches and athletes when we integrate our recommender system into u4fit.

Class 0:
(589 occurrences, 14.49% of the workouts that we assigned to this class)
Weekly workouts: ≤ 3.5
Average historical hours from last workout: > 63.98
Absolute distance (meters): ≤ 4,325.06
Absolute duration (seconds): > 609.5
Weekly duration (seconds): ≤ 3,824.5

Class 1:
(365 occurrences, 9.76%)
Weekly workouts: > 3.5
Weekly workouts: > 5.5
Average historical hours from last workout: ≤ 30.37
Weekly duration (seconds): > 21,644.0
Avg. historical pause time (seconds): ≤ 52.04

FIGURE 2. Example decision trees for athletes who are likely to stop training (class 0) and those who are not (class 1). The number of weekly workouts, shown in green, can be an important indicator of whether an athlete will stop.

In this article, we advanced the state of the art in the areas of health recommendation systems, e-health, e-coaching, and persuasive technologies with the first approach that can predict if an athlete will soon stop training due to a motivation decline. The approach can exploit these predictions to provide notice of problematic behaviors and help the coach figure out how to persuade an athlete to continue training.

Our study is advanced compared to our previous work. Our classification algorithm and results are different, we use more features, and the dataset covers three more years of training. In addition, the generated recommendations now come with explanations for coaches. Moreover, this approach could generate two types of recommendations: short-term ones when there is an urgent need to motivate a user because she's likely to stop training in a matter of days and long-term ones when users are gradually losing motivation.

We must still investigate this approach's potential via large-scale studies. It would be useful for coaches—as well as health professionals—to know in advance when and why a certain user is losing motivation to exercise. We also want to develop and evaluate various human and automated methodologies for counteracting motivation loss.

With this in mind, we plan to integrate the results of the system's recommendations into the u4fit commercial platform to start evaluating our approach's effectiveness on a large sample of u4fit users. We will also develop new automated strategies to have u4fit's virtual personal trainer act on the recommendations.

ACKNOWLEDGMENTS

The authors thank Alessio Demurtas, Marika Cappai, and Stefano Amico for their contribution to this research. This work is partially funded by Regione Sardegna under Project NOMAD (Next-generation Open Mobile Apps Development) through PIA (Pacchetti Integrati di Agevolazione) Industria Artigianato e Servizi (annualità 2013).

REFERENCES

1. S.J. Curry, "eHealth Research and Healthcare Delivery beyond Intervention Effectiveness," *American J. Preventive Medicine*, vol. 32, no. 5, 2007, pp. 127–130.
2. F. Mulas et al., "Linking Human-Computer Interaction with the Social Web: A Web Application to Improve Motivation in the Exercising Activity of Users," *Proc. 4th IEEE Int'l Conf. Cognitive Infocommunications* (CogInfoCom 13), 2013, pp. 351–356.

ABOUT THE AUTHORS

PAOLO PILLONI is a researcher in the University of Cagliari's Computer Science Department. His research focuses on novel e-coaching and persuasive methodologies to help people in their daily physical activities. He received a PhD in computer science from the University of Cagliari. Contact him at paolo.pilloni@unica.it.

LUCA PIRAS is a scholarship researcher at the University of Cagliari. His main field of interest is machine learning, with a focus on recommender systems, text mining, and social-network analysis. He received a bachelor's degree in computer science from the University of Cagliari. Contact him at lucapiras@unica.it.

SALVATORE CARTA is an associate professor of computer science at the University of Cagliari. His research interests include recommender systems, the social web, human–computer interaction, and e-coaching. He received a PhD in electronics and computer science from the University of Cagliari. Contact him at salvatore@unica.it.

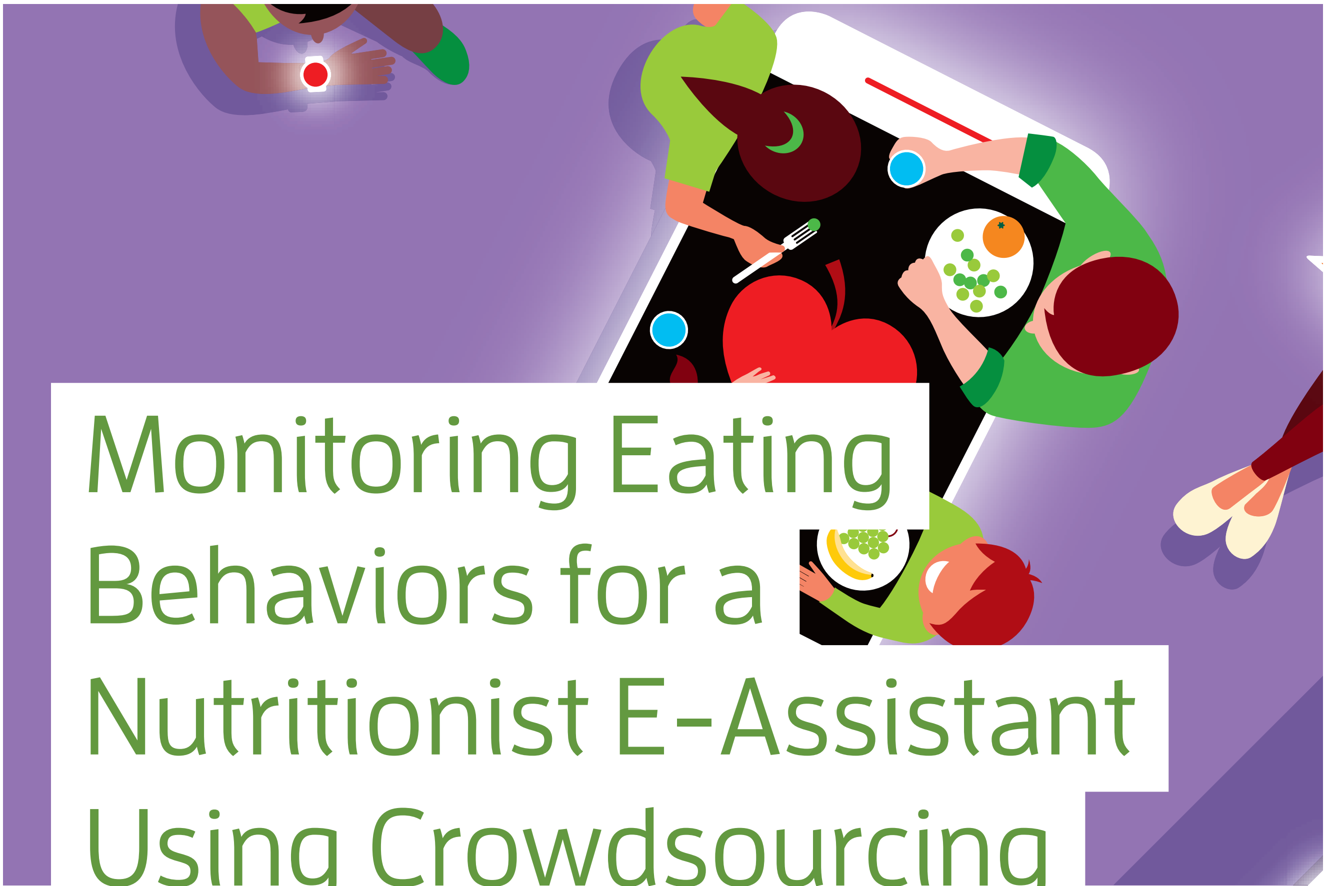
GIANNI FENU is an associate professor of computer science at the University of Cagliari. His research interests include human–computer interaction, ubiquitous computing, and the social web. He received a master's degree in engineering from the University of Cagliari. Contact him at fenu@unica.it.

FABRIZIO MULAS is a researcher at the University of Cagliari. His research interests include persuasive technologies, human–computer interaction, serious games, and the social web. He received a PhD in computer science from the University of Cagliari. Contact him at fabrizio.mulas@unica.it.

LUDOVICO BORATTO is a researcher in Eurecat's Data Science and Big Data Analytics Group. His research interests include recommender systems and data mining in social networks. He received a PhD in computer science from the University of Cagliari. Contact him at ludovico.boratto@acm.org.

3. L. Boratto et al., "An E-Coaching Ecosystem: Design and Effectiveness Analysis of the Engagement of Remote Coaching on Athletes," *Personal and Ubiquitous Computing*, vol. 21, no. 4, 2017, pp. 689–704.
4. A. Calero Valdez et al., "Recommender Systems for Health Informatics: State-of-the-Art and Future Perspectives," *Machine Learning for Health Informatics: State-of-the-Art and Future Challenges*, Andreas Holzinger, ed., Springer, 2016, pp. 391–414.
5. P. Pilloni et al., "Recommendation in Persuasive eHealth Systems: An Effective Strategy to Spot Users Losing Motivation to Exercise," *Proc. 2nd Int'l Workshop Health Recommender Systems*, 2017, pp. 6–9.
6. F. Mulas et al., "Everywhere Run: A Virtual Personal Trainer for Supporting People in their Running Activity," *Proc. 8th Int'l Conf. Advances in Computer Entertainment Technology* (ACE 11), 2011, pp. 70:1–70:2.
7. D.M.W. Powers, "Evaluation: From Precision, Recall, and F-Measure to Roc, Informedness, Markedness, and Correlation," *J. Machine Learning Technologies*, vol. 2, no. 1, 2011, pp. 37–63.
8. D.M.W. Powers, "The Problem with Kappa," in *Proc. 13th Conf. European Chapter of the Assoc. for Computational Linguistics*, (EACL 12), 2012, pp. 345–355.

See **www.computer.org/computer-multimedia** for multimedia content related to this article.



Monitoring Eating Behaviors for a Nutritionist E-Assistant Using Crowdsourcing

Mario O. Parra and Jesus Favela, Ensenada Center for Scientific Research and Higher Education
Luis A. Castro, Sonora Institute of Technology
Arturo Morales, Ensenada Center for Scientific Research and Higher Education

The authors describe Lucy, a digital assistant that monitors eating behaviors to help users lose weight. Lucy's design was informed by a study of clients in a nutrition clinic, as well as by crowdsourcing to evaluate six approaches to assessing nutritional content or caloric intake based on meal photos.

Advances in behavior recognition and monitoring have fostered the development of applications aimed at supporting behavior change. While technologies for tracking behaviors such as caloric expenditure or sleeping habits have matured, tracking eating habits remains essentially unsolved.

Eating behaviors of interest might include eating late at night, prolonged fasting, or regular fast-food consumption. Estimating food intake is particularly relevant to health-related issues such as diabetes and obesity. Nutrition specialists usually tailor dietary regimens to patients' preferences and needs. However, patients might change dishes or ingredients due to availability or preferences. If changes in the dietary regimen are to occur (for example, no eggs available for breakfast), then it is important to make proper adjustments to the diet. Using a crowd to assess what patients are eating can be a convenient way to validate whether patients are following their diet.

RELATED WORK: MONITORING EATING BEHAVIORS

Monitoring eating behaviors has been traditionally addressed with various strategies, including paper-based logs, computational systems that facilitate meal logging, and automatic-recognition approaches.

Paper-based methods for monitoring eating behaviors include food records, food-frequency questionnaires, and forms that include a meal description and time of intake. Mobile phones have been increasingly used to support this task with apps such as Noom (noom.com) or Calorific (calorificapp.com) that help individuals record their meals and coach them on healthful habits. Both of these approaches rely on self-report, meaning that individuals must explicitly enter detailed data regarding their food intake on a daily basis. In the long run, this can be burdensome for individuals being monitored.

Ubiquitous technologies have also been proposed for automatic or semiautomatic recognition of food intake and eating habits, including when or where you are eating, identifying eating disorders, recognizing healthful food, portion size, and so on. Early work in the area addressed the problems of estimating portion sizes.[1] Other efforts include using crowdsourced rating of photos of meals to assess food intake and composition[2] and rating food on a "healthiness" scale.[3] However, the former requires multiple steps before a crowd-based rating can be obtained. On the other hand, the latter only supports rating food in a healthiness continuum (for example, fat versus fit), which can be subjective and limited in terms of food nutritional content. Anne Moorhead and her colleagues[4] compared the ability of experts and nonexperts to accurately estimate calorie count in photos of meals. The findings indicate that nonexperts achieved very low accuracy, while experts had an average error of 8 percent in their estimation.

Additionally, wearables have been used recently for detecting when individuals are eating. For instance, Haik Kalantarian and his colleagues used a sensing necklace to detect throat motion and infer periods of food intake.[5] Also, Yujie Dong and his colleagues implemented a wristband to track eating periods.[6] Other works have aimed to automatically recognize food intake through sensors such as a microphone.[7] Of particular interest to our work are studies aimed at detecting the contents of meals. For example, Austin Meyers and his colleagues presented a system that can recognize the contents of meals from a single image, and then predict its nutritional contents.[8]

Conversely, some commercial mobile phone apps have aimed to help users achieve particular goals. DietApp is a mobile phone app that provides advice for a healthy diet according to age, clinical history, and physical condition.[9] Finally, Gerasimos Spanakis and his colleagues' Think Slim is a mobile phone app that can warn people prior to unhealthy eating events by using mood or location data.[10]

Despite these efforts, estimating what people are really eating has been remarkably challenging. This particular aspect is fundamental to the success of increasingly individual-centric approaches to healthcare (precision medicine), wherein individuals' health and behavior can be automatically monitored to provide patients and caregivers with better care and more accurate and timely information, and, ultimately, to improve quality of life.

We propose the use of crowdsourcing for monitoring dietary intake. In this article, we describe six crowdsourced approaches based on 51 participants' assessment of photos of meals and compare them on accuracy, latency, and cognitive load. We also describe the integration of one of these approaches into a conversational coaching agent to assist individuals who want to change their eating behaviors. For a brief review of related work, please see the sidebar.

RESEARCH PROCEDURE

We worked in collaboration with the "En línea" nutrition clinic, established in 2000 in Ensenada, Mexico. A physician nutrition specialist leads the clinic, supported by another physician and two nurses. In this clinic, approximately 100 patients receive treatment, all with the goal of losing weight.

We carried out the following studies for the design of a nutritionist e-assistant:

- *Stage 1 (S1):* a mixed-method study for understanding nutritionists' and patients' practices

in weight-loss programs to inform the design of an e-coach to assist patients.
- *Stage 2 (S2):* an empirical study for evaluating six photo-based, crowdsourced approaches for inferring food intake.
- *Stage 3 (S3):* development and evaluation of Lucy, a conversational software agent, for assisting patients who participate in the nutritionist program.

S1 and S3 were carried out in the clinic, while S2 was carried out with individuals with no training in nutrition.

STAGE 1: INFORMING THE DESIGN OF A NUTRITIONIST E-ASSISTANT

We conducted a 3-month study in the clinic to better understand the needs of the nutritionist and the individuals undergoing treatment and how to best support them. To this aim, we collected data from various sources. For instance, we interviewed the nutrition specialist, and surveyed 95 patients of the clinic to determine their technology usage and behaviors related to eating, such as whether they took pictures of meals. Moreover, two of the authors attended five weekly sessions at the clinic, as well as participated in the program for three weeks to better understand the challenges faced by the patients. Data were aggregated and later analyzed in interpretation sessions by the authors and the nutritionist to obtain design insights for the nutritionist e-assistant.

Briefly, the patient treatment process followed by the clinic is as follows. At the start of treatment, new patients must complete a short questionnaire about their eating habits. Then, the patients attend an informative session about the clinic methods, and the nutritionist sets a target body weight. After that, the nutritionist designs a one-week meal plan for each patient to follow for the next three weeks. Throughout treatment, patients must attend a weekly group session for nutrition counseling and weight measurement, until the target weight is reached. Patients usually enquire about their diets through occasional calls to the clinic.

The study yielded design insights, which were turned into the following functional (FR) and nonfunctional requirements (NFR):

- *FR1: Meal logging*. Allow the patient to keep a log of meals.
- *FR2: Diet compliance*. System must oversee dietary regime compliance, without demanding additional effort from patients or the nutritionist.
- *FR3: Ingredient exchange*. Provide alternatives for ingredient and dish substitutions.
- *FR4: Meal assessment*. Estimate nutritional content, without giving the nutritionist additional burden.
- *FR5: Nutritional counseling*. Provide nutritional counseling for diet compliance.
- *FR6: Weight tracking*. Oversee progress in weight loss.
- *FR7: Motivation*. Deliver motivational messages based on the patient's goals.
- *NFR1: Reliability*. Answers to patients should be reliable.
- *NFR2: Rapidness*. Answers should be provided quickly.
- *NFR3: Cost-effective*. System should run on hardware available to the patient.
- *NFR4: Mobility*. Patients should be able to access the system anywhere.
- *NFR5: Engaging*. The agent should be engaging either by voice or text.

Informed by these requirements, we propose the design of a conversational agent that provides personalized support to patients during the intervention. The application is aimed at reducing some of the burden of counseling from the nutritionist. In addition, to achieve FR2 and FR4, we propose the use of semiautonomous, crowdsourced-based food intake assessment of patient-taken photographs of meals. The following scenario illustrates how the system can benefit patients:

> *Sara is a 56-year-old teacher concerned about being overweight. Her older daughter is planning to marry in eight months and this has motivated Sara to attend the nutrition clinic. On her first visit, the physician asks her to take photographs of her meals before the second appointment. This information is used by the nutritionist to complete her assessment and recommend a meal plan. Sara also participates in an introductory session in which general advice on the intervention and suggestions for adhering to the diet are provided. She is also asked to download and install "Lucy," an app that can help her follow the program. The clinic personalizes her Lucy account with her current diet, motives, and goals. For the next four months, Sara follows the program, attending sessions, personal consultations every two weeks, and interacting with Lucy. She has found it*

more practical to ask Lucy for the diet of the day than to look at the hardcopy plan. The system also suggests changes she can make when eating out and serves as an assistant to record her weight and remind her of her main motivation and the progress she is making toward it. A feature that Sara particularly enjoys is that sometimes after taking a photo of a meal, Lucy will compliment her for adhering to the plan by having a healthful meal. She understands that Lucy relies on other people to assess the nutritional content of the meals she photographs and is considering volunteering to classify the photos posted by fellow patients.

