xAPI and temporal analytics: Designing and developing open standards to store and analyze temporal learner data in and around an LMS

Danny Y.T. Liu, Edward Moore, James Hamilton, Yvonne-Noemi Nemes
Macquarie University NSW 2109 Australia
[danny.liu,ed.moore,james.hamilton,yvonne.nemes]@mq.edu.au

There is a wealth of data already captured by learning management systems (LMSs), especially from courses that are well-designed to take advantage of a variety of online activities. Our preliminary conversations with instructors and students have highlighted desires to understand what particular groups of students are 'doing' in their courses and how they are interacting with resources, in the context of engagement and course outcomes. With increasing focus on the importance of instructional and other contexts in learning analytics (LA), our approach to temporal analytics works at the level of individual courses to discover and explore interesting patterns. For instructors, this means that the temporal information can be used to assess intended learning designs, perhaps in conjunction with a semantic overlay around resources, tasks, and supports. For students, salient patterns could be used to provide more meaningful feedback compared to aggregate counts of LMS activity.

These aggregate, atemporal counts, averages, and sums of activity are typical of practically-available LA tools. This is, in part, a result of the unwieldy, difficult-to-interrogate, and often outright arcane database structures in modern day LMSs. We therefore present our approach to temporal analytics which combines nascent open standards for the storage and analysis of learner data. Moving beyond the limited data present in log tables, we adopted the open Experience API (xAPI) standard and built a tool to extract, transform, load (ETL) and analyze Moodle and other data via a learning record store (LRS). We see the benefits of xAPI for temporal analytics being that each statement is a self-contained and human- and machine-readable account of granular learning interactions. Also, the statement structure is sufficiently flexible to possibly store novel multimodal data into the future, and is largely independent of LMS platforms. This parallels other LA work to store learner data in xAPI format and build LMS-agnostic analytical tools.

As a proof of concept, we used our tool to ETL more than 30 million rows of recent Moodle data into the Learning Locker LRS. However, because the Learning Locker querying engine was insufficiently powerful or complex for temporal analysis, we built our own analysis engine which directly accessed the underlying database. One of our proof-of-concept visualizations uses Sankey diagrams to combine Riemann's event- and variable-centred temporal analytics with possibilities for Zhou et al.'s student-, session-, and object-based sequence modelling ideas. These visualizations allow instructors and students to customizably see the sequences of interactions with learning activities/objects/events, overlaid with variables such as study strategies and course outcomes. These allow users to perform post-hoc discovery on the approaches of, for example, high-performing vs poorer-performing students, or internal vs external students, or strategic vs surface learners. We envisage this to be useful to, for example, instructors in understanding the ordering of learning events and activities taken by students, possibly with the view to understand if their instructional designs are effective, and being used as intended.

Analyses of temporal data are in need of open and widely-adopted standards for the storage of informative learning data. The increasing adoption of the xAPI standard means that more and more learner data are being stored in this meaningful format. Additionally, open source analysis tools that read xAPI data can then be applied adopted by other practitioners and researchers to further the progress of temporal analytics. In this ecosystem, the xAPI standard levels the data playing field to enable interoperability of learning analytics, and the sharing of enabling tools. To promote this, our aim is to release all our code as open source to help catalyze the collaborative development of temporal analytics.