Learning Analytics as AI Extenders in Education: Multimodal Machine Learning versus Multimodal Learning Analytics

Mutlu Cukurova, PhD
University College London, United Kingdom, m.cukurova@ucl.ac.uk

I. Introduction

The nature of the appropriate role for AI is a topic of great interest in many disciplines, and Education is no exception. Perhaps, the initial focus of AI in Education research was on attempts to create systems that are as perceptive as human teachers [1]. Therefore, the majority of early research in the field focused on designing autonomous intelligent tutoring systems. However, more recently, there have been AI technologies embedded in non-autonomous systems, used by educators to support their practice. Non-autonomous systems, in which AI is used to extend human cognition and enhance teacher and learner capabilities, differ significantly from approaches that aim to create fully automated AI systems. These non-autonomous approaches might even be considered as 'incomplete' or 'inadequate' in AI research. Here, I argue that, in educational contexts, AI systems should be considered a continuum with regards to the extent they are decoupled from humans, rather than only an approach to provide full-automation. AI can be used to externalize, internalize or extend human cognition [2], and these different conceptualizations and implementations of AI can each have a valuable role to play in the support of learning and teaching. In this paper, I will briefly describe the distinctions between AI as a fully autonomous system versus AI as part of a non-autonomous supportive system, and present examples of both in educational contexts. I will particularly focus on multimodal learning analytics where the human cognition is internalised or extended with AI tools, rather than externalised. For educational research, where the ultimate purpose is to improve education rather than improving the state-of-the-field in AI, AI extenders as exemplified in learning analytics research should be distinct from research on fully autonomous AI designs; and they require attention from the field *at least* in equal measure.

II. Relevant Theories

A. *AI to Extend Cognition and Enhance Capabilities*

The idea of computers as interactive systems to support and potentially augment human capabilities is not new [3], yet, recently, it is re-visited in the context of AI systems designed to augment human intelligence [4]. The general idea of technology helping humans become more capable through expanding their cognitive capacities is usually associated with the thesis of extended cognition [5]. This argues that human cognition can be partially represented by artefacts that are able to exemplify the right kinds of computations. It maintains that the artefacts we use to help us complete cognitive tasks can become integrated into our biological capacities and can play a functional role in triggering our cognitive abilities. Therefore, rather than considering human cognition as limited to intracranial biological activities, the theory proposes that the wider processes that take place in the surrounding environments of humans can actually constitute human cognition [5]. These "cognitive artefacts" [6] can replace some of the functions of the brain through extended cognition, but also can move beyond them to potentially enhance cognitive capabilities. Based on earlier work [6], more recent researchers defined a cognitive extender as "…an external physical or virtual element that is coupled to enable, aid, enhance, or improve cognition, such that all – or more than – its positive effect is lost when the element is not present" [2, p.3]. Recently, learning analytics research generated valuable examples of cognitive extenders in various forms (i.e teacher and learner dashboards) to support and extend teacher and learner cognition. Positioning AI systems as cognitive extenders is significantly different than positioning AI systems as fully autonomous external systems. The first situates "the intelligence" within the human-artificial coupled systems, whereas the latter, in the artificial. There is relatively less research undertaken in non-autonomous AI systems coupled with humans, however particularly in areas of social sciences such as Education, focus on these systems might lead to more productive outcomes.

III. Technological Advances and Real World Applications

The literature on such theoretical and philosophical considerations on the different conceptualizations of AI might appear too abstract for researchers and practitioners of the AI in Education field. Nevertheless, a significant question to ponder upon and investigate for the field is w*hat might different conceptualizations of AI systems as a continuum with regards to their autonomy mean for AI in Education research and practice?* AI systems can vary considerably in the extent of their autonomy and symbiosis with humans, and design decisions about the autonomy of an AI system have significant social and ethical implications. Colleagues in [2], make a valuable contribution to this discussion with their arguments suggesting to consider non-autonomous AI systems distinctively from autonomous, human decoupled AI systems. However, general recommendations naturally fall short in terms of their particular implications in specific research areas. In the next section, I attempt to make some of the discussed distinctions more concrete within the context of multimodal AI systems in Education.

A. *Multimodal Machine Learning vs Multimodal Learning Analytics*

Modality can be defined as the type of communication channel used by two agents to convey and acquire information that defines the data exchange [7]. Some modalities used in AI in Education research include, but are not limited to, video, audio, text, click-stream, eye tracking, electroencephalography (EEG), and electro-dermal data. Multimodality has great potential to help us understand the complex world around us and it has been studied for around three decades in the context of social semiotics. Its potential to interpret complex social phenomena led AI researchers to try and build models that can process information from multiple modalities through machine learning techniques and social signal processing [8]. The literature on multimodal machine learning is rich with examples of audio-visual speech recognition; multimedia content indexing and retrieval, as well as multimodal affect recognition [9]. Learning from multimodal data provides opportunities to acquire an in-depth understanding of complex processes and, for AI research to make progress, it makes sense to focus on multimodal machine learning models that can process and relate information from multiple modalities [8]. However, in multimodal machine learning research, the ultimate aim is to create fully-autonomous systems that in essence replicate the human cognition through decoupling humans from the system by making machines capable of processing multimodal data at a similar accuracy. In this sense, they can be considered on the one extreme of the AI continuum towards high autonomy. For instance, specifically in learning contexts, there are recent attempts that aim to interpret various modalities of data including click-stream data, eye-tracking, EEG, video, and wristband data to automatically predict learning performance in game contexts [10]. These kinds of studies are great examples of showcasing the potential of machine learning approaches to unravel complicated manifolds in complex educational data including non-linear effects and multivariate interactions. They are also good examples to present the superiority of multimodal over unimodal approaches in terms of predicting learning performance automatically. The full-automation is particularly useful for the provision of personalised support to learners through intelligent tutoring systems and adaptive learning platforms.