A key challenge in the design of the e-assistant is the proper assessment of food intake to estimate whether the individual is following the proposed regimen. To address this, in S2 we investigated six crowdsourced, photo-based approaches to analyze the nutritional content of meals.

STAGE 2: CROWDSOURCED ASSESSMENT OF FOOD INTAKE

Assessing food intake and nutritional content involves estimating certain nutritional properties of the food in a photograph. Some properties are difficult to detect; for instance, it is challenging for a person to estimate the amount of carbohydrate, proteins, or fats just by looking at a photograph. Nonetheless, a person can identify salient ingredients, estimate quantities, select similar images, or estimate calories. These activities were the basis of the design of our approaches. Additional considerations included ensuring that individuals can perform the task with no expertise in nutrition and within a few seconds, and that accuracy increases as more people rate the photograph.

We used the following six crowdsourced, photo-based approaches to assess the nutritional content of meals:

- *A1: Number of calories*. The individual estimates the number of calories contained in a meal by looking at the photograph.
- *A2: Food groups*. The individual selects the food groups perceived from the photograph. The participant has to estimate the quantities (none, some, adequate, plenty) of each of the following food groups: fruits, vegetables, cereals, legumes, and animal origin. This is based on the Official Mexican Standard NOM-043-SSA2-2005 for a balanced diet.
- *A3: Healthfulness scale*. The approach is designed to assess the healthfulness of the food in the images. The user rates the photograph on a scale from 1 (not healthful) to 7 (very healthful).
- *A4: Caloric range*. This approach is similar to A1—calories in the meal shown in the photo are estimated by selecting one of six 200-calorie intervals (for example, 401–600).
- *A5: Ingredients*. The user types in all the ingredients that he or she thinks are contained in the meal shown in the photograph, even those not visible, such as salt or cooking oil.
- *A6: Similar images*. The participant selects from a set of nine images the one he or she believes is the most similar to the photograph presented. This action is repeated twice, with the second set of images depending on the first image selected.

We compared these approaches with regard to latency, cognitive load, and accuracy. The ideal crowdsourcing strategy would minimize latency and cognitive load (cost) while maximizing accuracy. We defined and estimated these constructs for our study as follows:

- *Latency*. We defined latency as the time taken to assess one image, measured from the moment the photograph is shown, until the user records the answer.
- *Cognitive load*. We used the NASA-Task Load Index (NASA-TLX) instrument,[11] which uses six subscales (mental demand, physical demand, temporal demand, performance, effort, and frustration measures) to measure the perceived amount of mental effort required by a person to execute a task, and a pairwise comparison between subjects. For this experiment, we removed the pairwise comparison and dropped the physical and temporal subscales because they were not relevant.
- *Accuracy*. We defined accuracy as the degree of confidence in the participant's answers with respect to the ground-truth values of the nutritional content of the meals shown in the photographs. To measure accuracy, we compared the individual values registered by the participants with the ground truth of the images. Because of the uniqueness of the approaches, this

process was different for each of them. For instance, to compute accuracy for A5, we cross-checked the list of ingredients typed in with the actual list of ingredients. We considered different names for the same ingredient (for instance, slice of bread versus wheat bread), typos, and incomplete names. Furthermore, we estimated the collective accuracy of the crowd for a particular photograph, meaning that we computed the mean (or mode, depending on the approach) per photograph and took that value as the answer of the crowd for that particular photograph.

Evaluating crowdsourced approaches for inferring food intake

A total of 51 subjects (8 females, 43 males) participated in the study. Participants were college students majoring in software engineering, with no formal training in nutrition. A database with 45 preclassified images was utilized for this study. We used FatSecret (fatsecret.com.mx) to obtain the nutritional information and ingredients of the food presented in the pictures.

The study was carried out over two days in a computer lab. Each participant required approximately 45 minutes to classify 90 photographs using all six approaches (15 images per approach). Each preclassified image was shown twice using two different approaches. To minimize threats to internal validity, we created 45 different arrangements in which images and approaches were shown to participants (a collection of 15 consecutive images per approach; in each arrangement, both lists of images and approaches were shifted). This was done for the images to receive approximately the same number of ratings across the different participants and approaches, and for the order in which the images were shown to not influence the rating (for example, being tired or bored by the last image). After each approach, the user answered the NASA-TLX questions.

Results

We compared the approaches on latency, cognitive load, and accuracy.

Latency. As Figure 1a shows, participants required, on average, less than 30 seconds to assess one photograph across all approaches. Nonetheless, four approaches had a latency under 15 seconds (see Table 1). As expected, A5 was the approach with highest latency. In contrast, participants took less than 7 seconds, on average, to assess A4 and A3.

Cognitive load. Table 1 shows the cognitive load results in its third column (range 0–100). All approaches scored 50 or less; thus, the approaches do not demand much mental effort from participants. A1 scored the largest cognitive load of all, which might have been due to the fact that estimating the number of calories from a photograph can be frustrating and requires considerable effort (Figure 1b). In contrast, A3 and A6 were perceived as requiring less cognitive load, influenced mainly by mental demand and effort.

Accuracy. In Table 1, the approaches related to calories, A1 and A4, had low accuracy scores. On the other hand, A5 and A6 registered acceptable accuracy, which improved with additional answers from the crowd. It was not possible to calculate the accuracy of A3 because we were not able to assess how healthful a meal was with no additional information; that is, what might be healthful for some might not be so for others. For the other approaches, ground truth was obtained from the source of the image.

Table 1 also shows how the accuracy of each approach changed as more participants (the crowd) classified the images. Again, the approaches related to counting calories, A1 and A4, showed poor results; interestingly, accuracy actually decreased as more answers were collected. Thus, these approaches are clearly not adequate for crowdsourcing. In contrast, the accuracy of the other three approaches (A2, A5, and A6) increased with additional answers from participants. Noticeably, A5 and A6 presented a crowd-based accuracy of 75 percent or higher with the answers of just five participants, reaching nearly 90 percent.

Although A2's accuracy increases, albeit not to acceptable levels, the information conveyed by this approach (the relative amounts of fruits, vegetables, cereals, legumes, and animal origin) is rather limited in terms of monitoring dietary intake. In contrast, from A5 and A6, we can obtain a good estimate of the main ingredients contained in the dish and, thus, obtain a calorie count.

An additional advantage of A5 and A6 is that criteria can be defined for hard-to-assess images, taking into account the varying ratings of the crowd. One such criterion for A6 is to accept a classification when the difference between the image with the most selections and the rest is of at least three votes, having a maximum of 15 votes per image. We used these criteria with the 45 images of the experiment with the following results (average of 30 random runs): the algorithm

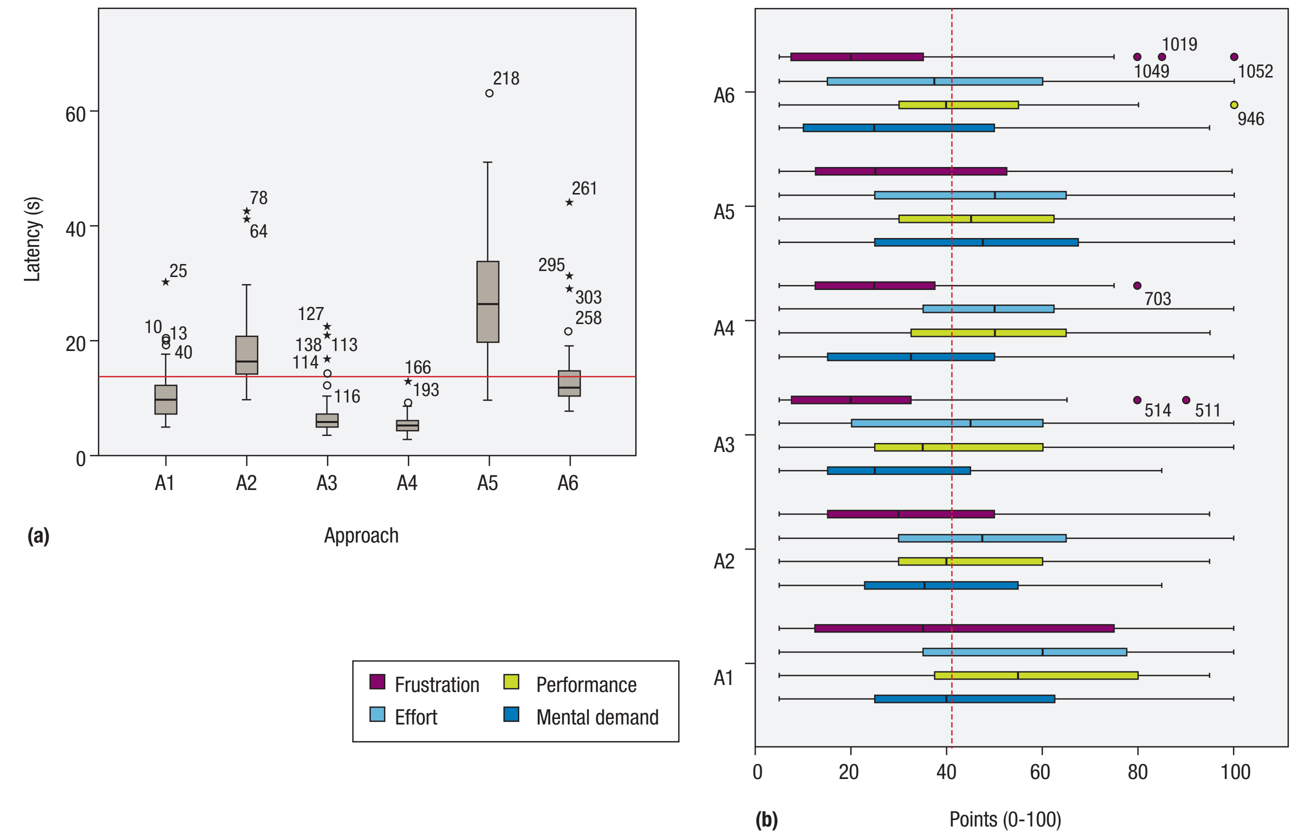


FIGURE 1. Comparison of six crowdsourced, photo-based approaches to assess the nutritional content of meals (see main text). (a) Latency per approach. (b) Cognitive load per approach across the four utilized subscales of the NASA-Task Load Index (NASA-TLX). Circles are suspected outliers; stars are outliers (with their corresponding observation number). Red lines show the overall mean across all approaches.

TABLE 1. Comparison of latency (LAT), cognitive load (CL), and accuracy (ACC) for crowdsourced approaches to assessing the nutritional content of meals.

	Individual			Crowdsourced		
Approach	**LAT (s)**	**CL (%)**	**ACC (%)**	**ACC, n = 5 (%)**	**ACC, n = 10 (%)**	**ACC, n = 15 (%)**
A1	10.60	50.40	28.81	27.63	24.37	22.22
A2	18.49	40.11	34.20	38.89	44.48	57.78
A3	7.04	**32.30**	—	—	—	—
A4	**5.61**	38.72	21.51	13.04	13.78	13.33
A5	28.02	42.39	**80.15**	**88.22**	**88.30**	**88.89**
A6	13.74	33.89	58.66	75.74	85.63	**88.89**

n = number of individual answers taken per image. Numbers in bold represent the best approach per condition.

stopped, on average, at 7.75 votes per image; 78 percent of the images were classified correctly; 5.4 percent were false positives; and the remaining 16 percent were labeled as uncertain, 6 percent of which were false negatives.

The images labeled as uncertain could potentially be sent to an expert for approximating nutritional content.

While some approaches are more suitable than others for certain applications, some of them provide very low accuracy or require considerable effort. For instance, A1 and A4 can easily provide an approximation of caloric intake, but do so with low precision, making them of limited use. On the other hand, A2 can help identify if the dish is well balanced, but is not useful in estimating caloric intake. Finally, A5 and A6 provide a good estimate of caloric intake as well as a good approximation of how balanced a dish is.

STAGE 3: LUCY—A NUTRITIONIST ASSISTANT CONVERSATIONAL AGENT

To assist the patients in the nutritionist program, we developed a conversational agent named Lucy, based on the requirements and scenario described above. The agent's main purpose is to answer frequent, simple questions related to diets and provide general nutritional advice. Lucy was designed to interact with the patient through speech, although it also includes a text interface. Figure 2a shows a simple interaction with Lucy in which the patient is told that today's planned lunch is a tuna sandwich, vegetables, and a piece of fruit. When the patient tells Lucy that she has no bread, Lucy proposes that she use tortillas instead. The dialogue is voice-based, but the app displays a transcript of the interaction as well. Figure 2b shows the day's log of photographs of the meals eaten by the patient that day.

To estimate eating behaviors, Lucy is designed to use a variation of the crowdsourced approach A5 (referred to as A5*). Since the patient has a diet designed by the nutritionist, the crowd member doing the assessment is initially presented with an ingredient list for that meal. The crowd member can mark if each ingredient is present, absent, or not sure. The rater can also type in ingredients, not originally listed, as done in A5. This variant helps deal with the main limitation of A5: latency. We evaluated A5* using the methods reported earlier but with images of meals taken by participants in the intervention. The latency recorded by 15 participants evaluating 15 images was 20 percent lower than for A5.

Lucy offers answers to the following type of queries:

- *Current diet*. What is for dinner today? What will I have for breakfast? (Figure 2a).
- *Ingredient/meal substitution*. The nutritionist offers substitute side items for meals. For instance, when the diet calls for one serving of fruit, this can include a small apple, or half a banana. At the patient's convenience, the meals for one day can be substituted for those indicated for tomorrow or the day after. Lucy can provide notification regarding these alternatives (see Figure 2a).
- *Recipes*. Lucy provides step-by-step instructions to cook all meals provided in the dietary regimen.
- *Log of meals*. The patient can register foods eaten by either telling Lucy or by taking a picture of the meal (Figure 2b). In the latter case, approach A5* will be used to validate that the patient is adhering to the diet. This can be used to make adjustments to the program. For instance, if the participant has eaten a heavy breakfast, the agent will suggest removing or reducing a portion during lunch or dinner.
- *Weight-loss monitoring*. A patient can share with Lucy his or her current weight, which can be queried later on. For instance, the user can ask, "How much weight have I lost this month?" or "How close am I to my target weight?"
- *Motivation*. As suggested in the intervention, the patient can register up to five reasons to lose weight. Lucy will occasionally remind the patient of these when reporting on the weight lost or when explicitly asked to

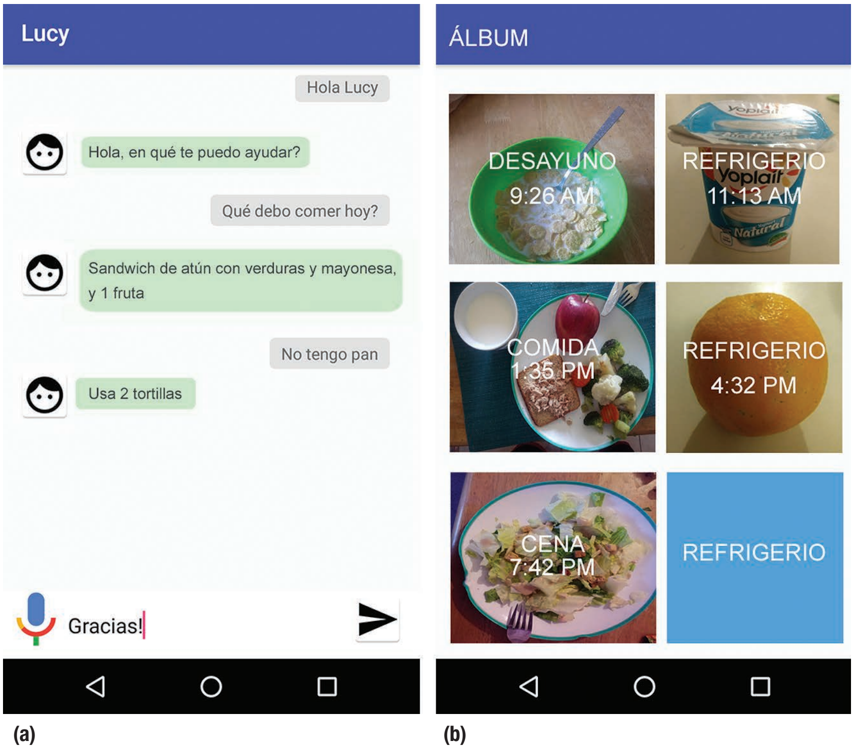


FIGURE 2. The Lucy conversational agent: (a) sample interaction with a patient, and (b) log of photographs of meals consumed in a given day.

ABOUT THE AUTHORS

MARIO PARRA is a graduate student at the Ensenada Center for Scientific Research and Higher Education (CICESE). His research interests include ubiquitous computing, mobile sensing, and pervasive healthcare. Parra received a BSc in software engineering from the Sonora Institute of Technology (ITSON). Contact him at marioparra375@gmail.com.

JESUS FAVELA is a professor of computer science at CICESE, where he leads the Mobile and Ubiquitous Healthcare Laboratory. His research interests include ubiquitous computing, medical informatics, and human–computer interaction. Favela received a PhD in civil engineering from MIT. Contact him at favela@cicese.mx.

LUIS A. CASTRO is a professor at ITSON, where he leads the Interactive and Intelligent Systems Laboratory. His research interests include ubiquitous computing, community informatics, mobile sensing, and human–computer interaction. Castro received a PhD in informatics from the University of Manchester. Contact him at luis.castro@acm.org.

ARTURO MORALES is a graduate student at CICESE. His research interests include human–computer interaction and ubiquitous computing. Morales received a BSc in computer science from the Meritorious Autonomous University of Puebla (BUAP). Contact him at amorales@cicese.edu.mx.

recall them.

The implementation of Lucy is supported by three cognitive services from IBM Watson: Speech to Text, Text to Speech, and Dialog, which is used to indicate different ways in which the user can provide a query. Lucy runs on Android smartphones or tablets.

