Amplifying Human Cognition: Bridging the Cognitive Gap between Human and Machine

Evangelos Niforatos
Università della Svizzera italiana
Lugano, Switzerland
evangelos.niforatos@usi.ch

Athanasios Vourvopoulos
Madeira ITI, Universidade da Madeira
Funchal, Portugal
athanasios.vourvopoulos@m-iti.org

Marc Langheinrich
Università della Svizzera italiana
Lugano, Switzerland
marc.langheinrich@usi.ch

UbiComp/ISWC'17 Adjunct , September 11–15, 2017, Maui, HI, USA

ACM ISBN 978-1-4503-5190-4/17/09.
https://doi.org/10.1145/3123024.3129266

Abstract

Evolution has always been the main driving force of change for both the human body and brain. Presently, in the Information era, our cognitive capacities cannot simply rely on natural evolution to keep up with the immense advancements in the field of Ubiquitous technologies, which remain largely uninformed about our cognitive states. As a result, a so-called "cognitive gap" is forming between the human (users) and the machine (systems) preventing us from fully harnessing the benefits of modern technologies. We argue that a "cognitive information layer", placed in-between human and machine, could bridge that gap, informing the machine side about aspects of our cognition in real time (e.g., attention levels). In this position paper, we present our vision for such a software architecture, we describe how it could serve as a framework for designing and developing cognition-aware applications, and we showcase some application scenarios as a roadmap towards human-machine convergence and symbiosis.

Author Keywords

Human Cognition, Physiological Computing; Human Augmentation; Software Architecture; Framework.

ACM Classification Keywords

H.5.2. Information interfaces and presentation: User interface management systems (UIMS);

Introduction

Over the millennia, the human brain has evolved to excel in collecting (perception) and processing information (cognition) in an incredibly efficient manner. Essentially, the way our brain has evolved and structured is what makes us so different from all other animal species, many of which have significantly larger brains with even higher number of synapses [14]. For the human brain to continue evolving, certain organs may have to become bigger including the human brain per se. For example, for processing more information, wider synapses are needed, resulting in a demand for greater in-brain blood flow and in turn a larger heart. However, even if the human brain and its supporting organs and networks will eventually grow in size, further evolution will certainly face limits imposed by laws of Physics, as well as diminishing efficiency after a certain brain size increase threshold [8]. Nevertheless, these changes cannot occur naturally and in a timely manner for satisfying the frenetically increasing demands in information processing of the today's era.

Evidently, the way people collect and process information has always been influenced by technology. For example, the introduction of stone-headed javelins during hunting would have greatly altered decision making and strategic thinking of prehistoric humans. Similarly, the introduction of smartphones has greatly influenced the way we seek and consume information, highly disrupting the entire spectrum of our cognitive processes (attention, memory, learning, decision making, problem solving etc.). Albeit technology so far had a beneficial role on how we perceive and process the world around us, in recent years it has undertaken a rather disruptive and double-edged role. The era of ubiquitous technologies and the Internet of Things (IoT) finds our brains unprepared for handling the sheer volume of information produced daily by a multitude of sources. Attention deficit disorders, the multi-tasking illusion, learning difficulties, sleep deprivation, weak memory, chronic stress, etc., are just a few examples of the negative side effects that modern technologies impose on everyday life. As modern technologies become increasingly pervasive and even addictive [11], while blurring the limits of personal and professional life, such negative side effects are expected to further exasperate. Some companies limit access to technology after a certain hour (e.g., the “right to disconnect”), while individuals decide to abstain from using smart devices or social media for a period of time or even completely. But why should one have to resort to such practices after all? Is technology per se not meant to help us improve our quality of life, and eventually realize our full potential as human beings?

We believe that modern technologies hold the potential to greatly amplify human cognition in the entirety of its spectrum, seizing in a way the role of natural evolution [13]. In this position paper, we argue that the negative effects observed by increased technological use are simply side effects of our inability to keep up with the pace that technology advances. We attribute this phenomenon to the fact that despite the overall technological proliferation, the devices and systems (i.e., machine side) we use daily, still remain oblivious of our cognitive and affective states, assuming always the maximum of our cognitive capacities. We name this discrepancy “cognitive gap” and we propose a software architecture, we call the “Cognitive Information Layer” (CIL), for bridging this gap, essentially paving the way towards human-machine convergence.

