Teacher's intervention in debugging activities

« 9 » The ARM results provide information for two types of interventions, at the individual level as well as at the classroom level. It is easier to understand how teachers can intervene at the individual level. However, the authors claim that ARM is potentially fruitful for analyzing conceptual changes at the classroom level. The questions then are: What kind of approach should the teacher take to intervene at the classroom level? In this situation, which activities can be done, and how can the teacher help students debug their misconceptions? (Q3)

References

Berland M., Baker R. S. & Blikstein P. (2014) Educational data mining and learning analytics: Applications to constructionist research. Technology, Knowledge and Learning 19(1–2): 205–220.

Guo Y., Xing W. & Lee H. (2015) Identifying students' mechanistic explanations in textual responses to science questions with Association Rule Mining. In: Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW). IEEE Computer Society, Washington DC: 264–268.

Imai M. (1986) Kaizen: The key to Japan's competitive success. McGraw-Hill, New York.

Papert S. (1980) Mindstorms: Children, computers and powerful ideas. Basic Books, New York.

Sussman G. J. (1975) Computer model of skill acquisition. American Elsevier, New York.

José A. Valente is a Professor at the Multimedia Department at the State University of Campinas (UNICAMP) in Brazil, teaching graduate and undergraduate students on topics related to programming and the development of digital narratives. He earned a PhD from the Department of Mechanical Engineering and Division for Study and Research in Education, at Massachusetts Institute of Technology (MIT). Currently he is developing research projects related to the use of active teaching methodologies such as maker activities and the integration of digital technologies and curriculum.

Received: 26 May 2019
Accepted: 10 June 2019

Knowledge Pieces, Causality in Complex Systems, and Computational Methods for Knowledge Analysis

Yu Guo
Northwestern University, USA
yuguo2012/at/u.northwestern.edu

> Abstract • This commentary responds to Hjorth and Wilensky's article by providing more literature and conversations on three themes: the piecemeal view of knowledge, the difficulties of identifying causality in complex systems, and some considerations on applying computational methods to knowledge analysis.

« 1 » Arthur Hjorth and Uri Wilensky's target article is about applying Association Rule Mining to the data collected from a constructionist learning activity. The article not only shows an innovative bridging between qualitative data analysis and machine learning methods, it also shows the authors' careful consideration about the grain size of knowledge pieces and how they changed from pre- to post-tests in university social policy classrooms.

« 2 » Below, I organize my comments to Hjorth and Wilensky's article into three themes: (a) literature and difficulties in conceptual change; (b) the difficulty of identifying causality in complex systems; and (c) considerations when applying computational methods to knowledge analysis.

Conceptual change and knowledge representation

« 3 » There is no agreement on the definition of conceptual change. Usually, conceptual change refers to learning that is more dramatic, compared to more mundane knowledge gain. Conceptual change is notoriously difficult for learners. Many researchers have been attracted to study this topic in a variety of subject areas in order to explain why it is so difficult to bring about. Examples of research in conceptual change include elementary-school students learning the infinite divisibility of number and matter (Smith, Solomon & Carey 2005); children's understanding of animal and human biological properties (Carey 1988); and college students' understanding of force and acceleration (diSessa & Sherin 1998). These researchers can be classified into two camps: the theory-theory camp and the knowledge-in-pieces camp. The former believes that learners hold misconceptions that are coherent and systematically organized, hence, robust and very hard to changes (e.g., Vosniadou 1992), whereas the latter believes that knowledge exists as tiny primitives that are formed through interacting with the world and are situationally constrained or activated in various combinations based on experience and context (e.g., diSessa & Sherin 1998).

« 4 » Bruce Sherin (2006) cites John Smith, Andrea diSessa, and Jeremy Roschelle (1993: 539) and argues that "any framework that entails wholesale replacement of intuitive knowledge is fundamentally at odds with constructivism." They proposed "an analytical shift from single units of knowledge to systems of knowledge with numerous elements and complex substructure that may gradually change, in bits and pieces and in different ways" (Smith, diSessa & Roschelle 1993: 148). The complex structure is called coordination class, hence, conceptual change involves changes in the way it coordinates the knowledge pieces across different situations (diSessa & Sherin 1998).

« 5 » Hjorth and Wilensky chose to use "causal nodes" as the unit of analysis in their article and provided validations in §§38–44 regarding why causal nodes are viable knowledge pieces. This validation is essential in making the claim about their machine learning method convincing. The validation could be further strengthened if the studies had a larger sample size, so that the statistical significance of the between-people/within-people consistency and percentages could be calculated. Currently, as mentioned in §16, only 41 students responded to both the pre- and the post-intervention questions. The appearance of a code in 10% of the responses would mean that only four students mentioned it.