Evaluating Lucy

We evaluated Lucy with clinic patients. We attended four weekly meetings coordinated by the nutrition specialist to present Lucy. We gave a 10-minute presentation about Lucy, described scenarios of use and demonstrated the system. Afterwards, we asked the participants to complete a questionnaire about user acceptance using the Technology Acceptance Model (TAM).[12] At the end, participants could suggest improvements or offer general comments about Lucy. Finally, they asked Lucy questions on their own.

Fifty-nine patients (70 percent female) participated in the evaluation. Seven participants (12 percent) were 20–29 years old, 17 (29 percent) were 30–39, 10 (17 percent) were 40–49, 16 (27 percent) were 50–59, and 6 were older than 60. Two participants did not provide their age in the questionnaire. Sixteen participants had been in the program for less than a month, 20 had attended for between 1 and 3 months, 9 between 3 and 6 months, and 12 had participated for 6 months or more.

Answers to the 10 TAM items were given on a 7-point Likert scale, with 1 indicating "completely disagree" and 7 "completely agree." We report on the constructs for technology adoption of the TAM model, namely ease of use and usefulness. The average scores indicate that the participants found Lucy to be useful (6.35) and easy to use (6.22) and seemed interested in adopting the system (6.33). The item with the lowest score (6.14) was "It will be easy for me to become an expert in the use of the system." In contrast, the item "Using Lucy would help me with tasks related to my diet" obtained the highest score (6.47). One of the participants gave a "neutral" answer to all questions, and two others had mostly negative ratings. All others seemed rather positive toward Lucy.

Thirty participants (51 percent) provided final comments. Seventeen made suggestions such as allowing personalization of the information and adding new features for Lucy to control water intake, give reminders, connect with a shopping list, and provide encouraging messages. The other 12 responses included mainly praises to the system and the initiative. Finally, a considerable majority (>76 percent) was interested in being notified when the system is available.

We presented Lucy, a digital assistant aimed at helping patients undergoing weight-reduction treatment. Lucy's design was informed by a study that we conducted in a nutrition clinic. To monitor eating behaviors for the assistant, we proposed and evaluated six approaches for assessing nutritional content or caloric intake by a crowd. The approaches offer different advantages in terms of latency, cognitive load, and accuracy. On the basis of our results, we devised a variant of A5 for Lucy, because it offered the highest accuracy with only a few raters required. Although A5 had the highest latency, the proposed variant reduced latency by 20 percent. Both Lucy and the crowd assessment have the potential to facilitate the work of nutrition experts in coaching multiple patients 24/7. Potential users were enthusiastic about adopting Lucy for future use.

Future work includes evaluating the use of Lucy in naturalistic environments to assess participants' performance when the digital assistant is incorporated into their nutritional program. We also plan to propose and evaluate incentives to encourage the crowd to rate photographs of meals.

ACKNOWLEDGMENTS

We thank M.D. Magdalena Gomez, head of the nutrition clinic, and her patients for their kind support. This work was partially funded by Mexico's PFCE 2017 federal program and CONACYT through scholarships provided to Mario Parra and Arturo Morales.

REFERENCES

1. D.A. Williamson et al., "Comparison of Digital Photography to Weighed and Visual Estimation of Portion Sizes," *J. Academy of Nutrition and Dietetics*, vol. 103, no. 9, 2003, pp. 1139–1145.
2. J. Noronha et al., "Platemate," *Proc. 24th Ann. ACM Symp. User Interface Software and Technology* (UIST 11), 2011, p. 1.
3. G.M. Turner-McGrievy et al., "The Use of Crowdsourcing for Dietary Self-Monitoring: Crowdsourced Ratings of Food Pictures are Comparable to Ratings by Trained Observers," *J. American Medical Informatics Assoc.*, vol. 95, 2014, pp. 1103–1108.
4. A. Moorhead, R. Bond, and H. Zheng, "Smart Food: Crowdsourcing of Experts in Nutrition and Non-Experts in Identifying Calories of Meals Using Smartphone as a Potential Tool Contributing to Obesity Prevention and Management," *Proc. IEEE Int'l Conf. Bioinformatics and Biomedicine* (BIBM 15), 2015, pp. 1777–1779.
5. H. Kalantarian et al., "Monitoring Eating Habits Using a Piezoelectric Sensor-Based Necklace," *Computers in Biology and Medicine*, vol. 58, 2015, pp. 46–55.
6. Y. Dong et al., "Detecting Periods of Eating During Free-Living by Tracking Wrist Motion," *IEEE J. Biomedical and Health Informatics*, vol. 18, no. 4, 2014, pp. 1253–1260.
7. Y. Bi et al., "AutoDietary: A Wearable Acoustic Sensor System for Food Intake Recognition in Daily Life," *IEEE Sensors J.*, vol. 16, no. 3, 2016, pp. 806–816.
8. A. Myers et al., "Im2Calories: Towards an Automated Mobile Vision Food Diary," *Proc. IEEE Int'l Conf. Computer Vision* (ICCV 15), 2015; doi.org/10.1109/ICCV.2015.146.
9. I. de la Torre Díez et. al, "A New mHealth App for Monitoring and Awareness of Healthy Eating: Development and User Evaluation by Spanish Users," *J. Medical Systems*, vol. 41, no. 7, 2017, p. 1–7.
10. G. Spanakis et al., "Machine Learning Techniques in Eating Behavior e-Coaching," *Personal and Ubiquitous Computing*, vol. 21, no. 4, 2017, pp. 645–659.
11. S.G. Hart, "NASA-Task Load Index (NASA-TLX): 20 Years Later," *Proc. 50th Ann. Meeting Human Factors and Ergonomics Society* (HFES 06), 2006, pp. 904–908.
12. F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, vol. 13, no. 3, 1989, p. 319.

Programming Spiking Neural Networks on Intel's Loihi

Chit-Kwan Lin, Andreas Wild, Gautham N. Chinya, Yongqiang Cao, Mike Davies, Daniel M. Lavery, and Hong Wang, Intel Labs

Loihi is Intel's novel, manycore neuromorphic processor and is the first of its kind to feature a microcode-programmable learning engine that enables on-chip training of spiking neural networks (SNNs). The authors present the Loihi toolchain, which consists of an intuitive Python-based API for specifying SNNs, a compiler and runtime library for building and executing SNNs on Loihi, and several target platforms (Loihi silicon, FPGA, and functional simulator). To showcase the toolchain, the authors describe how to build, train, and use an SNN to classify handwritten digits from the MNIST database.

Spiking neural networks (SNNs) represent a new model of computing—one that casts computation as the time-dependent, state evolution of a dynamical system, composed of many parallel computing elements (neurons). These elements have local memory, exhibit local state dynamics, and communicate solely via spike impulses. Taking cues from biology, the discipline of neuromorphic computing attempts to adopt the brain's locality, fine-grain parallelism, and event-driven operation in building highly efficient, scalable computing machines, by realizing SNNs in dedicated hardware. To date, there have been a number of efforts to build large-scale neuromorphic systems—ranging from analog designs, such as HICANN[1] and NeuroGrid,[2] to digital ones, such as SpiNNaker[3] and IBM's TrueNorth[4] processor—all exploring different architectures to accelerate and scale up the computation of SNNs.

Recently, Intel unveiled Loihi,[5] our state-of-the-art, fully digital, asynchronous, neuromorphic processor. Loihi features a manycore mesh composed of 128 neuromorphic cores (hereafter, "neuro cores"), three embedded Intel Architecture (IA) cores for managing the neuro cores and controlling spike traffic into and out of the chip, and an asynchronous network-on-chip (NoC) for transporting messages between cores. Most importantly, it is the first of its kind to feature on-chip learning via a microcode-based

0018-9162/18/$33.00

learning rule engine within each neuro core. Programming this engine and the SNNs that use the microcodes requires toolchain features (such as dynamic rule swapping) that are not supported by current end-to-end software frameworks for neuromorphic hardware, such as PyNN,[6] Nengo,[7] and TrueNorth Corelets.[8]

In this paper, we introduce our toolchain (Figure 1) to configure SNNs for Loihi. We present a Python-based API that can be used to specify complex SNN topologies and to program custom learning rules; a compiler and runtime library for building and executing SNNs on Loihi; and, three backend targets—Loihi silicon, an FPGA emulator, and a software functional simulator.

We showcase this toolchain by providing an end-to-end example of a single-layer SNN that is capable of learning and classifying images of handwritten digits from the well-known MNIST database.[9] To acquaint readers to this new model of computing, we will walk through the example step-by-step: from the encoding of the input images into spike patterns, to the design of the SNN topology and supervised learning rules, to their implementation in our API, and to extracting and interpreting the classifier's results.

It is our hope that this toolchain will ease the task of programming SNNs on Loihi and, as a result, be a catalyst for greater participation in neuromorphic computing from the broader research community.

PROGRAMMING MODEL

Existing SNN frameworks[6,7,10,11] all share similarities in the high-level abstractions they present to the programmer. The core primitives we have chosen for this programming model are: neuronal *compartments* and synaptic connections (or *synapses*, for short) as a means of defining SNN topology; *synaptic traces* and a *neuron model* to describe SNN dynamics; and synaptic *learning rules*. We describe these in detail below.

SNN Topology

At the most basic level of abstraction, we describe our SNN as a weighted, directed graph $G(V, E)$ where the vertices V represent compartments, and the weighted edges E represent synapses.

Our programming model also supports the more advanced abstraction of a neuron, which is a set of compartments organized in a tree topology (analogous to dendritic tree[12] structures in biology), with the compartment at the root being the only one permitted to emit spikes. All other compartments participating in the dendritic tree can output state values to parents and integrate such inputs from their children to update internal state. This mechanism provides a way to construct a rich set of nonlinear neuronal input/output mappings.

SNN Dynamics

Both compartments and synapses maintain internal state and communicate information only via discrete-time spike impulses, explicitly incorporating the relative timing of such spike events into a computation. This necessitates some form of temporal information summarization to combine events from different times in a computation. We achieve this through the exponential smoothing

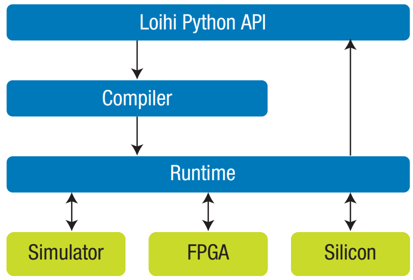


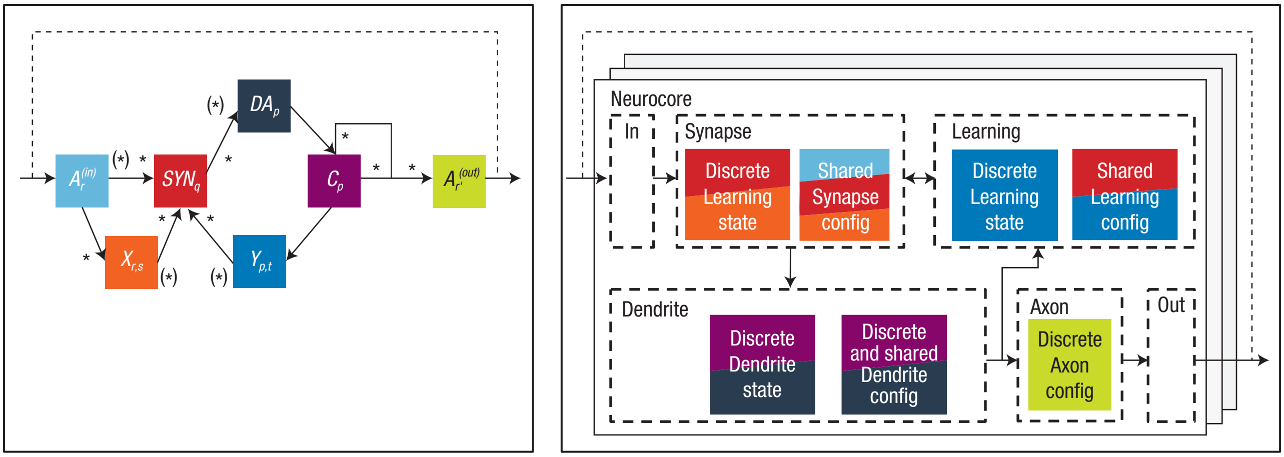
FIGURE 1. The Loihi toolchain consists of a Python API, a compiler and runtime library, and several backend targets, namely, a functional simulator, a field-programmable gate array (FPGA) emulator, and the Loihi silicon.

of spike sequences, the results of which we call traces. For compartments, we maintain "conductance" and "potential" traces, while for synapses, we distinguish "pre-synaptic" (for spikes arriving at a synapse) and "post-synaptic" (for spikes emitted by a compartment) traces.

The specific way in which compartment and synapse states evolve over time and in response to spikes are governed by a neuron model. In our programming model, we assume a variant of the well-known current-based (CUBA) model,[13] defined as a set of first-order differential equations in terms of the above traces and evaluated in discrete algorithmic time steps (the exact formulation is beyond the scope of this paper, but interested readers are referred to the work by Mike Davies and his colleagues[5]).

Learning Rules

The essence of learning for an SNN is the adjustment of its synaptic state variables (for example, weights) in



(a) (b)

FIGURE 2. Computational graph underlying Loihi. Logical view (a): Logical entities of the system. Each box represents an entity-set such as synapses SYN_q or compartments C_p, while arrows represent their associations. $A \rightarrow */* \rightarrow$ annotation represents a one-to-many/many-to-one association between the entities at both ends of an arrow. When connection sharing is enabled, (*)-relations hold as well, namely that multiple axons reference the same set of synapses. Architectural view (b): dashed boxes within a neuro core represent independent asynchronous circuit blocks. Boxes within these circuit blocks reflect the various state and configuration registers. The color coding indicates how logical entities are decomposed and regrouped within architectural implementation. The In/Out blocks are interfaces to the network-on-chip (NoC) and have no correspondence in the logical view. $A^{(in)}$: input axons; $A^{(out)}$: output axons; C: Compartments; DA: dendritic accumulator; SYN: synapses; X: pre-synaptic traces; Y: post-synaptic traces.

response to input and output spike patterns. The way in which synaptic states are updated is governed by a learning rule. In our programming model, this is expressed as a finite-difference equation with respect to a synaptic state variable. The equation must follow a sum-of-products form:

$$Z(t) = Z(t-1) + \sum_m S_m \Pi_n F_n \qquad (1)$$

where Z is the synaptic state variable defined for the source destination neuron pair being updated; t is the discrete time step; S_m is a scaling constant; and F_n may be a synaptic state variable, a pre-synaptic trace, or a post-synaptic trace defined for the neuron pair. Although this sum-of-products constraint does limit the space of possible learning rules, we make this trade-off because it naturally translates into an efficient implementation in hardware.

ARCHITECTURAL OVERVIEW

The Loihi architecture supports the programming model we have just outlined above by providing hardware support for complex multicompartment neurons; multivariable synaptic connections; connection sharing (for efficiently representing convolutional network topologies); and a microcode-programmable learning engine offering different modes of on-chip learning, such as supervised, unsupervised, and reinforcement learning. The Loihi architecture implements this model as a fully digital, asynchronous, event-driven, manycore, compute-in-memory system that supports multichip scalability and utilizes barrier synchronization to deterministically advance algorithmic time. Individual neuro cores in a Loihi chip, are organized in an asynchronous NoC, which also includes several IA cores that are used to process spike I/O and manage system state. Note that the barrier synchronization mechanism abstracts any asynchronous implementation considerations away from the programming model.

Figure 2 illustrates Loihi's computational graph from two perspectives: a high-level view that shows its main logical entities (Figure 2a) and another view of their major architectural counterparts (Figure 2b). Each box in the logical view of Figure 2 corresponds to an entity-set (for example, synapses (SYN_q) or compartments (C_p)) and represents a table containing the various dynamic states and static configuration parameters per discrete entity.

- **Compartments.** Compartments (C_p) are the main building blocks used to configure simple, single-compartment neurons as well as complex, multicompartment neurons. Both integrate incoming spikes as well as input from other compartments. Compartments at the root of a dendritic tree produce a spike when they cross an activation threshold. When learning is enabled, spikes may propagate backwards to update postsynaptic spike traces ($Y_{p,t}$). For each such compartment there may be many output axons ($A_r'^{(out)}$) that deliver spikes to their destinations, reminiscent of multiple axonal arbors in biology.
- **Axons**. Axons connect neurons to synapses. Here, we distinguish between output and input axons ($A_r'^{(out)}$ and $A_r^{(in)}$). In our model, spikes arrive at input axons either from within the network or from external sources of stimulus. An input axon indexes a variable-length list of synapses (SYN_q) and, when learning is enabled, multiple synaptic pre-traces ($X_{r,s}$). Note that the axon construct is not present in the programming model and is introduced here at a lower level of abstraction because it is particular to the hardware implementation.
- **Synaptic traces.** There are multiple pre-($X_{r,s}$) and postsynaptic ($Y_{p,t}$) traces for each input axon and compartment, respectively; they may all use different exponential smoothing parameters. These traces enter the learning engine as proxies of elapsed time between spikes for spike timing dependent plasticity (STDP) or to correlate pre- to post-synaptic activity.
- **Synapses**. Synapses (SYN_q) represent the connectivity in the neural network. They reference the target compartment's dendritic accumulator (DA_p) and support a weight, as well as optional delay and tag states. Synaptic delays enable advanced temporal codes by delaying the accumulation of an incoming spike while tags are useful as an additional scratch variable within the learning engine. All synaptic states can be modified by the learning engine as a function of synaptic states and pre-/post-traces.
- **Dendritic accumulator.** Input spikes trigger an accumulation of a synaptic weight by a compartment, which we refer to as a synaptic activation. The dendritic accumulator (DA_p) acts as a temporal buffer for activations scheduled to be processed at a future algorithmic time step based on the synaptic delay. It also releases the current total activation to the corresponding compartment (C_p).

The architectural implementation in Figure 2b partitions the above logical entities across neuro cores, since each core has limited resources and can hold only a finite number of logical entities. In addition, logical entities are typically decomposed into multiple registers and separated into different independent asynchronous circuit blocks (dashed boxes in Figure 2b) based on circuit optimization considerations. Dynamic state is stored per logical entity while static configuration parameters are stored in shared profile registers per core and are referenced by individual logical entities for resource and algorithmic efficiency. The color coding reflects how logical entities in Figure 2a are realized through the various discrete and shared registers in Figure 2b.