Background

Already by 1965 (the year of the first space-walk), people envisioned a future where humans are linked to computers in a symbiotic manner for enhanced cognition. The 1965 Sunday comic strip "Our New Age" stated: "By 2016, man's intelligence and intellect will be able to be increased by drugs and by linking human brains directly to computers!"[1]. Underexplored, equally like outer-space, the study of the human brain gave slowly birth to neurotechnology - in the early 1970s -, proposing the use of electroencephalography (EEG) as a new way for directly linking brain activity with computers [16].

Interestingly, to date, computer systems and the human brain have already formed a basic -unidirectional- communication channel through Brain-Computer Interfaces (BCIs). Disabled people can now learn to control robotic limbs by the sheer power of their mind [19], stroke survivors can manipulate virtual limbs in virtual-reality [18] up to brain-controlled computer games designed for entertainment [12]. Undoubtedly, BCIs have driven a revolution in the areas of assistive and rehabilitative technologies, increasing people's quality of life.

Despite this technological advancement, however, the transition from using BCI technology in research labs to everyday life is still slow and far from the 1960s vision. The idea of closing the "cognitive gap" between Human and Machine has for long been discussed in the circles of academics, futurists, practitioners and science (fiction) enthusiasts. So far, certain approaches have focused on augmenting particular aspects of cognition, such as memory [4,5,10], attention [6,15] and learning [9].

Several endeavors promise to bring closer together the human mind with technology, in what has been named "Human Machine Confluence", essential the vision in which the human brain converges with the machine [7]. An EU project under this title has attempted to showcase that the concept may be viable in the future, identifying a set of research challenges 10 years ago in which very few advancements have happened since then. In the US, the BRAIN initiative[2] received initial funding of approximately $110 million from the Defense Advanced Research Projects Agency (DARPA), the National Institutes of Health (NIH), and the National Science Foundation (NSF). The EU Human Brain project[3], involving researchers from over 100 institutions, received funding over one billion Euros, together with criticism from Europe's leading neuroscientists. More recently, SpaceX and Tesla CEO Elon Musk has joined the BCI venture with a newly founded company called Neuralink[4]. This company is centered on creating implantable interfaces in the human brain, with the eventual purpose of helping human beings merge with software for a true human-machine symbiosis. Facebook unveiled a project on a BCI that could also be used by patients with severe paralysis. This will be a system that allows one to type even faster than one's physical hands, at upwards of 100 words per minute. The focal point of the debate is primarily the level of invasiveness of the employed cognitive intervention, with some arguing for brain implants, while others prefer non-invasive methods such as electrodes or information

[1] http://www.smithsonianmag.com/history/sunday-funnies-blast-off-into-the-space-age-81559551/

[2] https://www.braininitiative.nih.gov
[3] https://www.humanbrainproject.eu
[4] https://neuralink.com