On the other hand, in educational contexts, there are also plenty of multimodal AI technologies that are embedded in non-autonomous systems used by educators as support tools. Most of this research is undertaken under the emerging area of multimodal learning analytics and they provide promising opportunities for the advancement of educational research and practice. For instance, researchers recently proposed a multimodal feedback approach for learner reflections based on posture, gaze, volume, and performance data [11]. Similarly, within the context of collaborative learning, data from verbal and physical interactions of students are used to provide insights into their collaborative actions around table-top computers [12]. Furthermore, there are examples of similar aimed research investigating the potential of multimodal data to support learners' self-regulation performance [13]; project-based learning [14], tutor evaluations in debating [15], and classroom orchestration for teachers [16]. In this line of research, certain tasks and activities are indeed automated with the help of AI approaches. Nevertheless, in essence, most multimodal learning analytics approaches aim to provide explicit and comprehensible ways of presenting information to learners and teachers to make them more informed decision makers. Therefore, these AI technologies are designed to be tightly coupled with humans to become part of their extended cognition, and ultimately enhance their capabilities in teaching and learning. This work requires a greater emphasis on tutor and learner interfaces of non-autonomous AI systems as they need to be smoothly internalized by humans. It is important to note that the interfaces do not necessarily explain how the subsystem AI works to teachers and learners, but they provide opportunities for cognitive processes to be instantiated and internalized by humans to make the most of the multimodal learning analytics tool. In this sense, they are less autonomous systems of the AI continuum. Due to the significant difference in their respective ultimate goals, the design of autonomous AI systems as exemplified in the multimodal machine learning research, and the design of multimodal learning analytics research should be considered as relevant but distinct initiatives in the AI in Education field.

IV. Summary

In this paper, I argued for the value of considering the autonomy of AI systems as a continuum in educational research. Within the context of AI in Education, I presented two paradigms of i) multimodal machine learning that aims to create AI through externalisation and replication of human cognition and ii) multimodal learning analytics that aims to design artefacts that involve AI technology, but also tightly coupled with humans to enable, aid or extend their cognition and enhance their capabilities. The first approach is particularly useful for the provision of personalised support for learners through intelligent tutoring systems. However, education systems are much broader than what intelligent tutoring systems can provide on their own. Moreover, to keep the interest of AI in Education research *only* on the design of systems that can mimic or replace human tutors, would limit the possibilities of AI in Education to supporting tutors to reach a "standard human tutor" level. On the other hand, if the focus is on non-autonomous AI systems that are not like human, but human-centred, there are greater opportunities to extend human cognition and enhance our capabilities in both teaching and learning.

References

[1] Self, J.A. (1998) The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. *International Journal of Artificial Intelligence in Education*, 10, 350-364.

[2] Vold, K., & Hernandez-Orallo, J. (2019). AI Extenders: The Ethical and Societal Implications of Humans Cognitively Extended by AI. *Proceedings of AAAI / ACM Conference on Artificial Intelligence, Ethics, and Society* https://doi.org/10.17863/CAM.36128

[3] Engelbart, D. C. (1962). *Augmenting human intellect: a conceptual framework*. Summary Report AFOSR-3233, Stanford Research Institute, Menlo Park, CA.

[4] Rouse, W. B., and Spohrer, J. C. (2018). Automating versus augmenting intelligence. *Journal of Enterprise Transformation*, 1–21.

[5] Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.

[6] Hutchins, E. (1999). Cognitive artefacts. *The MIT encyclopedia of the cognitive sciences*, 126-127

[7] Kress, G. (2009). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

[8] Vinciarelli, A., Pantic, M., & Bourlard, H. (2009). Social signal processing: Survey of an emerging domain. *Image and vision computing*, *27*(12), 1743-1759.

[9] Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *41*(2), 423-443.

[10] Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management*, *48*, 108-119.

[11] Ochoa, X., Domínguez, F., Guamán, B., Maya, R., Falcones, G., & Castells, J. (2018). The rap system: Automatic feedback of oral presentation skills using multimodal analysis and low-cost sensors. *Proceedings of the 8th international conference on learning analytics and knowledge*, 360–364.

[12] Martinez-Maldonado, R., Dimitriadis, Y., Martinez-Monés, A., Kay, J., & Yacef, K. (2013). Capturing and analyzing verbal and physical collaborative learning interactions at an enriched interactive tabletop. *International Journal of Computer-Supported Collaborative Learning*, *8*(4), 455-485.

[13] Di Mitri, D., Scheffel, M., Drachsler, H., Börner, D., Ternier, S., & Specht, M. (2017). Learning pulse: A machine learning approach for predicting performance in self-regulated learning using multimodal data. *Proceedings of the seventh international learning analytics & knowledge conference*, 188–197.

[14] Spikol, D., Ruffaldi, E., Dabisias, G., & Cukurova, M. (2018). Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *Journal of Computer Assisted Learning*, *34*(4), 366-377.

[15] Cukurova, M., Kent, C., & Luckin, R. (2019). Artificial Intelligence and Multimodal Data in the Service of Human Decision-making: A Case Study in Debate Tutoring, *British Journal of Educational Technology*, *1-22*.

[16] Prieto, L. P., Sharma, K., Kidzinski, Ł., Rodríguez-Triana, M. J., & Dillenbourg, P. (2018). Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *Journal of Computer Assisted Learning*, 34(2), 193–203.