During execution of an SNN, the system of cores operates in three sequential phases within each algorithmic time step:

1. Neuronal updating and spike processing—neuronal and synaptic state is updated locally and in parallel within each neuro core, without the need to access off-chip memory and with neurons only communicating via spikes messages transmitted through the NoC.
2. Learning—synaptic states are updated using a neuro core's integrated learning engine. The learning engine executes microcode that encodes programmable rules specified as a sum of product terms as in Equation (1).
3. System management (optional)—at the end of each

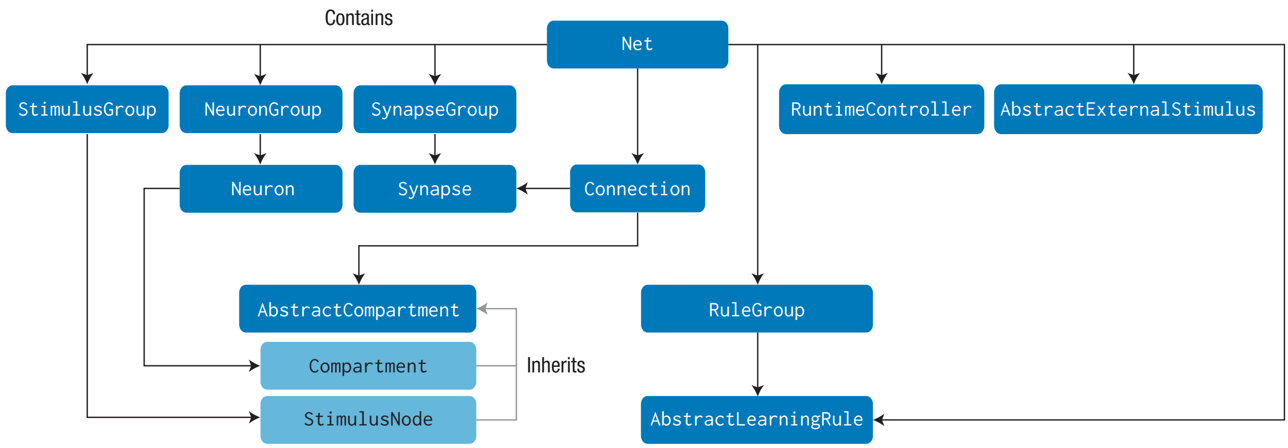


FIGURE 3. Loihi API class hierarchy.

algorithmic time step, there is the option to execute management commands, for example, to reconfigure or inspect state within neuro cores. This is useful for monitoring and debugging the behavior of an SNN.

LOIHI'S API

As described above, Loihi's Python API is an implementation of the programming model and allows the programmer to implement SNNs in an intuitive way, without requiring intimate knowledge of its architectural details. The API provides ways to define a graph of neurons and synapses and configure their parameters (such as decay time constants, spike impulse values, synaptic weights, refractory delays, spiking thresholds); inject external stimulus spikes into the network; implement custom learning rules; and monitor and modify the network state during runtime.

Figure 3 gives an overview of the API class hierarchy. Readers already acquainted with existing SNN simulators such as Brian[10] or libraries such as PyNN[6] will find this API to be comfortably familiar. This is an intentional design choice to lower the barrier of entry to using Loihi.

Specifying an SNN

Loihi's API abstracts the elements comprising an arbitrary SNN topology into the following Python classes:

- `Net`. The root class for specifying an SNN in the API and the top-level container for all elements of the SNN being configured, those containing instances of all the other classes described below. Each Loihi API program must have one instance of `Net`. For convenience, it provides several methods for creating and fully connecting layers of neurons, initialized with reasonable default parameter settings.
- `Compartment`. A `Compartment` models a dendritic compartment of a `Neuron` and is the basic computing element of the SNN.
- `Neuron`. A `Neuron` contains one or more `Compartments`, arranged in a binary dendritic tree. The compartments of a `Neuron` are each configured by calling `configureCompartment(i)` on the *i*-th `Compartment`. This includes setting `Compartment` parameters such as decay constants, refractory delays, and spiking threshold values, which are components to our variant of the CUBA neuron model.
- `NeuronGroup`. A container class for logically grouping `Neurons`. This is often useful for organizing the SNN topology into neuronal layers.
- `StimulusNode`. A `StimulusNode` models a logical point source that injects spikes into the network.
- `StimulusNodeGroup`. Similar to a `NeuronGroup`, a `StimulusNode-Group` is a container for logically

grouping `StimulusNodes`, for example, when representing an input layer that injects stimulus from the external world (including sensors or an image-to-spike encoder and so on) into the `Net`.
- `Synapse`. Each `Synapse` object logically belongs to a `Compartment` and is the end point of a `Connection`. When on-chip learning is enabled for a `Synapse`, its states are adjusted according to its assigned learning rule (see `AbstractLearningRule`, which is described below).
- `SynapseGroup`. A container class for logically grouping `Synapses` for different functional purposes.
- `Connection`. A `Connection` explicitly models a directed edge in the SNN graph where the source vertex is a `Compartment` and the destination vertex is a `Synapse`.

Injecting external stimulus

To control external spike injection into the SNN, the API provides the following:

- `AbstractExternalStimulus`. An abstract base class that models an external stimulus signal source that injects spikes over designated `Connections` at specific time steps. The intended use case is for programmers to subclass `AbstractExternalStimulus` for their specific application, including when they interface with custom signal-to-spike encoders or with physical sensors such as a dynamic vision sensor (DVS) camera.[14]

Incorporating learning

The programmer can implement a wide variety of learning rules through the following:

- `AbstractLearningRule`. An abstract base class that models an on-chip learning rule, which is expressed in sum-of-products form, as we described above. The programmer subclasses this for her own custom learning rules.
- `RuleGroup`. A container class for logically grouping subclasses of `AbstractLearningRule`. Some networks, including our example described below, operate by dynamically swapping learning rules during runtime; we group these rules into a `RuleGroup`.

Runtime monitoring and control

Loihi's API also provides the following facilities for the programmer to record the state of an SNN and control its runtime behavior:

- `Monitor`. A `Monitor` records the state evolution of a SNN during execution. The programmer specifies which network entities' (`Compartments`, `Synapses`, and so on.) state(s) will be recorded and at what interval. After execution, the programmer can inspect the recorded values for data analysis. This is a useful tool for debugging and validating the behavior of a particular SNN implementation.
- `RuntimeController`. The programmer can require runtime control of the SNN while it is executing. `RuntimeController` allows the programmer to issue commands to modify the state of the SNN at a specific time step.

LOIHI'S COMPILER AND RUNTIME LIBRARY

Loihi's compiler takes an SNN implementation and produces a common binary byte stream for various backend targets, including our functional simulator, FPGA, and of course the chip itself. Compilation proceeds in three phases: preprocessing, resource allocation, and code generation.

In preprocessing, the compiler validates that the network parameters specified are supported by the hardware; extracts the necessary information for producing the shared configuration register entries required by the architecture; and extracts the parameter values required to generate a rule's microcode from the sum-of-products form of that learning rule.

Next, the compiler performs resource allocation by greedily assigning network entities (compartments, synapses, learning rules) in the SNN graph to available core resources. To do so, the compiler statically partitions the SNN graph onto as few cores as possible. Assuming there are enough core resources to accommodate the specified SNN, the compiler then performs code generation by walking through all the network entities in the SNN and outputting the byte stream that encodes each network entity onto a specific neuro core, as well as the shared configuration entries required by that neuro core. The conversion into the binary byte stream is performed by calling into the Loihi runtime library, which provides a low-level interface to control the runtime configuration of all the neuro and IA cores that comprise Loihi.

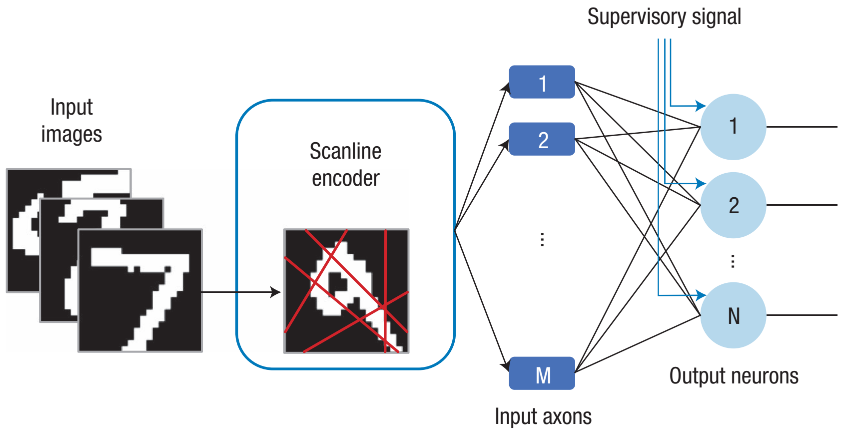


FIGURE 4. Preprocessing steps and spiking neural network (SNN) model for our image classifier. The scan line encoder is an image edge detector that outputs spikes.

LOIHI'S TARGET PLATFORMS

The Loihi compiler supports several backend targets: our functional simulator, a FPGA implementation of the Loihi architecture, and of course Loihi silicon (Figure 1). Each of these target platforms play vital roles in the creation of and application development on the Loihi architecture:

- The Loihi functional simulator implements in software the computational graph described in Figure 3. It is functionally bit-accurate with respect to the FPGA and silicon implementations of Loihi and offers full visibility into the dynamic system state. Thus, it not only supports and eases hardware validation but also enables rapid algorithmic feature exploration as well as future hardware prototyping.
- The FPGA platform emulates the Loihi silicon and is thus crucially important for validation and hardware feature exploration. Beyond this development aspect, it can also be integrated with the Loihi silicon for rapid generation and deployment of architectural Loihi variants that are optimized for specific workloads, that is, trade off efficiency of the hardware for programmability.
- The silicon target ultimately implements the Loihi architecture in the most energy-efficient, performance-optimized, and memory-dense way possible.

EXAMPLE OF A SINGLE-LAYER SNN FOR SUPERVISED IMAGE CLASSIFICATION

With the Loihi toolchain description complete—including its API, runtime monitoring and control, and target platforms—we now showcase an end-to-end example that uses it to program an SNN onto Loihi. This example SNN has a single layer of output neurons (Figure 4) and can be trained in a supervised manner to classify images from the MNIST database.[9] We will describe the model and then review the training and inference implementations.

Model description

All images for both training and inference are first preprocessed and encoded into spikes before being injected into the SNN. The preprocessing consists of binarizing each image vector and encoding it with a so-called scan line encoder, an edge detection procedure inspired by human saccadic eye movements. This encoding exploits the large degree of topological redundancy in natural images by sampling each image sparsely along certain directions ("scan lines" are shown in red in Figure 4), while remaining robust against small local image deformations such as feature shifts or stretches. The encoder produces spikes along each scan line where the scan line is incident on an edge in the image. We present each encoded training image to a single-layer SNN and train it by applying a supervisory signal (via the learning rules) to the output neuron with the index matching the ground truth label of the input image. During inference, the index of the most active output neuron (as identified by spike count) is the predicted class of the current input.

As a supervised learning rule, we use an event-based formulation of the Widrow-Hoff rule,[15] which we call supervised STDP (S-STDP) and write in the sum-of-products form understood by the learning engine (as in Equation (1)):

$$W_{i,j}(t) = W_{i,j}(t-1) + y_{i,0} \cdot S_1 \cdot x_{j,1} + u_K \cdot S_2 \cdot x_{j,1} \qquad (2)$$

$$W_{i,j}(t) = W_{i,j}(t-1) + y_{i,0} \cdot S_1 \cdot x_{j,1} \qquad (3)$$

Here, t is the time step, $W_{i,j}$ is the synaptic weight between the j-th input (pre-synaptic) and i-th output (post-synaptic) neurons, and $S_1 < 0$, $S_2 > 0$ are constants. The spike output of the network is represented by the post-synaptic (namely, output) spike sequence $y_{i,0} \in 0, 1$ and the supervisory signal is encoded into a sequence of periodic training impulses, $u_\kappa \in 0, 1$ with periodicity $\tau^{(epoch)} \cdot 2_\kappa$. The execution of both learning rules is gated on the pre-synaptic trace $x_{j,1} > 0$, specifically, the j-th input neuron being stimulated by the current input spike signal.

Synapses whose post-synaptic neuron corresponds to the class of the current input image use the potentiating rule in Equation (2), whereas all others use the depressive-only rule in Equation (3). As a result, each neuron develops a selectivity for the difference between the average in-class and the average out-of-class image vectors.

Training

We implement this SNN using Loihi's API as follows. First, we create a `Net` container for the entire network and instantiate a `StimulusNodeGroup` representing the input layer and a `NeuronGroup` for the output layer in Figure 4 (see sequence in Figure 5a).

Here the parameters to `createNeuronGroup` correspond to the activation threshold, the two compartment trace decay constants, and the refractory period for each `Neuron` being instantiated in the `NeuronGroup`. These parameters are required by the variant of the CUBA model that Loihi implements.

Next we instantiate the two learning rules described by Equations (2) and (3), and connect the input with the output layer (see sequence in Figure 5b).

As can be seen by the syntax of the initialization parameter `productTerms`, the learning rules directly reflect the sum-of-product terms in Equations (2) and (3), respectively. In both cases, we have also used learning rule parameters for $S_1 = -2^{-5}$ and $S_2 = 2^{-5}$, which are specified as the `coeff` and `coeff_exp` parameters.

We encode all the training set images via `LineScanStimulus`, a subclass of `AbstractExternalStimulus` that implements the scan line encoder described above (see Figure 5c).

For training, we will need to dynamically reassign the appropriate learning rule (2) or (3) to the synapses of different output neurons at runtime. To achieve this, we create and initialize a `RuntimeController` with the `groundTruthLabels` of the training set. This assignment allows the `RuntimeController` to issue the proper network reconfiguration commands for each image during execution (see Figure 5d). For example, an input image with ground truth label *i* results in the `RuntimeController` assigning the `p_rule` to the synapses

```
snn = Net()
numExtInputs = 1920
inputLayer = snn.createStimulusNodeGroup( numExtInputs)
numNeurons = 10
outputLayer = snn.createNeuronGroup(numNeurons,
                                    threshold=1024,
                                    decayU=410,
                                    decayV=410,
                                    refractory=2)
```

(a)

```
p_rule = SSTDPPotentiationRule(productTerms=[('Y0', 'X1'), ('X1',)],
                               coeffs=[-1, 1],        # S_1, S_2
                               coeff_exps=[-5, -5],
                               preTraceDecay=[10,],
                               preTraceAcc=[127, ])
d_rule = SSTDPDepressionRule(productTerms=[('Y0', 'X1')]
                             coeffs=[-1],            # S_1
                             coeff_exps=[-5],
                             preTraceDecay=[10,],
                             preTraceAcc=[127, ])
rules = SSTDPRuleGroup( ruleList=[d_rule, p_rule])
snn.fullyConnect(inputLayer, outputLayer, rules)
```

(b)

```
lss = LineScanStimulus(training_images, inputLayer, snn.connections)
snn.addExternalStimulus(lss, lss.getDuration())
```

(c)

```
if rtctl is None:
        rtctl = RuntimeController(timeoffset=GUARD_INTERVAL,
                                  timestepping=TIMESTEPS_PER_IMAGE)
        snn.addRuntimeController(rtctl)
rtctl.appendLabels(groundTruthLabels)
```

(d)

```
mon = Monitor([synapse for synapseGroup in snn.synapseGroups
                       for synapse in synapseGroup],
                       [SynapseState.WEIGHT],
                       dt=num_instances*TIMESTEPS_PER_IMAGE
                           +GUARD_INTERVAL)
snn.addMonitor(mon)
```

(e)

```
sim = PySimulator.Simulator(snn)
sim.simulate(nsteps, reportProgress=True)
wgts = mon.getStateValues(sim, SynapseState.WEIGHT)
numpy.savetxt(wgtsFilename, wgts.values()[0], fmt="%d")
```

(f)

FIGURE 5. Example Loihi API program for supervised training of an SNN for image classification.

```
snn.fullyConnect(inputLayer, outputLayer,
                 weight=numpy.loadtxt(wgtsFilename,
                 dtype=int).reshape(numExtInputs,numNeurons))
neuronIdx = [neuron.compartments[0] for neuron in outputLayer]
spikemon = Monitor(neuronIdx, [CompartmentState.SPIKES])
snn.addMonitor(spikemon)
```
(a)

```
lss = LineScanStimulus(test_images, inputLayer, snn.connections)
snn.addExternalStimulus(lss, lss.getDuration())
rtctl = RuntimeController(timeoffset=GUARD_INTERVAL,
                          timestepping=TIMESTEPS_PER_IMAGE)
rtctl.clearAllCx(range(GUARD_INTERVAL,
                 len(imglist)*TIMESTEPS_PER_IMAGE,
                 TIMESTEPS_PER_IMAGE))
snn.addRuntimeController(rtctl)
sim = PySim.Simulator(snn)
numSteps = GUARD_INTERVAL+ len(imglist)*TIMESTEPS_PER_IMAGE
sim.simulate(numSteps)
out = spikemon.getStateValues(sim, CompartmentState.SPIKES)
```
(b)

FIGURE 6. Example Loihi API program that uses a trained SNN for inference in image classification.