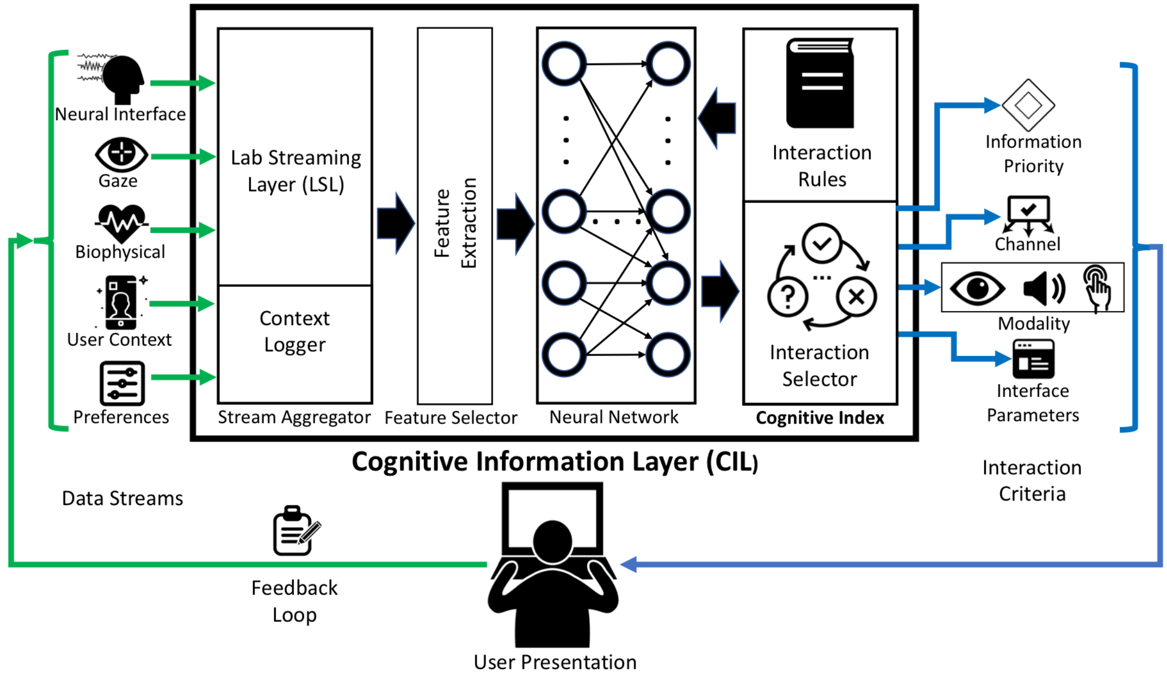

Figure 1: The Cognitive Index Layer (CIL) receives inputs of multimodal data streams to a stream aggregator component responsible for the data synchronization. The feature selector extracts meaningful features from the amassed multimodal data and delivers them as inputs to the neural network model that has been initially trained on a set of interaction rules. The cognitive index then outputs the adequate interaction criteria for adjusting user presentation to user's current cognitive state and the feedback loop closes with behavioral measures (e.g., task completion times etc.). (icons: thenounproject)

from multiple sources. Current technology fails to provide reliable brain-to-computer interaction. Many brain processes are still unknown even to modern neuroscience. Thus, to fully understand how to decode brain information, as captured by current technology, and infer the user's cognitive state in a less invasive manner, an intermediate layer is necessary for not only connecting human and machine, but doing so in a symbiotic manner, synergistically rather antagonistically.

Cognitive Information Layer

The main component of the envisioned architecture is an intermediate layer between the human and the machine, responsible for continuously monitoring the user's cognitive and affective states and informing the machine side for performing adequate interaction interventions. As "machine", we consider systems with which we interact frequently and increasingly depend on, such as personal computers, smartphones, smartwatches, public/personal displays and soon any IoT device. Interaction interventions regulate the information flow between human and machine to match the current cognitive and affective state of the user, user context and task at hand. An overview of the proposed cognitive information layer is shown in Figure 1. In short, the envisioned architecture is based on real time multimodal data stream inputs that reveal both current cognitive and affective states, task at hand and user context. Next, features are extracted and classification takes place for selecting the adequate interaction criteria based on a constantly updated "cognitive index". Finally, the user presentation is adapted on the fly according to the cognition-aware interaction criteria outputted. Below, we describe in more detail the primary stages and components of the envisioned architecture.

Multimodal Data Streams

We are currently experiencing an upheaval of data that characterize one's physiological states, mainly due to the miniaturization and reducing cost of sophisticated wearable physiological hardware [17]. Mobile eye-trackers (e.g., Tobii), affordable portable neural interfaces (e.g., OpenBCI) and biophysical monitoring wristbands (e.g., Empatica E4), produce a sheer volume of physiological data. Gaze behavior, Electroencephalography (EEG), Heart Rate Variability (HRV), and Electrodermal Activity (EDA) are just a few examples of physiological data that can reveal one's cognitive or affective states, stress levels and other intrinsic information [2]. This plethora of cognitive descriptive information can be utilized in conjunction with user context, task at hand and preferences to greatly increase the accuracy at which user cognitive and affective capacities can be estimated [1]. As shown in Figure 1, all multimodal data streams are aggregated continuously and then pushed into the feature selection stage, where data is filtered and useful features are extracted. Feature extraction aims at describing the acquired data streams with as few relevant values as possible. Such features should capture the information embedded in those data streams that is relevant to describe the mental states, while rejecting the noise and other non-relevant information. The next step, denoted as "classification" assigns a class to a set of features extracted from the data streams. This class corresponds to the kind of mental state identified. Classification algorithms are known as "classifiers". Typically, for learning which kind of feature vector corresponds to which class (or mental state), classifiers try either to model which area of the feature space is covered by the training feature vectors from each class or they try to model the boundary between the areas covered by the training feature vectors of each class mostly used in BCIs.

The Cognitive Index

When first installed, the Cognitive Index (CI) is simply a catalog of appropriate interaction rules for efficient management of one's cognitive resources, extracted from guidelines and best practices available in literature. However, the CI is constantly updated to match the cognitive capacities, preferences, task at hand and context of each individual user. Next, the CI is responsible for selecting the appropriate interaction criteria in terms of information prioritization (e.g., e-mails over social media notifications when at work), adequate channel (i.e., mobile vs. stationary), modality (i.e., visual, audible or tactile) and certain interface parameters (e.g., font size), as shown in Figure 1. Ultimately, the CI will be enriched by interconnecting the CIs of different users for creating a universal pool of cognitive interaction mapping.

