A Visual Approach towards Knowledge Engineering and Understanding How Students Learn in Complex Environments

Lauren Fratamico, Sarah Perez, Ido Roll
University of British Columbia
Online Learning Environments are Complex

- Diversity of instructional activities
- Variety of learners, goals, and engagement patterns
- Some support user-driven exploration (vs pre defined learning trajectories)

Would like to be able to use log data to:
- Interpret student actions
- Label student interaction patterns
- Infer intentions
- Assess learning
- Evaluate quality of engagement
For Example, PhET DC Circuit Construction Kit

- Exploratory learning environment
- Hundreds of actions available
- CCK used 4 million times/year
- Translated into 60 languages

- How to use log data to evaluate students when the design space is unlimited and the solution space is underdefined?
- How can we account for learners with diverse backgrounds and goals?
For Example, PhET DC Circuit Construction Kit

- Exploratory learning environment
- Hundreds of actions available
- CCK used 4 million times/year
- Translated into 60 languages

- How to use log data to evaluate students when the design space is unlimited and the solution space is underdefined?
- How can we account for learners with diverse backgrounds and goals?
Current approaches to make sense of student log data

- **Knowledge engineering (top-down)**
  - Based on expert analysis
  - More challenging as event space grows
- **Knowledge discovery (bottom-up)**
  - Extract patterns from data via machine learning and statistical approaches
  - Often hard to interpret the detected trends and inform theory (Aleven 2016, Roll 2005)
  - Effective for skills that are easy to label, but less for divergent strategies (Baker 2013, Sao Pedro 2013)
  - The detected models may be overly specific to context and populations (Conati 2015)

- Ideal to combine the two
Goal of Research

- Create a hybrid approach that combines:
  - Data-driven, bottom-up insights
  - Human-initiated, top-down understanding
- Allow others to easily interpret their own log data through exploratory analysis
Goal of Research

Create a visual approach that:
- Highlights potential patterns of related actions (using data)
- Helps its users raise hypotheses about these actions (when combined with their knowledge of theory)
- Allows them to quickly test their hypotheses in an exploratory way by:
  - Grouping actions
  - Visualizing the relationship between behaviors and other student-level factors (such as knowledge level)
tempr

a visualization tool for exploratory analysis of temporal log data
What does tempr allow you to do?

Identify features in your log data that:
- Differentiate groups of learners temporarily
- Abstract beyond surface differences
- Are informative with respect to common learning strategies
3 main panels of tempr

- Heatmap
- Merging
- Visualization

bit.ly/tempr_tool
Overview
Exploring Log Data with tempr
Exploring Log Data with tempr

**Data:** PhET Circuit Construction Kit log data

**Groups:** High Learners and Low Learners

**Question:** How do students learn by testing circuits?
Data Input

what item

what action
What does tempr allow you to do? - Heatmap Panel
What does tempr allow you to do? - Heatmap Panel
Surface big picture patterns ~ Compare groups of learners
Hypothesis: HL and LL differ in use of ammeter and voltmeter
What does temp allow you to do? - Merging Panel
What does tempr allow you to do? - Merging Panel

Supports exploratory grouping of related actions
What does temp allow you to do? - Merging Panel
Supports exploratory grouping of related actions
Quickly test hypothesis by merging actions
What does tempr allow you to do? - Visualization Panel

Theory

Data

Hypothesis Raising

Hypothesis Testing

Grouping Actions

Visualizing Actions Temporaly
What does tempr allow you to do? - Visualization Panel

Visualize learning over time - Compare groups of learners - Contrast actions

Testing - User dragging nonContactAmmeter:
user.nonContactAmmeter.drag + user.nonContactAmmeter.endDrag + user.nonContactAmmeter.startDrag

Percent of Log Events

Median

75th Percentile

25th Percentile

Percent of Interaction

HL

LL
HL increase use of ammeter over time

HL decrease use of voltmeter over time
HL test slightly more than LL, fairly constant usage over time

Splitting behaviors into subtypes can reveal important nuances
Conclusion

Tempr aids in temporal analysis of log data by:

- Theory
- Data
- Hypothesis Raising
- Hypothesis Testing
- Grouping Actions
- Visualizing Actions Temporally

Download tempr from GitHub and try it out! bit.ly/tempr_tool