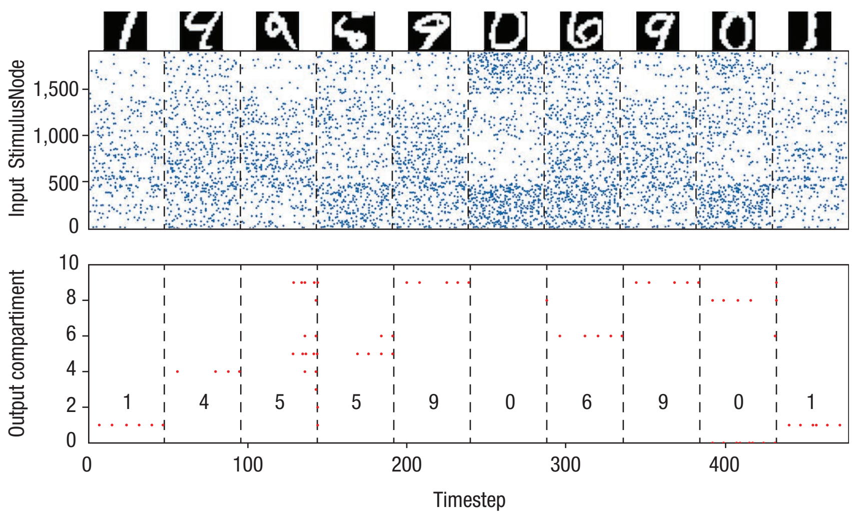


FIGURE 7. Inference results for the single-layer SNN previously trained on the MNIST data set. The top panel shows the input image and the corresponding spike encoding (blue). The bottom panel shows the output spiking activity (red) with the classification result overlaid. In this example, the SNN correctly classifies all but the third image.

of the *i*-th output neuron whereas all the other synapses are assigned the d_rule. Thanks to indirection in Loihi's learning rule assignments, the rule changes can be applied with just a single register write per neuron (see Figure 5d).

Finally (see Figure 5e), because we will want to export the trained model (namely, the synaptic weights) after the SNN is trained, we set up a Monitor on the synapses and set the recording interval dt such that only the last time step is recorded (it is, of course, possible to set it to a shorter interval to record intermediate weights during training).

To execute this SNN for training, we first invoke the compiler on the Python program. This produces a byte stream (in our case a binary file) that can be loaded and run on Loihi or, for example, the functional simulator. Below, we launch the simulator and, after the simulation completes, we save the trained weights for future use during inference (see Figure 5f).

Inference

For inference, we instantiate the same SNN topology as above, but instead load the previously trained weights. We have also added a spike Monitor to record output spikes in order to determine the activity of the neurons in the output layer and thus the classification result (see Figure 6a).

During the inference phase, we encode all test images using LineScanStimulus. After each image's spike train is injected into the SNN, we reset compartment state; this is scheduled by providing the appropriate reset time steps to the RuntimeController.clearAllCx() method (see Figure 6b).

The inference results are shown in Figure 7. The top panel shows the scan line spike encoding corresponding to each input image as it is injected into the network (blue dots indicate spikes). The bottom panel shows the SNN's output spikes in red, overlaid with the corresponding classification result. The trained network correctly classifies most inputs, achieving a classification accuracy of 96.4 percent on the standard MNIST test set.

We have presented an end-to-end toolchain for programming Loihi, our novel neuromorphic processor. The toolchain includes a Python-based API that allows the programmer to quickly implement SNNs capable of both learning and inference; a compiler and runtime for building and executing these programs on our platform; and several target Loihi platforms. We envision that this toolchain will not only lower the barrier for the research community to develop neuromorphic applications on the current generation of Loihi, but also serve as a means to collaboratively explore new features for future hardware generations. C

ABOUT THE AUTHORS

CHIT-KWAN LIN is a researcher with Intel's Microarchitecture Research Lab. Contact him at chit-kwan.lin@intel.com.

ANDREAS WILD is a researcher in Intel's Neuromorphic Computing Lab. Contact him at andreas.wild@intel.com.

GAUTHAM N. CHINYA is a researcher and Principal Engineer in Intel's Microarchitecture Research Lab. Contact him at gautham.n.chinya@intel.com.

YONGQIANG CAO is a researcher in Intel's Microarchitecture Research Lab. Contact him at yongqiang.cao@intel.com.

MIKE DAVIES is director of Intel's Neuromorphic Computing Lab. Contact him at mike.davies@intel.com.

DANIEL M. LAVERY is a principal engineer at Intel. Contact him at daniel.m.lavery@intel.com.

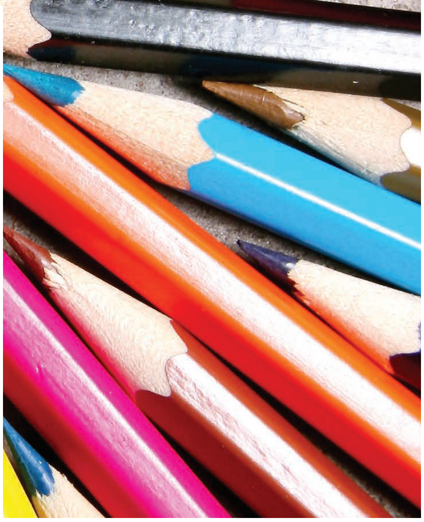
HONG WANG is an Intel Fellow and director of Intel's Microarchitecture Research Lab. Contact him at hong.wang@intel.com.

REFERENCES

1. J. Schemmel et al., "A Wafer-Scale Neuromorphic Hardware System for Large-Scale Neural Modeling," *Proc. IEEE Int'l Symp. Circuits and Systems* (ISCAS 10), 2010, pp. 1947–1950.
2. B.V. Benjamin et al., "Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations," *Proc IEEE*, vol. 102, no. 5, 2014, pp. 699–716.
3. S. Furber et al., "Overview of the Spinnaker System Architecture," *IEEE Trans. Computers*, vol. 62, no. 12, 2013, pp. 2454–2467.
4. P. Merolla et al., "A Million Spiking-Neuron Integrated Circuit with a Scalable Communication Network and Interface," *Science*, vol. 345, no. 6197, 2014, pp. 668–673.
5. M. Davies et al., "Loihi: a Neuromorphic Manycore Processor with On-Chip Learning," *IEEE Micro*, vol. 38, no. 1, 2018.
6. A.P. Davison et al., "Pynn: a Common Interface for Neuronal Network Simulators," *Frontiers in Neuroinformatics*, vol. 2, no. 11, 2009.
7. T. Bekolay et al., "Nengo: A Python Tool for Building Large-Scale Functional Brain Models," *Frontiers in Neuroinformatics*, vol. 7, no. 48, 2014.
8. A. Amir et al., "Cognitive Computing Programming Paradigm: A Corelet Language for Composing Networks of Neurosynaptic Cores," *Proc. IEEE Int'l Joint Conf. Neural Networks* (IJCNN 13), 2013.
9. Y. LeCun, C. Cortes, and C. Burges, "The MNIST Database of Handwritten Digits," 1998; http://yann.lecun.com/exdb/mnist.
10. D. Goodman and R. Brette, "Brian: A Simulator for Spiking Neural Networks in Python," *Frontiers in Neuroinformatics*, vol. 2, 2008.
11. M. Beyeler et al., "Carlsim 3: A User-Friendly and Highly Optimized Library for the Creation of Neurobiologically Detailed Spiking Neural Networks," *Proc. IEEE Int'l Joint Conf. Neural Networks* (IJCNN 15), 2015.
12. B. Mel, "Synaptic Integration in an Excitable Dendritic Tree," *J. Neurophysiology*, vol. 70, no. 3, 1993, pp. 1086–1101.
13. T. Vogels and L. Abbott, "Signal Propagation and Logic Gating in Networks of Integrate and-Fire Neurons," *J. Neuroscience*, vol. 25, no. 46, 2005, pp. 10786–10795.
14. P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 db 15μs Latency Asynchronous Temporal Contrast Vision Sensor," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, 2008, pp. 566–576.
15. F. Ponulak and A. Kasiński, "Supervised Learning in Spiking Neural Networks with Resume: Sequence Learning, Classification, and Spike Shifting," *Neural Computation*, vol. 22, no. 2, 2010, pp. 467–510.

The Maker Revolution

Scooter Willis, TechGarage



The maker revolution, in which 3D printers and other innovative technologies let us produce many types of prototypes and products quickly and easily, is profoundly affecting research and industry.

We are experiencing a modern-day version of the industrial revolution known as the maker revolution. Gone are the days where an idea for a new product requires a long design and manufacturing lifecycle just to get a sample. You no longer need to meet with a tool and die maker and commit to a large number of widgets before beginning production.

The maker revolution began in 2009 with the introduction of a low-cost printing kit from MakerBot Industries. Once you assembled and tested your RepRap 3D printer, MakerBot encouraged you to print the parts needed to build another 3D printer for a friend. In 2013, Stratasys, a leading commercial 3D-printing company, acquired MakerBot for $604 million.

With the recent advances in rapid prototyping that target the maker revolution, I want to share what we are using at the TechGarage makerspace and what we can't wait to purchase.

3D PRINTING STARTED IT ALL

A couple of years ago, we went through a dark period during which 3D printers weren't used much because the odds of failure were high. Based on product reviews from *Make:* magazine (makezine.com/comparison/3dprinters), we at TechGarage decided to invest in Lulzbot Taz 6 3D printers. We now have six. They stay busy and have proven to be reliable. We've also had good results from the Prusa i3 printer.

At January's International Consumer Electronics Show (CES) in Las Vegas, I came across XYZprinting's da Vinci Color 3D printer (www.xyzprinting.com/en-US/product/da-vinci-color), which uses a colorless polylactic acid (PLA) filament and lays down color ink as part of the printing process. Typically, 3D printers use filaments to produce color. At CES, XYZprinting showcased a selection of 3D printer models with full-color features.

In the same way that we went from black-and-white to color copiers, I anticipate a future with highly reliable and affordable color 3D printers that allow product printing on demand.

 0018-9162/18/$33.00

EDITOR **SCOOTER WILLIS**
Avera Cancer Institute;
willishf@gmail.com

LASER CUTTING FOR EVERYONE

One of 3D printers' shortcomings is the print time needed to produce large objects. Two years ago, we received a grant that let us purchase a 60-W laser cutter. It was transformative because we could cut parts out of 1/4-inch-thick acrylic or wood with dimensions up to 18" x 24". Cuts are fast and precise and enable a wide range of rapid parts manufacturing. The need for a grant illustrates that commercial 60-W laser cutters aren't cheap; they typically cost from $15,000 to $30,000. From my perspective, though, it was one of our best investments in the past 18 months, as we have used it to cut 2,000 different parts.

Fortunately, Glowforge Inc. has introduced, via Kickstarter, a $3,995 40-W laser cutter targeting the home market. It has numerous important features and is easy to use. The professional version costs $6,995. The integrated camera determines where to start cutting by detecting the location of the pattern that was just cut. The camera also identifies a barcode on materials so that it can automatically establish settings for wood, acrylic, leather, rubber, cardboard, fabric, and other materials.

5-AXIS CNC ON A BUDGET

3D printers make plastic parts that lack aluminum's durability. Although there are examples of exotic 3D print materials with the look and feel of metal or wood, parts made with these materials aren't as durable as if they were made of aluminum.

Computer numerical control (CNC) is the standard process for making an aluminum part. It starts with a solid block of aluminum and carves away metal. This differs from 3D printing, which fabricates parts by adding material. CNC machines start at $5,000 for hobby-level versions and $50,000 for production-manufacturing versions.

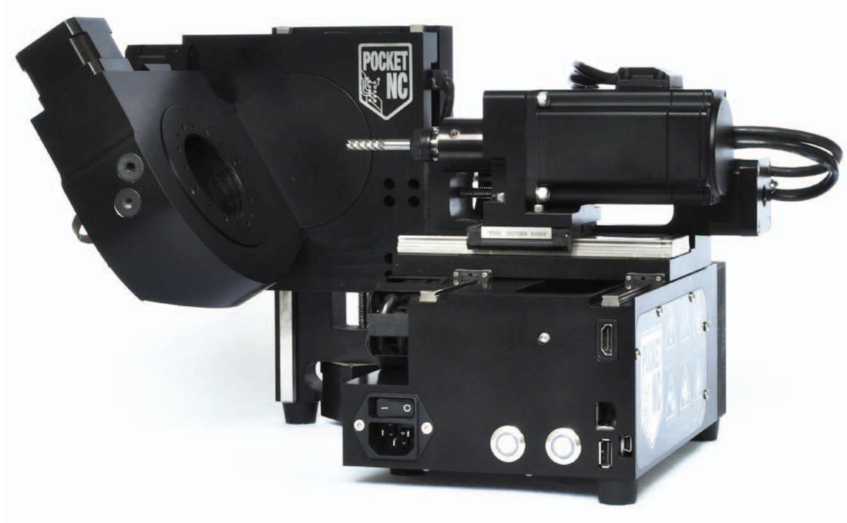


Figure 1. Pocket NC version 2. This desktop computer numerical control (CNC) tool can produce parts from aluminum, wood, plastic, and other materials.

CNC machines use G-code, a series of commands that move the carving bit through 3D space via a series of tool paths. Until recently, generating G-code was difficult because highly specialized software was necessary to generate the tool paths necessary to produce 3D objects. However, Autodesk has introduced Fusion 360, a 3D computer-aided design, manufacturing, and engineering tool that requires minimal custom coding to generate working G-code. Also, with Fusion 360, a CNC machine profile can simulate tool paths to validate G-code before testing on a CNC machine.

People who want to make complex parts with fewer steps need a 5-axis CNC machine. In the past, though, if you needed to ask the price of such a machine, you probably couldn't afford it. In 2015, however, Pocket NC started a Kickstarter campaign to sell a desktop 5-axis CNC machine for $5,000 that could produce parts from aluminum, wood, plastic, and other materials. It also works with Fusion 360. Pocket NC had $355,000 in pre-sales, exceeding the company's goal of $70,000 to start the project. We recently purchased the Pocket NC version 2, shown in Figure 1, which now sits on our back workbench waiting to make its first part.

LOW-COST CNC

A recent Kickstarter project that caught my attention was the Maslow CNC, whose goal was to develop a $500 4' x 8' computer-controlled router table that could cut intricate designs in full sheets of plywood (www.kickstarter.com/projects/1830738289/maslow-cnc-a-500-open-source-4-by-8-foot-cnc-machi). The device, shown in Figure 2, uses two motors in the upper left and right corners of a vertical stand attached to a wood router via two chains. As the motors rotate along the length of the two chains, they move the router to different parts of the plywood sheet. Using an Arduino microcontroller running G-code, you can cut precise shapes.

The project was a hit and resulted in pre-sales of $314,547. For Maslow CNC's version 2, the maker community has provided code changes and an innovative router-sled design that helps

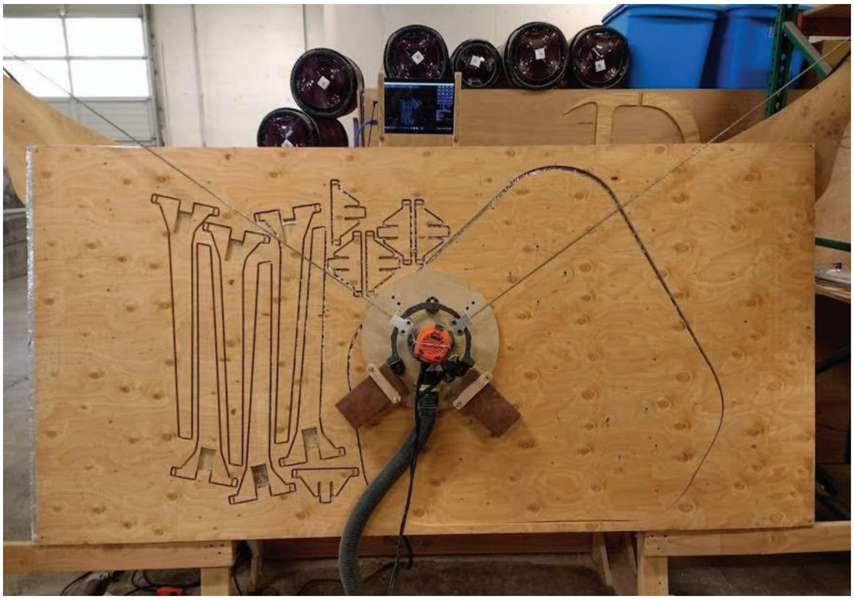


Figure 2. Maslow CNC. This computer numerical control (CNC) device is a low-cost, large-format product for cutting intricate designs out of plywood sheets.

Figure 3. Shaper Origin. This innovative power tool cuts pieces of plywood into various shapes. It uses an integrated camera to track its position on the wood via pieces of tape with dots on them. This helps the operator cut along the desired path. This tool dramatically reduces the craftsmanship required for users to make precision cuts.

compensate for positioning errors when making cuts.

This is a good example of what makers can do to lower costs and create new platforms for the devices they use.

POWER TOOLS' FUTURE

People who really enjoy simple but clever solutions to a problem will love Shaper Origin (shapertools.com), shown in Figure 3. This is a new category of high-tech power tool that works with a traditional router design but that also has an integrated computer screen and video camera. This dramatically reduces the craftsmanship required for a precision cut.

The first step is to apply multiple strips of tape with dot patterns to your sheet of plywood. The camera integrated into the router uses the tape to track its position on the board. You then download the pattern to be cut, and the integrated computer screen on the router tells the human operator the direction in which to move the router to follow the virtual pattern that must be cut. The tool's adjustment capabilities move the router bit left and right to compensate for virtual tracing errors by the operator.

They tool's demonstrated precision (www.youtube.com/watch?v=DekAjAOIVvQ) justifies the $2,399 purchase price.

DESKTOP STEEL-CUTTING WATER-JET FOR THE OFFICE

My most anticipated future purchase is a $4,995 desktop water-jet cutter called Wazer, shown in Figure 4, that enables precise cuts out of almost any material, including steel, aluminum, titanium, carbon fiber, and ceramics. Anticipated delivery is December 2018.

The company's founders are four engineering students from the University of Pennsylvania's School of Engineering and Applied Science, where they met in 2012 and saw the need for a low-cost water-jet cutter for parts fabrication. Typical commercial water-jet cutters cost at least $100,000 and are too big for a desktop. The Wazer team started building prototypes and in 2015, launched their Kickstarter campaign. The project received $1,331,936 in pre-sales, indicating customer demand for the product.

RAPID-BUILD LEGOS FOR BIG KIDS

RobotZone's Actobotics (www.servocity.com/actobotics) is a precision system that provides users with many types of parts they can use to build numerous mechanical systems with just an Allen wrench.

goBILDA by Modern Robotics Inc. (modernroboticsinc.com/gobilda) is a new metric-based system that takes machine building to the next level. It provides building-block components that users can put together to create robots.