The Feedback Loop

Finally, the outputted interaction criteria have resulted in tangible interaction interventions that are delivered in real time to the user, by appropriately altering user presentation to better match current cognitive capacities and needs. At this point, a closed feedback loop informs in situ the cognitive information layer about the effect that the modified user presentation has on user's cognitive state via behavioral measures (e.g., task completion times, motion patterns etc.). This serves as a real-time assessment mechanism for evaluating the effect of the selected interaction intervention, and adapting the interaction criteria accordingly.

Cognitive Application Framework

Knowledge obtained from the deployment of the CIL will be the basis for eliciting requirements and guidelines for informing the design and development of cognition-aware applications. "Cognitive applications" will consider user cognitive states for adapting information flow, interaction techniques and interface parameters. Design guidelines, requirements and best practices will be incorporated into a cognitive application framework along with a dedicated cognitive Application Programming Interface (API). The cognitive API will connect cognition-aware applications with the CIL for providing interaction criteria that facilitate communication between human and machine. For instance, the cognitive API will inform cognition-aware applications such as a text editor, an (mobile) e-mail client and a music player, that user's attention levels are currently scarce. In turn, the cognition-aware text editor will increase window and font size, and contrast, for keeping user focused to a text typing task. Simultaneously, the cognition-aware e-mail app will suppress incoming e-mail notifications for avoiding disrupting user's attention, and the cognition-aware music player app will opt in for music that helps one focus. This application scenario showcases that the CIL, via a dedicated cognitive API, could synergistically amplify human cognition by informing and adapting information flow and interaction across multiple applications and channels (mobile/stationary) simultaneously and surreptitiously. This CIL characteristic may be vital for our cognition, rectifying and regulating the fierce competition of modern applications over our limited cognitive capacities.

Challenges

Clearly, the proposed approach is difficult to realize. From a software architecture perspective, challenges lie throughout the entire stack of the cognitive information layer. First, multimodal data streams can be largely heterogeneous with wildly varying sampling rates and fundamentally divergent acquisition logic. For example, an EEG signal is continuously obtained from 250 - 500 Hz, whereas a location transition can be sampled asynchronously, yet both data types need to be synchronized and processed together. Even so, accurately inferring user's actual cognitive state (e.g., for ground truth acquisition) remains a conundrum. Selecting the right classifier for features derived from a plethora of multimodal data streams, could be a solution but also a considerable challenge. In Figure 1, we assume a neural network (NN) as the most adequate classifier candidate due to the NN's ability to deal with highly heterogeneous input. Finally, the outputted interaction criteria should be implemented within the cognitive application per se. This means that (mobile) applications that one uses daily (e.g., an Internet browser) will need to be re-programmed to accommodate for changes that the outputted interaction criteria recommend. However, we believe that the biggest challenge is a shift in how designers and developers create software for everyday use. Cognition-aware software will have to consider the user's cognitive states and adapt its functionalities and presentation accordingly.

From a sociotechnical perspective, challenges lie primarily in user adoption. Physiological sensing hardware (e.g., an EEG system) is still expensive, cumbersome to use, tiresome to wear for extended periods of time, and socially unacceptable. Furthermore, current hardware solutions are prone to noise generated from movement or muscle activity, heavily degrading the quality of the acquired signal. However, as hardware miniaturization progresses, with wearables and IoT becoming more prevalent, we expect that physiological sensing will become gradually mainstream, perhaps even in an implant level (e.g., Neuralink).

From a user privacy perspective, the streams of data supplied to the CIL are highly sensitive, able to reveal apart from user's cognitive state, also health and affective states, and thus data protection measures and policies will be of paramount importance. Also, the diffusion of knowledge about our cognitive states to a network of interconnected objects (IoT) may also raise unexpected ethical and security concerns [3].

Despite the considerable challenges, we believe that the proposed architecture, or a similar one, will be implemented and seamlessly integrated in an operating system level in the mid-term future. This integration will form and enforce design and development policies for creating cognition-aware applications.

