Towards high-bandwidth computer-human understanding: Cognitive chunking theory and micro-behavioural signals.

Peter C-H. Cheng[0000-0002-0355-5955]

Department of Informatics, University of Sussex, Brighton, BN1 9QJ, UK

Abstract. Chunking theory from cognitive science provides a basis for analyzing micro-behaviours in human performance in order to build models of individuals' understanding of domain content that are richer than those available from current methods used for human-machine communication in AI systems.

Keywords: Human-machine communication, user models, knowledge structure, chunking theory, micro-behavioural signals, pause analysis.

Introduction. How can we build artificial intelligent agents that can teach, support and collaborate with humans at human-like levels? To possess such abilities an AI agent must be able to rapidly model what a human knows about the task that they are sharing. Such models must be specific to each human, so that an agent's interaction is personally relevant to the individual. Further, the AI agent should continuously update the model, because cognitively we are first and foremost organisms that adapt and learn. Without such models humans will perceive agents as rigid, uncooperative or simple-minded.

Human teachers and coaches set the standard. They quickly gauge learners' competence and rapidly assess their knowledge by exploiting myriad subtle behavioural signals beyond the overt content of verbal responses. Unfortunately, the information that current AI systems use to construct user models are relatively low-bandwidth; for instance, where AI tutors record keystrokes the data is subsequently aggregated so analysis is typically rather coarse grained.

Cognitive science theory. AI agents may be imbued with the critical ability to rapidly and accurately model individual human competency and knowledge by exploiting ideas from cognitive science. We focus upon *chunking theory*, a long-established and empirically well-supported theory [1, 2, 3]. According to the theory humans process information as chunks, which are strongly associated clusters of concepts. During learning we build hierarchies of chunks in memory. The size of chunks grows with expertise. When we engage in any knowledge-based task we mentally traverse our specific hierarchy for that task. But, significantly, our central executive is essentially a serial processor, so the retrieval of information from memory produces spatial and temporal microbehaviours that can serve as natural signals about the structure of chunk hierarchies. These chunk-signals are apparent in writing, drawing and speaking, at times scales between 100ms to a few seconds. So, the capture and analysis of chunk-signals can expand the current bottleneck in machine-human collaboration, by providing a rich stream of task relevant information about what concepts an individual knows and how well they are known.

We have designed and evaluated tools and methods to exploit chunk-signals, including studies that: probe task competence and knowledge in real domains (e.g., mathematics [4, 5], English and Dutch [6, 7], programming [8], diagrammatic reasoning [9]); span adults and children; use diverse interfaces (e.g., freehand writing [4, 6, 7, 8], element selection with a mouse [5], natural drawing [9]); use a variety of tasks (e.g., transcription [4, 5, 7, 8], recall from memory [6, 9]). The results show that spatial and temporal chunk-signals can provide good measures of an individual's memory organization and measures of competence with educationally useful levels of accuracy.

Sample study. One experiment [5] recorded students' handwritten transcriptions of mathematical formulas of varying levels of difficulty, from simple arithmetic equations through to partial differential equations. Their writing was captured with a graphics tablet and pause durations between pen strokes were computed. Fig. 1 shows typical performance graphs for participants with low and high mathematical competence. The pauses reflect the amount of mental processing required prior to the production of each stroke; e.g., small for the second stroke of a character (e.g., '+') and large for transitions between chunks. The novice's many long pauses (>500ms) implies that they processed the equation using many chunks. In contrast, the expert used just three chunks, separated by two peaks that occurred at the breaks between terms (at '+'s). Thus, the profiles of pauses reveal the students' very different levels of familiariy with this equation.

We have developed various performance measures based on the distributions of pauses [4, 5, 8]. The strength and robustness of the underlying temporal chunk-signals means that these measures: (a) are strongly correlated

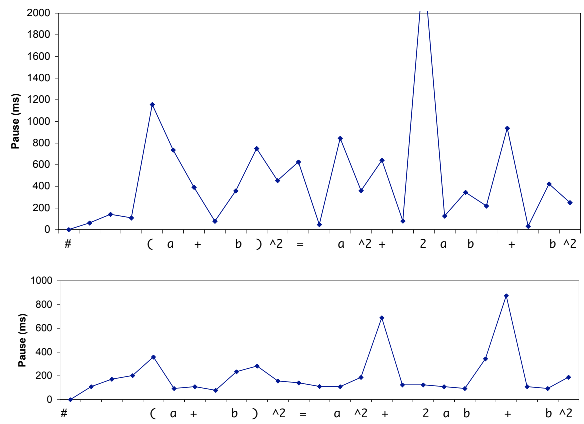

Fig. 1. Durations of pauses between strokes of handwritten transcriptions of a formula (quadratic expansion): top – novice, bottom – expert; ^2 is ^{2}.

with independent measures of domain competence; (b) are sensitive to the relative difficulty of stimuli for participants at different levels of competence [4-8]; (c) are obtainable from quick tests that just include a handful of items.

Implications for machine-human understanding. Our application of chunking theory to the analysis of human micro-behaviours demonstrates the feasibility of creating novel measures of performance to assess *what* humans know, how *well* they know it, and how it is *structured* in memory. Because these measures are based on sub-second signals, they provide information at a granularity far finer than available in current human-machine communication methods. For example, in tutoring systems, assessment tasks may be conducted in a shorter time and with fewer items. Significantly, learning activities in a tutoring system may be judiciously adapted to enable the inconspicuous continuous recording of such measures, thus eliminating the need for overt testing and creating the potential to model changes to students competence in detail, possibly at a concept by concept level, in real time.

All this highlights the need for collaboration between AI developers and cognitive scientist in order to utilize fully our rich knowledge of human cognition.

Acknowledgement. This work was support by EPSRC grant EP/R030642.

References

1. Simon, H. A. (1974). How big Is a chunk? Science, 183(4124), 482.
2. Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Science, 24(1), 87-114.
3. Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C.-H., Jones, G., Oliver, I., & Pine, J. (2001). Chunking mechanisms in human learning. Trends in Cognitive Science, 5(6), 1236-1243.
4. Cheng, P. C.-H. (2014). Copying equations to assess mathematical competence: An evaluation of pause measures using graphical protocol analysis. In Proc. of the 36th Ann. Conf. of the Cognitive Science Society (pp. 319-324). Austin, TX: Cognitive Science Society.
5. Cheng, P. C.-H. (2015). Analyzing chunk pauses to measure mathematical competence: Copying equations using 'centre-click' interaction. In Proc. of the 37th Ann. Conf. of the Cognitive Science Society (pp. 345-350). Austin, TX: Cognitive Science Society.
6. van Genuchten, E., & Cheng, P. C. H. (2009). Missing working memory deficit in dyslexia: children writing from memory. In Proc. of the 31st Ann. Conf. of the Cognitive Science Society (pp. 1674-1679). Austin, TX: Cognitive Science Society.
7. Zulkifli, M. (2013). Applying Pause Analysis to Explore Cognitive Processes in the Copying of Sentences by Second Language Users. University of Sussex (Unpublished PhD Thesis).
8. Albehaijan, N., & Cheng, P. C.-H. (2019). Measuring programming competence by assessing chunk structures in a code transcription task. In Proc. of the 41st Ann. Conf. the Cognitive Science Society. Austin, TX: Cognitive Science Society.
9. Roller, R., & Cheng, P. C.-H. (2014). Observed strategies in the freehand drawing of complex hierarchical diagrams. In Proc. of the 36th Ann. Conf. of the Cognitive Science Society (pp. 2020-2025). Austin, TX: Cognitive Science Society.