In my after-school community robotics program at TechGarage, Actobotics and goBILDA parts are key components. They let middle- and high-school students build amazing robots without having to drill or cut metal. This is safer and enables high-quality builds.

Figure 4. Wazer. This desktop water-jet device can precisely cut many materials, including steel.

CAD IN THE CLOUD

When using any of the above approaches, we first must implement CAD. This suggests that we should expect innovation in CAD software and modeling. An upstart that is changing CAD is Onshape Inc. Its Onshape system works via the software-as-a-service model and focuses on real-time collaboration, parametric modeling, and the rapid introduction of new features. It uses WebGL—a JavaScript API for producing interactive graphics in a browser without plug-ins—for CAD rendering. Onshape is free for users who keep their models and designs public, and costs $125 per month for a professional version.

The maker economy is in full swing, and there are exciting times ahead for people who have a product idea and need to develop a prototype that can be used in a Kickstarter campaign to measure customer interest. With a local makerspace filled with the latest and greatest in rapid prototyping equipment, the maker revolution will help drive an age of rapid innovation. ■

SCOOTER WILLIS is a founder of TechGarage, a centralized robotics workspace for middle- and high-school students in Palm Beach County, Florida. Willis is also director of computational biology at the Avera Cancer Institute. Contact him at willishf@gmail.com.

Trolling Pathologies

Hal Berghel, University of Nevada, Las Vegas

The netizen watch phrase is: don't feed the trolls. A better guide would be: learn to recognize the trolls.

The phrase "online trolling" denotes the practice of anonymously interrupting normal and customary information exchange in order to lure the recipient into reacting to the message. It shares this interference function with signal jamming, network blocking, network filtering, and a host of other technologies. The important difference lies in trolling's anonymity and pathological allure. Unlike more primitive interrupts, trolling attempts to either engage or inflame the receiver, usually through misinformation, lies, distortions, and so on. Of course, it weren't for the false, misleading, or inflammatory content, and the desire for anonymity, trolling wouldn't be necessary—ordinary communication would suffice.

ONLINE TROLLING

Trolling can be understood as a virulent form of online interruption. It makes an important and weaponized contribution to the Internet's negative space,[1] the digital repository at which unprepared and impressionable minds feed. In this capacity, trolling is an epistemological companion to fake news, alt-facts, canards, fallacies, prevarications, hyperbole, illogic, BS, exaggerations, and hate speech.

Of the various forms of online trolling, some can be easily dismissed as pedestrian information interferons. These include hit-and-run posting, sh*tposting, meatpuppeting, and other social media snarkiness that serve the subcerebral regions of troll-space. In my view, the most interesting and worthy of study are those that have the greatest potential for social disruption and/or are the most suggestive of underlying personality disorders—especially those that seem to be in co-morbidity with dramatic personality disorders (Cluster B) such as narcissism and histrionics, as described in the latest edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; www.psychiatry.org/psychiatrists/practice/dsm).

It's not surprising that trolling, as a weaponized form of communication, is popular with today's tyrants, dictators, and demagogues. We've recently become acutely aware how comfortably trolling complements such modern political perversions as pathocracy (rule by a maladjusted minority, primarily psychopaths and narcissists)[2]

EDITOR HAL BERGHEL
University of Nevada, Las Vegas; hlb@computer.org

and kakistocracy (rule by the least competent).[3] Of course, Aldous Huxley and George Orwell spoke forcefully on this phenomenon in the past century. Trolling adds nothing conceptual to villainy, just a new digital veneer.

For classification purposes, I offer the following (nonexhaustive) list of definitions of online trolling.

- *Provocation trolling*: to elicit a particular response (for example, hostility) from one or more participants of an online forum
- *Social engineering trolling*: to incite participants to activities that they would normally not have undertaken (closely related to provocation trolling)
- *Partisan trolling*: to de-legitimize a political opponent or opposing party for political advantage
- *Ad hominem trolling*: to harass, defame, or de-legitimize individuals or groups
- *Nuisance trolling*: to derail the thread of an online forum (blog, chatroom, and so on)
- *Snag trolling*: to evoke responses to satisfy curiosity
- *Jam trolling*: to disrupt a communication channel (the trolling equivalent to a denial-of-service attack)
- *Sport trolling*: to gratify the troll (just for the fun of it)

Of course, all forms of trolling should be taken seriously, for the consequences can be severe and disruptive to the prevailing social order or the civility of the participants to one another. Online trolling can manipulate public opinion precisely because it so easily escapes detection by the uninitiated. But not all trolling is equal in terms of effect—especially in politics. Trolling frequently says something particularly revealing about the psychology of the troll.

WHEN TROLLING BECOMES PATHOLOGICAL

"Don't feed the trolls" is a well-worn cliché for netizens. And there's certainly something to be said for this: responding to trolls encourages further trolling much like answering robocalls tells marketers that your phone line is active. Nicole Sullivan[4] argues that trolls tend to be people who seek conflict. Thus, feeding trolls only energizes them. Yet this analysis is incomplete.

It seems to me that pathological trolls aren't driven by the need for attention at all. Rather, they seek to exercise power over others. In other words, they're first and foremost control freaks. And since the online world allows them to hide behind the cloak of anonymity, they're malignant control freaks. Let's look at this from the point of view of what I'll call stimulus-response trolling theory, which builds on research by Claire Hardaker.[5]

We could develop a suite of trolling methodologies based on networking principles—rectified trolling, bi-directional trolling, relayed trolling, and so on—but we'll limit ourselves here to the basic types of exchanges, in which the tilde indicates that no trolling was involved in that component of the exchange:

1. ~TROLL/~TROLL: ordinary exchange
2. ~TROLL/TROLL: indicating a troll/trollop that injects him/herself into an otherwise ordinary exchange
3. TROLL/~TROLL: sucker/victim—the troll "catches" the unwary victim off-guard and manipulates to effect
4. TROLL/TROLL: troll warfare; provocative in its potential deviance, this has yet to mature

Case 3 predominates in the trolling landscape of the Internet's negative space and is the primary subject of Sullivan and Hardaker's observations. In fact, as it underlies much of the misinformation promulgated during the 2016 US presidential campaign and accounts for most categories of trolling described above, it fits well within the "discipline" of disinformatics.[6] Case 3 is also becoming a cornerstone of Machiavellian manipulation of public opinion in the digital age.[7,8] Case 4 has yet to fully emerge, but the prospect of future political troll wars seems likely.

Online trolling can manipulate public opinion precisely because it so easily escapes detection by the uninitiated.

TWO CASES OF AD HOMINEM TROLLING

Having outlined the underlying "logic" of online trolling, I offer a couple of noteworthy exemplars.

iPad Steering Wheel Mount

The first example illustrates a snag (or possibly social engineering) trolling stimulus and ad hominem trolling responses. It emerged during an online discussion of a fictional iPad Steering Wheel Mount (www.youtube.com/watch?v=AdAkKKxOvu4). This fantastic piece of trolling was engineered by my colleague (and former fellow *Computer* editor) Chuck Severance as part of a social media experiment to investigate how viral videos propagate. Along with the YouTube video, Chuck also provides a fictitious website (www.steering-wheel-ipad.com/index.html) replete with faux user testimonials and bogus ordering instructions.

The iPSWM purports to be safer and less distracting than other hands-free devices by attaching directly to the steering wheel. The idea alone is priceless, but the critical responses to the video are particularly noteworthy:

- Dana: "'you shouldn't be reading on a turn anyway' lol, OR BEHIND THE WHEEL AT ALL!"
- phoulmouth: "was waiting, no, hoping, that this moron crashed."
- Samuel: "Funniest thing ever … If this isn't a joke bro you got ambition. Lol"
- Miorakira: "You're an idiot. Plain and simple."
- Aerolift: "I can't wait tto see the airbag popping :D ipad inside your brain"
- Gerry: "Really stupid. What else ya got?"
- TheFroman007: "This is single handedly the dumbest f*cking idea ever. You cant look at anything on the dash now like the speedometer, tachometer, or even if you have gas in your car. And how about when that airbag goes off, have fun having a smashed ipad in your f*cking face. Why the f*ck do you need to read while driving. Do that when you stop."
- Alex: "This is one of the most idiotic and stupid things to do.. DONT DRIVE AND USE THIS OBJECT,!.,!please,,,"
- Danielkaas94: "Only a true american would 'invent' this kind of garbage you're an id[i]ot case closed."
- GanceDavin: "You must be a special kind of stupid"
- Kevlandy: "What a retarded invention."

To be sure, some respondents picked up on the gag. But all too many did not, and this makes the study fascinating—and alarming. Of course, the hostility of the responses is transparent. But at a subtler level the respondents betray an intellectual shallowness by not making any effort to understand the stimulus in context. Rather than being thoughtful, the reactions seem to be a gut-level state of hyperarousal or acute stress response by the sympathetic nervous system. If language weren't involved one might be tempted to explain these reactive mechanisms by way of analogy to the primitive behavior of lower primates.

> Rather than being thoughtful, troll reactions seem to be a gut-level state of hyperarousal or acute stress response by the sympathetic nervous system.

As an aside, the trolling phenomenon gets at the heart of many of society's deepest problems, not the least of which is the inability to either anticipate or understand irony. The scariest part of this exchange is that some of these respondents undoubtedly carry weapons and vote.

It's important to note that while Chuck didn't highlight the fact that he's on the faculty of the University of Michigan, he didn't go to any great effort to conceal it either. The iPSWM website contained links to both a personal profile and other university-related activities. The photo of the corporate headquarters ("the Base 441 design and incubation studio in the Lansing, MI designated renaissance zone") appears to be an abandoned warehouse, and the link to the iPSWM patent takes you to the patent for duct tape. Confirmation of the gag was never much more than a mouse click away from the original YouTube video and website, yet so many missed it. The reason for this should be intoxicating to mental health professionals. Trolling pathology must be understood in terms of this willful ignorance.

Right Angle Flex Shaft

My second trolling example is provided by Amazon. In the "Comments" section under an ad for the Dewalt DWARAFS 12" Right Angle Flex Shaft, a drill bit extension for hard-to-reach places, a customer posted the following question: "The description says it is a 'right angle flex shaft' but in the picture it appears to be pointing to the left—is this the right or left version?" (www.amazon.com/ask/questions/Tx2ZFG8IP7LBYNZ/ref=ask_dp_dpmw_al_hza). This predictably energized a platoon of trolls:

- S. Whitaker: "Seriously!?!? Lol. Right angle is a term for 90 degree angle. But I don't think you should be using any form of power tools."
- Lauren: "I've got a left-handed hammer, if you need one."
- Richard: "I had the same problem with a box of nails. I would grab one and it would point to the right wall, not the left wall. I was only able to use about half the box."
- Marty: "I think that would depend on whether you are north or south of the equator. If you are at the north or south pole, it will not work at an angle at all."
- K. Sianecz: "Depends if you live on the left or right side of the Prime Meridian."
- Bruce: "You should go into a circular room, find a corner to sit, and think about it."

- Octoberken: “It’s configurable. If you rotate it 180 degrees along the vertical axis it will point to the right.”
- Brendan: “Right angle refers to the degree of the angle. 90 degrees is a right angle. There’s no left or right version.”

All but the last two are ad hominem responses to what’s pretty clearly a snag or sport trolling stimulus. As with the iPSWM, the initial critical responses show clear evidence of pedestrian herd mentality as the respondents try to outdo one another at the questioner’s expense. This case differs from the iPSWM in only one important respect: this snag troller didn’t leave any clues behind. That said, the ad itself, especially the product photo, would make the gag obvious upon even minimal reflection. But once again, the first-order effect is an ad hominem attack on the source presumably triggered by some primitive primate instinct. Perhaps ad hominem trolling is a linguistic counterpart to a fight-or-flight response.

WHAT’S GOING ON HERE?

Note that in both examples the first-order response to the trolling stimulus is caustic. The reasons are probably best left to social scientists. But some superficial conclusions are obvious. The reaction suggests a variety of unhealthy personality traits that might even qualify as disorders. I’m convinced that whatever the causes might be, they’re likely to be found in DSM-5. As one untrained in psychotherapy, I can only suggest the mundane: the caustic responses suggest myopia, egoism, poor self-control, and a perverse feeling that diminishing others can elevate self-worth—not to mention a shallow intellect, and absence of contextual awareness. If Freud were alive today, what a field day he would have with trolling.

If I’m right, pathological trolling illustrates the worst aspects of human nature. The parallel between malicious trolling and disseminating fake news, lies, and such seems obvious. What’s not obvious are the mental machinations that drive this antisocial behavior. I would speculate that those who engage in malignant online commentary are in some sort of cognitive decline as yet to be defined by social scientists and mental health professionals. The fact that trolls seem to become immediately polarized by such a simple thing as an advertisement or a tweet shouldn’t be overlooked. Their behavior is rife with unhealthy impulses.

In Scandinavian mythology, trolls are mischievous and sometimes dangerous creatures that live in caves or underground. The name—derived from the Old Norse word for “fiend”—is an appropriate descriptor for the legions of anonymous pathological posters who contribute to the polarization of society by propagating misinformation and venom. At this point online trolling is but one stage of a disinformation wash cycle that seems to infect every important aspect of our life. In a future column, I’ll discuss what we might be able to do about it. Meanwhile, remember that the price to avoid being the “victim of trolling” is eternal vigilance.

REFERENCES

1. G. Fuller, C., McCrea, and J. Wilson, eds., “Trolls and the Negative Space of the Internet,” *The Fibreculture J.*, issue 22, 2013; fibreculturejournal.org/wp-content/pdfs/FC22_FullIssue.pdf.
2. E. Mika, “Who Goes Trump?,” B. Lee, ed., *The Dangerous Case of Donald Trump: 27 Psychiatrists and Mental Health Experts Assess a President*, St. Martin’s Press, 2017, pp. 298–318.
3. E.J. Dionne, N.J. Ornstein, and T.E. Mann, *One Nation after Trump: A Guide for the Perplexed, the Disillusioned, the Desperate, and the Not-Yet Deported*, St. Martin’s Press, 2017.
4. N. Sullivan, “Don’t Feed the Trolls,” video, O’Reilly Fluent Conf., 30 May 2012; www.youtube.com/watch?feature=player_embedded&v=ulNSlES1Fds.
5. C. Hardaker, “Trolling in Asynchronous Computer-Mediated Communication: From User Discussions to Academic Definitions,” *J. Politeness Research*, vol. 6, no. 2, 2010, pp. 215–242.
6. H. Berghel, “Disinformatics: The Discipline behind Grand Deceptions,” *Computer*, vol. 51, no. 1, 2018, pp. 89–93.
7. H. Berghel, “Lies, Damn Lies, and Fake News,” *Computer*, vol. 50, no. 2, 2017, pp. 80–85.
8. H. Berghel, “Alt-News and Post-Truths in the ‘Fake News’ Era,” *Computer*, vol. 50, no. 4, 2017, pp. 110–114.

HAL BERGHEL is an IEEE and ACM Fellow and a professor of computer science at the University of Nevada, Las Vegas. Contact him at hlb@computer.org.

The Cyber-Physical Systems Revolution

Dimitrios Serpanos, ISI/ATHENA and University of Patras

Cyber-physical systems constitute a disruptive technology across many industries, with a strong impact on economies and social processes. Their applications in many domains, from manufacturing to agriculture and from critical infrastructure to assistive living, will challenge technology, business, law, and ethics.

Cyber-physical systems (CPS) integrate computational and physical components to implement a process in the real world. Since the beginning of computing, there have been efforts to automate complex processes through the inclusion of computing systems, as demonstrated by the well-known Apollo Guidance Computer in the 1960s, which was used for guidance, navigation, and control. The technological advances of the last decades on fronts such as sensors, instrumentation, networking, and embedded computing have enabled us to develop systems and applications that have changed our daily lives. Automatic building power management, automated insulin delivery pumps, self-braking cars, and surveillance drones are just a few examples of conventional CPS that affect everyday life in diverse application areas.

CPS constitute a disruptive technology, bringing innovation to many industries because of their potential to integrate technologies from various sectors, transform traditional processes in several application areas, and enable new processes. Consumers experience transformations in many of the services they receive through conventional devices, from home automation and energy management to health services and entertainment. Personalized healthcare is one of the first changes, especially as medical innovation increasingly offers personalized treatments to citizens, with the support of sophisticated health devices at home and in health centers and hospitals; examples of CPS health applications are systems that monitor patients continuously and deliver medication on the fly; systems that enable remote monitoring of patients; and systems that support the movement of individuals with disabilities by, for example, sensing and adjusting the actions of artificial limbs.

Energy systems are being transformed at all levels because of CPS. Consumers, individuals, and organizations are already experiencing significant advantages from

 0018-9162/18/$33.00 © 2018 IEEE

EDITOR DIMITRIOS SERPANOS
ISI/ATHENA and University of Patras; serpanos@computer.org

the exploitation of smart systems that manage energy consumption in increasingly smart buildings that sense power needs and enable delivery. Energy producers exploit CPS to identify the need for energy and adjust their plants to deliver the required energy where and when needed, saving significant resources. Electric power distribution operators increasingly employ CPS to monitor and manage their distribution systems in real time to avoid outages and appropriately satisfy their consumers' needs.

The electric grid, already being transformed to a smart grid, constitutes part of the critical infrastructure of many countries, and is managed and controlled by CPS. Its continuous and reliable operation is necessary for these countries' economies as well as for the well-being of their citizens. Its service disruptions can lead to significant problems in everyday life, influencing aspects such as safety, productivity, and health.

Water management, transportation, manufacturing, and communications also constitute significant components of critical infrastructure and of emerging smart cities. Smart traffic management, autonomous vehicles, automated water leakage detection and prevention systems, and self-healing networks are transforming critical infrastructures' management and effectiveness. Several innovations are changing traditional processes in disruptive ways, for instance, allowing the personalization of mass-produced products in manufacturing—considered by many to be the fourth industrial revolution. Smart agriculture, enabled by CPS management of agricultural production, is another example of a CPS application with a strong, widescale impact. CPS are already having a visible impact on many fronts, including robotics, security, safety, and military.