Summary

Human brain evolution is thought to have already reached its apex together with our cognitive capacities. Some argue that one way forward is through achieving human-machine symbiosis, the notion of human converging with the machine. While sounding like a Sci-Fi scenario, we argue that not only could it soon be a reality, but also a necessity. In this position paper, we introduced the cognitive information layer as an inset between human and machine, for informing the machine side about current user's cognitive state and facilitating human-machine interaction. Thus, we define human cognition amplification as the optimization of existing cognitive resources, rather than extending human abilities beyond the humanly possible. We illustrated how by supplying a set of multimodal data streams to the proposed layer, it can output a set of interaction criteria as a pivot for manipulating user presentation with cognition-aware applications. We identified a range of challenges, including the need to reform traditional software (and not only) design thinking so to create software that respects our cognitive capacities. We believe that the human brain and technology can and should be able to work more closely in tandem for amplifying our cognitive capacities in the era of distractions and information overload.

References

1. P. D. Bamidis, C. Papadelis, C. Kourtidou-Papadeli, C. Pappas, and A. B. Vivas. 2004. Affective computing in the era of contemporary neurophysiology and health informatics. *Interacting with Computers* 16, 4: 715–721.
2. C. Berka, D. J. Levendowski, M. N. Lumicao, et al. 2007. EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory

Tasks. *Aviation, Space, and Environmental Medicine* 78, 5: B231–B244.

3. J. Bohn, V. Coroamă, M. Langheinrich, F. Mattern, and M. Rohs. 2004. Living in a World of Smart Everyday Objects—Social, Economic, and Ethical Implications. *Human and Ecological Risk Assessment: An International Journal* 10, 5: 763–785.
4. N. Davies, A. Friday, S. Clinch, et al. 2015. Security and privacy implications of pervasive memory augmentation. *Pervasive Computing, IEEE* 14, 1: 44–53.
5. T. Dingler, P. El Agroudy, H. V. Le, et al. 2016. Multimedia Memory Cues for Augmenting Human Memory. *IEEE MultiMedia* 23, 2: 4–11.
6. J. Dostal, P. O. Kristensson, and A. Quigley. 2013. Subtle gaze-dependent techniques for visualising display changes in multi-display environments. *Proceedings of the 2013 IUI*, ACM, 137–148.
7. A. Gaggioli, A. Ferscha, G. Riva, S. Dunne, and I. Viaud-Delmon. 2017. *Human Computer Confluence, Transforming Human Experience Through Symbiotic Technologies*. De Gruyter Open, Berlin.
8. M. A. Hofman. 2014. Evolution of the human brain: when bigger is better. *Frontiers in Neuroanatomy* 8.
9. K. Kunze, M. Iwamura, K. Kise, S. Uchida, and S. Omachi. 2013. Activity recognition for the mind: Toward a cognitive‟ Quantified Self.” *Computer* 46, 10: 105–108.
10. E. Niforatos, C. Cinel, C. C. Mack, M. Langheinrich, and G. Ward. 2017. Can Less be More?: Contrasting Limited, Unlimited, and Automatic Picture Capture for Augmenting Memory Recall. *Proceedings of the ACM IMWUT* 1, 2: 1–22.
11. A. Oulasvirta, T. Rattenbury, L. Ma, and E. Raita. 2012. Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing* 16, 1: 105–114.
12. R. Scherer, E. C. V. Friedrich, B. Allison, et al. 2011. Non-invasive Brain-Computer Interfaces: Enhanced Gaming and Robotic Control. *Advances in Computational Intelligence*, Springer, Berlin, Heidelberg, 362–369.
13. A. Schmidt. 2017. Augmenting Human Intellect and Amplifying Perception and Cognition. *IEEE Pervasive Computing* 16, 1: 6–10.
14. P. Thomas Schoenemann. 2006. Evolution of the Size and Functional Areas of the Human Brain. *Annual Review of Anthropology* 35, 1: 379–406.
15. R. Vertegaal, J. S. Shell, D. Chen, and A. Mamuji. 2006. Designing for augmented attention: Towards a framework for attentive user interfaces. *Computers in Human Behavior* 22, 4: 771–789.
16. J. J. Vidal. 1973. Toward Direct Brain-Computer Communication. *Annual Review of Biophysics and Bioengineering* 2, 1: 157–180.
17. A. Vourvopoulos and S. B. i Badia. 2016. Usability and Cost-effectiveness in Brain-Computer Interaction: Is it User Throughput or Technology Related? ACM Press, 1–8.
18. A. Vourvopoulos, A. Ferreira, and S. B. i Badia. 2016. NeuRow: An Immersive VR Environment for Motor-Imagery Training with the Use of Brain-Computer Interfaces and Vibrotactile Feedback: SCITEPRESS - Science and Technology Publications, 43–53.
19. J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. 2002. Brain-computer interfaces for communication and control. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 113: 767.