Causality in complex systems

« 6 » Hjorth and Wilensky analyzed students' conceptual change in the context of urban planning and city growth. Paul Krugman (1996) used cities as an example to illustrate the phenomenon of self-organizing – a hallmark behavior of complex systems. The term "complex system" does not only mean that the system is complex, it refers to a class of scientific approaches that complexity scientists use to investigate phenomena that are difficult to understand using traditional scientific methods. Through a complex systems lens, natural and social scientists see systems as composed of numerous elements that follow simple rules. The behavior of a complex system at the macro level emerges from the interactions among its constituting elements at the micro level (Epstein 1999; Wilensky & Resnick 1999). These parts (or agents) obey very simple rules. The system-level patterns at a very large scale can be strikingly different from the behavior of the constituting elements at the local level, making these complex systems unintuitive and hard for people to understand (Wilensky & Resnick 1999; Chi 2005; Penner 2000). The emergence of larger patterns from agents' simple local behavior appears disconnected and is usually surprising, precisely because it does not seem to follow the causality that people are familiar with, or the types used in traditional scientific methods.

« 7 » Usually, causality is explained by providing mechanisms. Mechanistic explanations require the identification of entities and activities that are involved in the phenomena, and the causal relationships among these entities and activities that are "productive of regular changes from start or set-up to finish or termination conditions" (Machamer, Darden & Craver 2000: 3). Peter Machamer, Lindley Darden and Carl Craver explain that mechanisms are composed of both entities – things that engage in activities (with their properties, which are required for certain activities to happen) and activities – the producers of change. The mechanistic explanations seem to work better with "clockwork" systems, in which causalities are more easily identified, rather than in complex systems, in which the emergent nature of larger patterns seems to be disconnected from the behavior of their constituting parts.

« 8 » In §14, Hjorth and Wilensky introduced the learning activity as design, grow, and results. Due to the emergent nature of complex systems, it would be difficult to reason out what city design had caused which emergent results, because the growing phase is a generative process that may not follow a defined path. In other words, if we could go back in time and regrow Chicago by giving it exactly the same initial design, would it grow into the same city as it is today? Because of the random factors involved in the growing phase, the city might go down a very different path and become an extremely different one. Therefore, it could be difficult to judge how much the students' design assuming direct causalities could explain the regrown city's properties.

Computational methods for knowledge analysis

« 9 » As described in §7 of the target article, the application of data-mining techniques in the field of education has been giving researchers and educators high hopes regarding large-scale data analysis, instant feedback to the learners and teachers, and so forth. However, the validity of computational methods for knowledge analysis seems to be especially difficult to justify, because of difficulties in aligning computational methods with knowledge theories. Sherin (2013) cautioned that machine-learning methods in the field of knowledge analysis should not be used for the sole purpose of reducing the labor of human analysis or scaling up. Instead, such methods should be used to provide convergent evidence, supporting the work of human knowledge analysis. According to Sherin (2013), to achieve these goals, computational analyses need to

- produce interpretable entities in light of chosen knowledge theories,
- identify knowledge elements at an appropriate grain size and level of abstraction,
- identify configurations of elements of the sort identified by the chosen knowledge theories, and
- capture the dynamics of knowledge change.

« 10 » In Hjorth and Wilensky's article, the Association-Rule-Mining method made use of causal nodes that were identified and coded by human coders. This method extended the human coders' abilities to discover relationships between codes and provided converging evidence of students' conceptual change. The pre-coded data produced easy-to-interpret results but constrains the scale of the data that can be fed into machine-learning methods, because the human coding process can be laborious and time-consuming. If processing a large amount of data remains a goal of using machine-learning methods for knowledge analysis, more work needs to be done to better address the four points that Sherin (2013) mentioned.

References

Carey S. (1988) Reorganization of knowledge in the course of acquisition. In: Strauss S. (ed.) Ontogeny, phylogeny, and historical development. Ablex, Norwood NJ: 1–27.

Chi M. T. (2005) Commonsense conceptions of emergent processes: Why some misconceptions are robust. The Journal of the Learning Sciences 14(2): 161–199.

diSessa A. A. & Sherin B. L. (1998) What changes in conceptual change? International Journal of Science Education 20(10): 1115–1191.

Epstein J. M. (1999) Agent-based computational models and generative social science. Chapter 1 in: Generative social science: Studies in agent-based computational modeling. Princeton University Press, Princeton: 4–46.

Krugman P. (1996) The self-organizing economy. Wiley-Blackwell, Hoboken NJ.

Machamer P., Darden L. & Craver C. F. (2000) Thinking about mechanisms. Philosophy of Science 67(1): 1–25.

Penner D. E. (2000) Explaining systems: Investigating middle school students' understanding of emergent phenomena. Journal of Research in Science Teaching 37(8): 784–806.

Sherin B. (2006) Common sense clarified: The role of intuitive knowledge in physics problem solving. Journal of Research in Science Teaching 43(6): 535–555.

Sherin B. (2013) A computational study of commonsense science: an exploration in the automated analysis of clinical interview data. Journal of the Learning Sciences 22(4): 600–638.

Smith J. P., diSessa A. A. & Roschelle J. (1993) Misconceptions reconceived: A constructivist analysis of knowledge in transition. Journal of the Learning Sciences 3: 115–163.