CPS's progress has enabled the emergence and fast growth of the Internet of Things (IoT). Exploiting Internet technology to transfer data over heterogeneous communication technologies, both wired and wireless, has provided opportunities to combine systems across distances in distributed applications and processes, collect data remotely for process analysis and optimization, and manage systems remotely and reliably at a low cost. IoT-connected CPS as "things," combined with data analytics and artificial intelligence, boost development of a wide range of smart systems, environments, applications, and services, from smart cities and governance to smart manufacturing and control of critical infrastructure.

CPS constitute a disruptive technology, bringing innovation to many industries because of their potential to integrate technologies from various sectors,

CHALLENGES

CPS differ from classical computational and information systems in many aspects. Their structure includes a wide range of heterogeneous technologies for computational components, networks, sensors, and actuators. Their interaction with the physical world raises issues about their robustness and resilience as well as about liability when CPS actions violate expectations or even regulations in applications, for example, in health or transport systems. Furthermore, their adoption for automating processes in domains ranging from manufacturing to agriculture is expected to change business models by taking humans out of the loop in many services, which will have major societal consequences. Thus, CPS's projected impact across industries and applications will create significant technical, legal, and societal challenges.

Considering the applications of CPS and their related technologies—computational, network, and instrumentation—we must address these challenges. CPS must be architected, designed, and implemented to be easily extendable and scalable. As systems are becoming increasingly interconnected, CPS must extend their functionality to integrate heterogeneous communication technologies and become part of larger systems that implement even more sophisticated processes. The design, operation, and management of such complex systems requires the ability to describe them at various levels of abstraction, enabling analyses and decisions for technical, operational, management, and business purposes. The need for scalability, modularity, and composability leads to significant technical issues in the area of systems-of-systems, where correctness and system verification are challenging problems, especially in light of applications with timing constraints and continuous operation requirements. An important challenge in this direction is the integration of traditional IT systems with operational technology (OT) systems. OT systems differ from traditional IT systems in many ways, from purpose, computing components, communication technologies, and interfaces to ownership and management. The integration of these technologies poses several challenges, especially when considering legacy OT systems; the challenges include software integration and upgrade, network interoperability, synchronization in light of real-time processes and applications, and, importantly, security.

Security is one of the major challenges in CPS, as several events have already demonstrated. Incidents such as the Aurora experiment that destroyed a power generator by cyber means in a controlled environment, the Stuxnet cyber-weapon against the Iranian nuclear facilities, and the 2016 Mirai distributed denial-of-service attack not only demonstrate the lack of security in CPS but also damage the trust of CPS users, organizations, and consumers.

The security challenges go beyond computations and communication in CPS. They also include reliable interaction with the physical world, especially when considering that the physical environment might "behave" in unexpected and unpredictable ways; independently, CPS have to adjust and operate reliably in a well-defined and predictable fashion with safety as a major priority, especially in environments where hazardous conditions might occur, such as power systems, autonomous vehicles, and health systems. Security is a prerequisite for safety because security breaches might lead to safety hazards—although not necessarily.

Another important aspect of security relates to the privacy of users. Security attacks on home devices, for example, can lead to sensitive information leakage. Moreover, the expected collection of (even anonymous) data from CPS operation over long time intervals might disclose behavioral patterns of users that enable their identification and characterization beyond the scope of applications, violating privacy rights.

Reliability and dependability are a necessity in CPS, especially in environments where continuous operation is required, as in the domains of critical infrastructure, manufacturing, health, autonomous vehicles, and so on. There is a need to address these properties combined with security, because malicious attacks often attempt to disrupt the continuous operation of the systems. Attacks on power systems, such as the recent attack on the Ukranian power grid, have demonstrated the urgency of addressing these aspects in a unified way.

> CPS have to adjust and operate reliably in a well-defined and predictable fashion with safety as a major priority, especially in environments where hazardous conditions might occur.

The integration of CPS across domains raises the challenge of interoperability. Data interoperability has been already identified as a major priority, because data have to be interpreted consistently within processes at all levels of the overall distributed CPS. However, data interoperability is only one aspect of the interoperability challenge, which needs to address computational aspects in addition to data. Composing complex CPS from components that already implement simpler processes requires consistent interpretation of the context and the semantics of operations, commands, and subprocesses, setting the foundation for machine interpretation, inferencing, logic, and knowledge discovery in order to enable seamless integration and compose a correct and resilient overall process.

The distributed structure of CPS also raises timing challenges. An integral part of many CPS applications is meeting real-time requirements. Many of the manufacturing processes, for example, require periodic actions and mandate time criticality; many monitoring systems detect conditions that need a reaction within a specific time interval. In many cases, violation of a deadline leads to unsafe states and results in undesirable conditions such as accidents. Combining systems with different clocks that operate at different granularities, and have different real-time requirements in terms of scale and priorities is a well-known challenge in computing and requires innovative methods to enable a unified timing scheme for the overall system.

Employment of CPS and their effect on emerging applications raises legal and societal challenges. Clearly, privacy concerns lead to legal actions and challenge lawmakers to establish legal frameworks for the appropriate storage, processing and use of sensitive data. Although there are significant efforts in progress, such as the General Data Protection Regulation (GDPR) in the EU, the unanticipated usage of collected data as well as the unexpected breaches of privacy and leakage of sensitive information pose a continuous challenge to evolve legal frameworks that protect consumers and organizations effectively.

Digital rights management is another aspect of CPS applications that has significant legal ramifications. Although the issue has been targeted for more than a decade at the online entertainment business, the pervasiveness of CPS makes the problem more acute, covering issues beyond media, including designs, software, and licenses on a large scale.

Additionally, despite efforts toward resilient CPS, unexpected circumstances are bound to occur, leading to violations of safety and accidents. Liability is a major issue in such cases—are human operators, manufacturers, or maintainers liable? What procedures are in place to investigate sources of problems and place liability appropriately? Differences in legal systems across the globe, in conjunction with emerging distributed CPS processes across borders, present unique legal challenges.

The adoption of CPS to automate processes is increasingly taking humans out of control loops, but creating new opportunities in the maintenance and management of the control loops. Thus, policies are needed that will help citizens transition effectively into these new employment roles. At the same time, ethical questions are raised about limits: where, how, and when automated processes should be replacing human operation and judgement. This relates not only to processes far from immediate human oversight but also to systems immediately used by people for their health and well-being, such as assistive living and artificial prosthetic parts.

INITIATIVES

National initiatives and large consortia efforts have emerged to develop frameworks that will enable fast and effective growth addressing the emerging challenges. Most notable among them are the Industrie 4.0 initiative in Germany, the Industrial Internet in the US, and the Society 5.0 initiative in Japan.

Industrie 4.0 is a strategic initiative to enable Germany to maintain a leading role in manufacturing.[1] The goal is efficient and low-cost production with flexible workflows. This will be achieved through widespread use of CPS in manufacturing and production processes, inserting intelligence in systems and processes, and coordinating them into more complex but flexible processes that produce high-quality products at low cost. The smart factory concept, which achieves the initiative's goals, employs a hierarchy of CPS with smart machines interconnected to establish smart plants, which are then combined to establish smart factories. The concept targets flexibility, autonomy, resilience, safety, efficiency, and low cost, enabling all actors—customers, operators, and manufacturers—to monitor the system parameters of their interest.

The Industrial Internet of Things (IIoT) focuses on the application of IoT to the industrial sector, including not only the industrial production process but other processes such as asset management and maintenance. IIoT addresses efficient and effective operations and interoperability, taking into account emerging and future services as well as the stakeholders involved in the related devices, CPS, communications, service provision, and business development. The need for standards and reference architectures is addressed by several efforts, most notably the International Telecommunication Union (ITU) ITU-T Y.2060 recommendation[2] and the Industrial Internet Consortium's (IIC's) Industrial Internet Reference Architecture.[3] The ITU recommendation addresses IoT, in general, but includes applications for IIoT such as a smart grid.

Importantly, the ITU recommendation also presents business models for IoT. The IIC is a detailed reference architecture that can be considered an elaboration of the ITU reference architecture addressing important aspects for all categories of stakeholders. Importantly, NIST has been developing a CPS framework to enable fast development and deployment of CPS in the nation's economy.[4]

Society 5.0, also called Super Smart Society, is a Japanese initiative to integrate CPS with IoT, big data technologies, and AI into every industry and all aspects of society to address societal challenges.[5] High-priority applications are in healthcare, transportation and mobility, infrastructure maintenance, and the financial sector.

CPS are revolutionizing economies and social processes. Their effective development, adoption, and use require breakthroughs on several fronts, including policy, law, business, and social sciences as well as technology. ■

REFERENCES

1. "Industrie 4.0 Smart Manufacturing for the Future," Germany Trade and Invest, July 2014; www.gtai.de/GTAI/Navigation/EN/Invest/Service/Publications/business-information,t=industrie-40--smart-manufacturing-for-the-future,did=917080.html.
2. "Series Y: Global Information Infrastructure, Internet Protocol Aspects, and Next-Generation Networks. Overview of the Internet of Things: Recommendation Y.2060/2012," Int'l Telecommunication Union, June 2012; www.itu.int/rec/T-REC-Y.2060-201206-I.
3. S.-W. Lin, M. Crawford, and S. Mellor, eds., *The Industrial Internet of Things Volume G1: Reference Architecture, Industrial Internet Consortium*, IIC:PUB:G1:V1.80:20170131, 2017; www.iiconsortium.org/IIC_PUB_G1_V1.80_2017-01-31.pdf.
4. *Framework for Cyber-Physical Systems Release 1.0*, NIST, May 2016; s3.amazonaws.com/nist-sgcps/cpspwg/files/pwgglobal/CPS_PWG_Framework_for_Cyber_Physical_Systems_Release_1_0Final.pdf.
5. *Toward Realization of the New Economy and Society*, Keidanren, 2017; www.keidanren.or.jp/en/policy/2016/029_outline.pdf.

DIMITRIOS SERPANOS is director of the Industrial Systems Institute/ATHENA and a professor of electrical and computer engineering at the University of Patras. He is a Senior Member of IEEE and a member of ACM, AAAS, and NYAS. Contact him at serpanos@computer.org.

Hacking's Brand-Equity Nexus

Nir Kshetri, University of North Carolina at Greensboro

Jeffrey Voas, IEEE Fellow

Cyberattacks can significantly affect brand reputation, but companies can take measures to repair the damage.

Companies experiencing cyberattacks are likely to suffer a loss in reputation. According to a January 2012 Ponemon Institute study, the average decline in brand equity from a data breach ranged from about $184 million to $330 million, or 17 to 31 percent of the brand's value.[1] A recent survey of retailers found that 55 percent considered reputation protection to be their top priority for IT security spending. According to the survey, 89 percent of retailers felt vulnerable to data threats and 51 percent had already experienced a data breach.[2]

LINKING CYBERSECURITY TO BRAND EQUITY

Brand equity denotes consumer perception of a product or service's value or attractiveness. Brand equity has two key determinants, perceived quality and brand loyalty,[3] and surveys indicate that consumers have a more negative perception and a less positive perception of companies that have been hacked, especially if consumer data has been compromised.

Products containing insecure software are perceived as being of lower quality, and there's growing awareness of the need to build cybersecurity into modern devices and systems.[4] For example, a recent *Wired* article chronicled the successful hack of a 2015 Jeep Cherokee by two researchers, who remotely changed the climate control system setting, turned on the radio and windshield wipers, and stopped the accelerator. "Automakers need to be held accountable for their vehicles' digital security," the author asserted. The lead researcher noted that "If consumers don't realize this is an issue, they should, and they should start complaining to carmakers. This might be the kind of software bug most likely to kill someone."[5]

Brand loyalty indicates consumers' tendency to repeatedly choose one brand over alternatives. Firms experiencing cybersecurity breaches will likely observe a decrease in brand loyalty. In a 2014 survey conducted for

0018-9162/18/$33.00 © 2018 IEEE

TABLE 1. Cyberattacks on companies and impact on brand equity.

Company	Cyberattack	Brand impact
Sony PlayStation Network	In 2011, hackers extracted personally identifiable information of more than 77 million users.	The YouGov BrandIndex Buzz score fell from 5 the week before to −14 one day after announcement of the breach.
Target	A security breach during the peak holiday shopping season of 27 November to 15 December 2013 compromised 40 million credit and debit card accounts and 70 million customers' personal data.	The Buzz score dropped from 26 the week before to −9 one day after announcement of the breach and another 10 points within a few days, for a total decline of 45 points. Target's sales fell by 46 percent in the fourth quarter of 2013.
Sony Pictures Entertainment	In 2014, hackers downloaded 100 Tbytes of data including unreleased films and TV shows and leaked more than 47,000 Social Security numbers (SSNs), details on salary negotiations, and other sensitive information.	The Buzz index dropped about 14 points after announcement of the breach. The company also fell from third to eighth place in the Brand Keys Customer Loyalty Engagement Index.
Anthem Blue Cross and Blue Shield	In early 2015, 80 million records were breached including customers' and employees' SSNs, birthdays, addresses, emails, employment information, and income data. The chief executive's information was also compromised.	The proportion of consumers who thought that Anthem Blue Cross and Blue Shield was a better brand than other insurers decreased from 51 percent before the breach to 45 percent afterward.
TalkTalk	In October 2015, hackers accessed details of 150,000 customers and 15,000 bank accounts.	In a two-week period following the hack, the Buzz score plummeted from −1 to −50. The company's share of new customers in the broadband home services market fell by 4.4 percent in the last quarter of 2015, and 7 percent of customers switched to a different provider.
Yahoo	In September 2016, Yahoo reported that hackers might have stolen information on at least 500 million users in 2014.	A day after the breach announcement, online mentions of the company increased 474 percent, with 70 percent of the mentions negative.

CreditCards.com by Princeton Survey Research Associates International, 45 percent of respondents with credit or debit cards indicated that they would "definitely or probably avoid" retailers that experienced hacks. Among households earning $75,000 or more annually, 31 percent would avoid such retailers compared to 56 percent of those earning less than $30,000.[6] Likewise, following the October 2016 botnet attack on DNS provider Dyn, more than 14,000 Internet domains stopped using its services.[7]

SUFFERING BRAND-IMAGE EROSION DUE TO ATTACKS

Table 1 lists some examples of businesses that experienced significant negative branding effects following cyberattacks. One widely used measurement of brand popularity is the "Buzz" score developed by the polling site YouGov BrandIndex (www.brandindex.com). A company's score, which ranges from 100 to −100, is calculated by subtracting negative feedback from positive feedback. As the table shows, companies victimized by high-profile cyberattacks in recent years including the Sony PlayStation Network,[8,9] Target,[9,10] Sony Pictures Entertainment,[11,12] Anthem,[13,14] TalkTalk,[15-17] and Yahoo[18,19] have suffered major declines in their Buzz scores as well as in other measures of brand equity and popularity.

Firms hit by cyberattacks usually also suffer a decline in brand loyalty. Three months after a major 2015 hack, TalkTalk lost 101,000 customers, 95,000 of them as a direct result of the cyberattack.[17] Consumers might also question the reliability of services provided by a firm that has lost sensitive information. About 20 percent of those who left TalkTalk cited "poor reliability" as the reason compared to less than 1 percent who had left in the previous quarter.[16]

POST-BREACH RESILIENCE AND REPARATION OF BROKEN TRUST

Recent studies have shown that reputation damage from a data breach can be irreparable. A KPMG survey of 448 consumers found that 19 percent would stop shopping with a retailer that had been a cyberattack victim even if the company took measures to fix the security issues.[20]

Further, companies that experience hacks and fail to disclose details are viewed as dishonest, suspicious, and untrustworthy and find it difficult to repair broken trust. TalkTalk, for example, was taken to task for keeping the details of its cyberattack secret for a week.[21] Yahoo faced similar criticism for not coming clean about the extent of its own massive data breach. Senator Richard Blumenthal noted: “As law enforcement and regulators examine this incident, they should investigate whether Yahoo may have concealed its knowledge of this breach in order to artificially bolster its valuation in its pending acquisition by Verizon.”[22]

It’s possible for a company to repair—at least partially—broken trust from consumers after a cyberattack.[23] The success of such an effort depends on both the speed and nature of the response.[24]

It’s especially important to have a detailed, well-developed plan in place for responding to cyberattacks, continual security training and education for employees, a designated spokesperson in the event of a breach, and prepared language for distribution to customers, social media, and news outlets. The company website and social media feeds should also have built-in emergency messaging capabilities.[25] Transparency is also critical—the company must honestly explain what went wrong. Finally, the company should consider offering discounts or complimentary services to customers, such as free credit monitoring.

The damage to a company’s reputation following a cyberattack isn’t just a function of the sensitivity of the compromised data and extent of the breach; it’s also based on consumers’ perception of the organization’s cybersecurity practices (or lack thereof) and the nature of the postbreach response. Firms that have been hacked can minimize the adverse effects on brand equity and repair trust with customers, but only if they respond quickly and aggressively.

REFERENCES

1. “How Data Breaches Harm Reputations,” blog, 17 Jan. 2012;www.experian.com/blogs/data-breach/2012/01/17/how-data-breaches-harm-reputations.
2. J. Goldman, “Acer Hacked,” eSecurity Planet, 20 June 2016; www.esecurityplanet.com/hackers/acer-hacked.html.
3. D.A. Aaker, *Managing Brand Equity*, The Free Press, 1991.
4. “Cybersecurity in the Age of Digital Transformation,” *MIT Technology Rev.*, 23 Jan. 2017; www.technologyreview.com/s/603426/cybersecurity-in-the-age-of-digital-transformation/?ref=rss.
5. A. Greenberg, “Hackers Remotely Kill a Jeep on the Highway—With Me in It,” *Wired*, 21 July 2015; www.wired.com/2015/07/hackers-remotely-kill-jeep-highway.
6. K.H. Queen, “Poll: Many Cardholders Will Avoid Stores Hit by Data Breaches,” CreditCards.com, 19 Oct. 2014; www.creditcards.com/credit-card-news/shopping-after-breach.php.
7. “Exclusive: Mirai Attack Was Costly for Dyn, Data Suggests,” The Security Ledger, 3 Feb. 2017; securityledger.com/2017/02/mirai-attack-was-costly-for-dyn-data-suggests.
8. “PlayStation Network Hackers Access Data of 77 Million Users,” *The Guardian*, 26 Apr. 2011; www.theguardian.com/technology/2011/apr/26/playstation-network-hackers-data.
9. T. Marzilli, “Target Perception Falls after Data Breach,” YouGov BrandIndex, 23 Dec. 2013; www.brandindex.com/article/target-perception-plummets-after-data-breach.
10. E.A. Harris and N. Perlroth, “For Target, the Breach Numbers Grow,” *The New York Times*, 10 Jan. 2014; www.nytimes.com/2014/01/11/business/target-breach-affected-70-million-customers.html
11. A. Picchi, “How Big of a Hit to Sony’s Brand?,” CBS MoneyWatch, 19 Dec. 2014; www.cbsnews.com/news/can-sonys-brand-recover-from-the-interview-fiasco.
12. C. Atkinson, “Sony Hack Hurts More Than Stars’ Egos—Brand Loyalty Takes a Hit,” *New York Post*, 15 Dec. 2014; nypost.com/2014/12/15/sony-hack-hurts-more-than-stars-egos-brand-loyalty-takes-a-hit.
13. R. Abelson and M. Goldstein, “Anthem Hacking Points to Security Vulnerability of Health Care Industry,” *The New York Times*, 5 Feb. 2015; www.nytimes.com/2015/02/06/business/experts-suspect-lax-security-left-anthem-vulnerable-to-hackers.html?_r=0.
14. J.K. Wall, “Anthem’s Brand Suffers Small Ding from Data Breach,” *Indianapolis Business J.*, 11 May 2015; www.ibj.com/blogs/12-the-dose-jk-wall/post/53151-anthems-brand-suffers-small-ding-from-data-breach.
15. R. Spillett, “TalkTalk Reveals Personal Details of 150,000 Customers and 15,000 Bank Accounts Were Accessed in Last Month’s Cyber Attack,” *Daily Mail*, 6 Nov. 2015; www.dailymail.co.uk/news/article-3306600/Hackers-accessed-details-150-000-TalkTalk-customers.html.
16. A. Sword, “Rebuilding Brand Trust: TalkTalk’s Path Back from Cyber Attack,” *Computer Business Rev.*, 22 Jan. 2016; www.cbronline.com/news/cybersecurity/data/rebuilding-brand-trust-talktalks-path-back-from-cyber-attack-4790671.
17. J. Lewin, “Cyber Attack Cost TalkTalk up to £60M and 101K Customers,” *Financial Times*, 2 Feb. 2016; www.ft.com/content/410f5477-d061-3cd4-a064-5563c91bd7fb.
18. J. Peters, “Yahoo and Others Face Cybercrime-Related Brand Damage,” blog, 2 Nov. 2016, blog.surfwatchlabs.com/2016/11/02/yahoo-and-others-face-cybercrime-related-brand-damage.
19. T. Dua, “Yahoo’s Data Breach Further Dents Its Already Flailing Brand,” Digiday, 23 Sept. 2016; digiday.com/brands/yahoos-data-breach-dents-already-flailing-brand.
20. “Cyber Attacks Could Cost Retailers One-Fifth of Their Shoppers: KPMG Study,” press release, KPMG, 23 Aug.

2016; home.kpmg.com/us/en/home/media/press-releases/2016/08/cyber-attacks-could-cost-retailers-one-fifth-of-their-shoppers-kpmg-study.html.
21. J. Staufenberg, “TalkTalk Cyber Attack: Call Centre Workers Arrested in India as Part of Hacking Probe,” *Independent*, 28 Jan. 2016; www.independent.co.uk/life-style/gadgets-and-tech/news/talktalk-call-centre-workers-arrested-in-india-in-cyber-hacking-probe-a6838381.html.
22. D. Volz, “Yahoo Faces Growing Scrutiny over When It Learned of Data Breach,” Reuters, 24 Sept. 2016; news.trust.org/item/20160923211729-stk2c.
23. W.P. Bottom et al., “When Talk Is Not Cheap: Substantive Penance and Expressions of Intent in Rebuilding Cooperation,” *Organization Science*, vol. 13, no. 5, 2002, pp. 497–513.
24. M.A. Korsgaard, S.E. Brodt, and E.M. Whitener, “Trust in the Face of Conflict: The Role of Managerial Trustworthy Behavior and Organizational Context,” *J. Applied Psychology*, vol. 87, no. 2, 2002, pp. 312–319.
25. K. Tynan, “Good Marketing Can Help Banks Survive a Cybersecurity Backlash,” *American Banker*, 12 Jan. 2015; www.americanbanker.com/opinion/good-marketing-can-help-banks-survive-a-cybersecurity-backlash.

NIR KSHETRI is a professor of management in the Bryan School of Business and Economics at the University of North Carolina at Greensboro. Contact him at nbkshetr@uncg.edu.

JEFFREY VOAS is a cofounder of Cigital, *Computer*’s Cybertrust column editor, and an IEEE Fellow. Contact him at j.voas@ieee.org.

See **www.computer.org/computer-multimedia** for multimedia content related to this article.

CS CONNECTION

TRAVIS DOOM RECEIVES 2018 COMPUTER SCIENCE & ENGINEERING UNDERGRADUATE TEACHING AWARD

Travis Doom, professor and associate chair in the Department of Computer Science and Engineering at Wright State University (WSU), was named the recipient of the 2018 Computer Science and Engineering Undergraduate Teaching Award.

Doom is recognized by the IEEE Computer Society (CS) "for tremendous contributions to computer science and engineering through a balanced career incorporating research, mentoring of students, curriculum and pedagogy development, and service."

Serving as WSU faculty president, Doom is a co-director of the university's bioinformatics research group and pursues research in the fields of undergraduate engineering education, data science, bioinformatics, and digital/computer systems. His education research focuses on the impact of active learning and culturally relevant teaching on student success in STEM.

Doom says he wouldn't mind being remembered as someone who spent "too much effort" on teaching. He was awarded the position of WSU's Robert J. Kegerreis Distinguished Professor of Teaching from 2014 to 2017. He also received the Outstanding Engineers and Scientists Award (Education) by the Affiliate Societies Council of Dayton in 2015 and the Southern Ohio Council for Higher Educations (SOCHE) Award for excellence in teaching (2014). Doom served on the Faculty Advisory Board for WSU's Center for Teaching and Learning from 2014 to 2016.

A WSU faculty member since 1998, Doom was the recipient of the WSU College of Engineering Excellence in Teaching award in 2000 and 2005, and he was an invitee to the Ohio Teacher's Excellence Program (OTEP) in 2001, and a sponsored attendee of the National Effective Teaching Institute (NETI) in 2002. Doom's work has been included in publications and materials from the the American Society for Engineering Education (ASEE), the ACM special interest group on computer science education (SIGCSE), and *IEEE Transactions on Education*. His teaching interests include both undergraduate and graduate computer science, computer engineering, and electrical engineering courses.

Doom received a PhD from Michigan State University.

The Computer Science & Engineering Undergraduate Teaching Award was created to recognize and honor outstanding contributions to undergraduate education through both teaching and service. The award consists of a plaque, certificate, and a $2,000 honorarium.

Further information about the award, including a list of past recipients, may be found at: www.computer.org/web/awards/cse-undergrad-teaching.

BARBARA LISKOV RECEIVES 2018 IEEE CS COMPUTER PIONEER AWARD

Barbara Liskov, institute professor at MIT, will receive the 2018 IEEE CS Computer Pioneer Award. Liskov is being recognized "for pioneering data abstraction, polymorphism, and support for fault tolerance and distributed computing in the programming languages CLU and Argus."

The award—which will recognize two individuals in 2018—is given for significant contributions to early concepts and developments in the electronic computer field, which have clearly advanced the state-of-the-art in computing.

Liskov's research interests include distributed and parallel systems, programming methodology, and programming languages. She is a member of the National Academy of Engineering, the National Academy of Sciences, and National Inventors Hall of Fame. She is also a Fellow of the American Academy of Arts and Sciences, ACM, and a charter Fellow of the National Academy of Inventors.

In 2003, Liskov was named one of the 50 most important women in science by *Discover* magazine. She was inducted into the National Inventors Hall of Fame in 2012.

She received the ACM Turing Award in 2008, the ACM SIGPLAN Programming Language Achievement Award in 2008, the IEEE John von Neumann Medal in 2004, and a lifetime achievement award from the Society of Women Engineers in 1996.

Liskov received her MS and PhD from Stanford University, becoming one of the first women to earn a doctorate in computer science in the United States.

The Computer Pioneer Award was established in 1981 by the IEEE CS Board of Governors to recognize and honor the vision of those whose efforts resulted in the creation and continued vitality of the computer industry. The award is presented to outstanding individuals whose main contribution to the concepts and development of the computer field was made at least 15 years earlier.

The award will be presented at the 2019 Computer Software and Applications Conference (COMPSAC) held in conjunction with SEMICON West in San Francisco, 9–11 July 2019.

Recipients of the award include Frances Allen, Grady Booch, Edgar Codd, Douglas Engelbart, Edward Feigenbaum, Tony Hoare, Robert Kahn, Jack Kilby, Dennis Ritchie, and David Wheeler. Further information about the award, including a list of all past recipients, may be found at www.computer.org/web/awards/pioneer.

0018-9162/18/$33.00

BJARNE STROUSTRUP RECEIVES 2018 IEEE CS COMPUTER PIONEER AWARD

Bjarne Stroustrup, managing director in the technology division of Morgan Stanley in New York, and visiting professor at Columbia University, was named the recipient of the 2018 IEEE CS Computer Pioneer Award. He is recognized "for bringing object-oriented programming and generic programming to the mainstream with his design and implementation of the C++ programming language."

This award—which will recognize two individuals in 2018—recognizes significant contributions to early concepts and developments in the electronic computer field, which have clearly advanced the state of the art in computing. Stroustrup is the designer and original implementer of C++, and he authored *The C++ Programming Language*, *A Tour of C++*, and *Programming: Principles and Practice using C++*, as well as many popular and academic publications.

Stroustrup accomplished much of his most important work in Bell Labs. His research interests include distributed systems, design, programming techniques, software development tools, and programming languages. To make C++ a stable and up-to-date base for real-world software development, he has been a leading figure with the ISO C++ standards effort for more than 25 years.

An IEEE, ACM, and Computer History Museum Fellow, Stroustrup is also a member of the US National Academy of Engineering (NAE). He is the recipient of the 2018 NAE Charles Stark Draper Prize for Engineering and the 2017 Institution for Engineering and Technology Faraday Medal.

He received a master's in mathematics from Aarhus University and a PhD in computer science from Cambridge University, where he is an honorary Fellow of Churchill College.

The Computer Pioneer Award was established in 1981 by the IEEE CS Board of Governors to recognize and honor the vision of those whose efforts resulted in the creation and continued vitality of the computer industry. The award is presented to outstanding individuals whose main contribution to the concepts and development of the computer field was made at least 15 years earlier.

The award will be presented at the 2018 COMPSAC, sponsored by the IEEE CS. The conference will be held 23–27 July 2018, in Tokyo.

The award consists of a silver medal.

Past recipients of the award include Frances Allen, Grady Booch, Edgar Codd, Douglas Engelbart, Edward Feigenbaum, Tony Hoare, Robert Kahn, Jack Kilby, Dennis Ritchie, and David Wheeler.

Further information about the award, including a list of all past recipients, can be found at www.computer.org/web/awards/pioneer.

SOREL REISMAN RECEIVES 2018 IEEE CS RICHARD E. MERWIN AWARD

Sorel Reisman, IEEE CS 2011 President and IEEE Senior Member, has been selected to receive the 2018 Richard E. Merwin Award for Distinguished Service. Reisman was recognized "for sustained contributions, leadership, and service to the IEEE CS, the IEEE, and the computing profession at large."

A professor of information systems in the Department of Information Systems and Decision Sciences at California State University, Fullerton, Reisman also serves as managing director of the international higher education, educational technology consortium, MERLOT (Multimedia Educational Resource for Learning and Online Teaching) at the California State University Office of the Chancellor.

Reisman has presented/published more than 120 articles including the books *Multimedia Computing: Preparing for the 21st Century*, and *Electronic Learning Communities – Current Issues and Best Practices*. He continues to write and review articles and columns for *IT Professional* and *Computer* magazines, and speaks internationally about his career-long passion on the use of technology for teaching computing.

In 2011, Reisman established the IEEE CS's Special Technical Communities (STCs) for which he received the IEEE CS Outstanding Contribution Award.

Reisman is chair of the Standing Committee of the IEEE CS COMPSAC conference, and he is also a member of the IEEE Computer Society Golden Core, Eta Kappa Nu, the IEEE Publications, Services, and Products Board, the IEEE Xplore Platform Guidance Committee, the IEEE Conference Publications Committee, and the Strategic Planning Committee of the IEEE Technical Activities Board. He has also served as a member of the Educational Activities Board, and he was editor in chief of the IEEE eLearning Library.

Reisman served as IEEE CS Publications Board vice president (2008–2009) and IEEE-CS Electronic Products and Services Board vice president (2006–2007). Prior to his executive committee positions, he chaired and served on many IEEE CS committees, including as member at large of the IEEE CS Publications Board, chair of the Magazines Operations Committee, founding member of the Computer Society Digital Library (CSDL) Committee, and chair of the Web Development Committee. He was an editorial board member of *IEEE Software*, a founding editorial board member of *IEEE Multimedia*, and *IT Professional* (for which he is currently Advisory Board chair).

Reisman received his EE degree, MA, and PhD in computer applications from the University of Toronto and was appointed a Fulbright Specialist in 2014.

The Merwin Award, the IEEE CS's highest level volunteer service award,

UNIVERSITY OF CALIFORNIA, IRVINE EXTENDS 15 PERCENT TUITION DISCOUNT TO IEEE COMPUTER SOCIETY MEMBERS

IEEE CS is pleased to announce that IEEE members will receive a 15 percent tuition discount for technical classes at the University of California, Irvine (UCI).

"This is an outstanding opportunity for Computer Society members to take advantage of the excellent courses offered by UCI," says Jean-Luc Gaudiot, 2017 IEEE CS President, and professor in the Department of Electrical Engineering and Computer Science, at UCI's Henry Samueli School of Engineering.

UCI is ranked among the top 50 universities nationally, and it is 10th among all public universities by *US News and World Report.* The Division of Continuing Education (DCE) at UCI has been offering education for adult learners for over half a century. Today UCI offers more than 60 convenient online programs with a focus in high job-demand industry sectors.

UCI's technical online course topics include:

- Embedded systems
- Communications systems engineering
- Device software engineering
- DSP systems engineering
- Ambient computing and IoT
- Predictive analytics
- Big data
- Data science
- Mobile applications
- Python programming
- Information systems security
- Systems engineering

To view all courses and programs that are of particular interest to IEEE members, visit https://ce.uci.edu/areas/ieee.

is given to individuals for outstanding volunteer service to the profession at large, including significant service to the IEEE CS. The award consists of a bronze medal and a $5,000 honorarium.

The award will be presented at the IEEE CS annual awards ceremony on 6 June 2018, in Phoenix.

Further information about the award, including past recipients, can be found at www.computer.org/web/awards/merwin.

MARGARET MARTONOSI TO RECEIVE 2018 IEEE COMPUTER SOCIETY TECHNICAL ACHIEVEMENT AWARD

Margaret Martonosi, an IEEE and ACM Fellow, has been selected to receive the 2018 IEEE CS Technical Achievement Award, "for contributions to power-aware computing and energy-constrained mobile sensor networks."

Martonosi is the Hugh Trumbull Adams '35 Professor of Computer Science and the director of the Keller Center for Innovation in Engineering Education at Princeton University.

Her research interests are in computer architecture and mobile computing, with particular focus on power-efficient systems. Martonosi's work has included the development of the Wattch power modeling tool and the Princeton ZebraNet mobile sensor network project for the design and real-world deployment of zebra tracking collars in Kenya. Her current research focuses on hardware–software interface issues in modern computing systems, including tools for quantum computing.

Martonosi has received numerous awards including the 2010 Princeton University Graduate Mentoring Award and the 2013 Anita Borg Institute Technical Leadership Award. Several of her papers have been honored for their long-term impact, including the 2015 International Symposium on Computer Architecture (ISCA) Long-Term Influential Paper Award, the 2017 ACM SIGMOBILE Test-of-Time Paper Award, and the 2017 ACM SenSys Test-of-Time Paper Award.

The IEEE CS Technical Achievement Award is given for outstanding and innovative contributions to the fields of computer and information science and engineering or computer technology, usually within the past 10, and not more than 15, years. Contributions must have significantly promoted technical progress in the field.

The award consists of a certificate and a $2,000 honorarium, and it will be presented at the 2018 Computer Software and Applications Conference (COMPSAC) sponsored by the IEEE Computer Society, which takes place 23–27 July 2018, in Tokyo.

Further information about the award, including a list of past recipients, is available at: www.computer.org/web/awards/technical-achievement.

IEEE COMPUTER SOCIETY: Be at the Center of It All

IEEE Computer Society membership puts you at the heart of the technology profession—and helps you grow with it.

Here are 10 reasons why you need to belong.

1 **Training** that sharpens your edge in Cisco, IT security, MS Enterprise, Oracle and more.

2 **Certifications** and exam preparation that set you apart.

3 **Industry intelligence,** including *Computer,* myCS, Computing Now, and myComputer.

4 Over 200 annual **conferences** and technical events held worldwide.

5 300-plus chapters and more than 30 technical committees **keep you connected.**

6 **Access** to hundreds of books, 13 technical magazines and 20 research journals.

7 **Deep discounts** on magazines, journals, conferences, symposia and workshops.

8 Opportunities to **get involved** through speaking, publishing and volunteering opportunities.

9 **Scholarships awarded** to computer science and engineering student members each year.

10 **A robust jobs board,** plus videos, articles and presentations to help you land that next opportunity.

IEEE Computer Society—keeping you ahead of the game. Get involved today.

www.computer.org/membership

IEEE computer society