Smith C. L., Solomon G. E. A. & Carey S. (2005) Never getting to zero: Elementary school students' understanding of the infinite divisibility of number and matter. Cognitive Psychology 51(2): 101–140.

Vosniadou S. & Brewer W. F. (1992) Mental models of the earth: A study of conceptual change in childhood. Cognitive Psychology 24(4): 535–585.

Wilensky U. & Resnick M. (1999) Thinking in levels: A dynamic systems approach to making sense of the world. Journal of Science Education and Technology 8(1): 3–19.

Yu (Bryan) Guo's research interest is in using complex-systems approaches and computer-based models to help people understand complex and controversial social phenomena. His dissertation investigates high-school students' learning processes and outcomes about wealth inequality in "Mind the Gap" – a computer-based participatory simulation curricular unit that he designed. He received his PhD in learning sciences from Northwestern University.

Received: 28 May 2019
Accepted: 10 June 2019

Authors' Response

New Questions About New Methods in Old Contexts

Arthur Hjorth
Aarhus University, Denmark
arthur/at/cs.au.dk

Uri Wilensky
Northwestern University, USA
uri/at/northwestern.edu

> Abstract • Designing, implementing and assessing the effects of classroom-based learning experiences spans across a wide variety of methodological, epistemological and design-related issues. Additionally, the use of data mining and computational methods for supporting qualitative data analyses is still new to the field. Potentially because of this, we received ten good, but quite different commentary questions, which we have organized under five headings. In this response, we address each of them to provide a more thorough background or reasoning behind the decisions we made in our target article.

« 1 » We wish to thank the four commentators, who asked a total of ten interesting and thought-provoking questions about our article. In our response, we have categorized these questions under five headings.

Choice of theory and approach

« 2 » **Jim Clayson** (Q1) raises a good question that we may have assumed the answer to, without making a case for it in our target article, i.e., why Knowledge-in-Pieces (KiP) can be applied in the social sciences as well as in math and physics.

« 3 » First, a bit of context to this question. Knowledge-in-pieces originated with Andrea diSessa's work on physics students' reasoning at college level. Andrea diSesssa, and others, haven taken KiP approaches to understanding how students reason about a wide variety of phenomena related to physics, such as gravity, resistance, movement and force, proper time, the shape of the Earth, Earth's four seasons, and the springiness of a spring (Andrea diSessa 1993 2002; Kapon & diSessa 2012; Levrini & diSessa 2008; Sherin, Krakowski & Lee 2012; Vosniadou & Brewer 1992). KiP has shown itself to be a powerfully generative approach to documenting and understanding students' knowledge of and conceptual change in physics. Of course, it does not necessarily follow that this framework would work outside of physics.

« 4 » However, more recently, researchers have started applying KiP to reasoning about non-STEM-related phenomena. Thomas Philip took a KiP approach to understanding the ideology underlying teachers' reasoning about students and race in his work on what he calls *ideology-in-pieces* (Philip 2011). Last year, Arthur Hjorth, one of the authors, participated in a symposium at the International Conference of the Learning Sciences 2018 called "Knowledge Analysis Outside the STEM Classroom" (Anderson et al. 2018). In the symposium, researchers presented their work on using Knowledge Analysis, which posits itself as the methodological approach associated with KiP (diSessa, Sherin & Levin 2016), to study reasoning about social issues ranging from restorative justice, and teachers' reasoning about pedagogy, to ethical sensemaking in engineering. While much of this work is still exploring the potential limits of using KiP outside of STEM, at this point a reasonable fit has been established by this emerging body of literature.

Data analysis and analytical methods

« 5 » **Clayson** emphasizes our statement that we used "a pragmatic and somewhat promiscuous approach to identifying the individual knowledge pieces" (§30), and asks us to explain this, and give a better account of our analytical approach and methods, and why we chose to focus our attention on student responses to the question that we did focus on, and not on one of the four other questions.

« 6 » First, a brief addendum to the phrase that **Clayson** quotes, and which may have not communicated what we intended. A central contribution of diSessa's work with KiP was the identification of *phenemonological primitives* (or "p-prims") that guide students' intuitive sensemaking of physics. The full definition of p-prims is long to recount here, but two aspects are important to address: First, their origin matters. They must have come about through the *everyday interactions* that people have with the physical world. Second, their size matters. They must be the *minimal* parts of an explanation at which students bottom-out, and for which there is no answer to "why do you think that?" When we say that our identification of causal nodes was promiscuous, we mean that we did not discriminate by origin or by size of the apparent underlying causal reasoning knowledge piece.

« 7 » Our primary goal during the qualitative coding phase of our analysis therefore aimed to identify common causal claims across student responses and to uncover patterns in how they are co-activated. As mentioned in the target article (§28), we asked students five different questions. The first three focussed on the relationship between commute time and income, the fourth asked students to reason about income and access to parks and leisure areas, and the fifth asked students to give examples of how one can use urban planning to improve on a city (but left the definition of "improvement" open). As we also mentioned (§16), some students did not answer the last questions due to technical problems, so we chose to focus on the questions for which we had the most