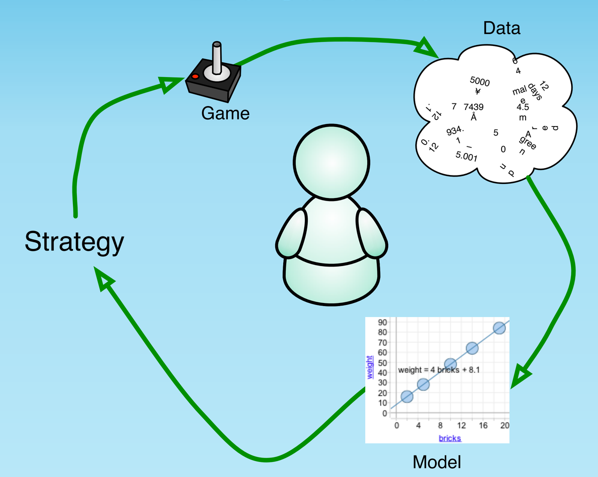
Datagames

Tools and Materials for Learning Data Modeling

Students improve their gaming by becoming data analysts of their own play.

Fall 2009–Fall 2012

The Core Idea



Students playing computer games generate large quantities of rich, interesting, highly variable data that mostly evaporates into the ether when the game ends. What if in a classroom setting, data from games students played remained accessible to them for analysis? In software and curriculum materials developed by this project, data generated by students playing computer games form the raw material for mathematics classroom activities. Students play a short computer game, analyze the game data, conjecture improved strategies, and test their strategies in another round of the game.

The Essence of a Data Game

The play of a data game consists of brief, suspenseful moves that generate data. The move parameters controlled by a player appear in the data. Examination of the data reveals patterns that are useful in improving a player's game score. At higher levels, the game requires increased sophistication in data modeling and automation of play.

Project Goals

- Improve students' ability to understand and work with data, with special emphasis on large data sets, data visualization, time series, informal inference, and data structures beyond rows and columns.
- Enrich students' understanding of mathematics through learning experiences based on using data generated by playing computer games.
- Expand research in students' understanding of data and chance and in students' ability to learn mathematical content when it is embedded in data-rich contexts.

Foundations—Prior Work

Fathom—Dynamic data analysis software for secondary and college math and science classrooms.

TinkerPlots—Dynamic data exploration software for middle school mathematics classrooms.

TinkerPlots and Fathom have made powerful data visualization tools available to students throughout the US and in a growing number of other countries. Both development teams have been increasingly focused on the need to get data into the hands of students more quickly, for example through scraping data from a web page whose URL is dropped into a document. TinkerPlots is soon to release a new version that includes simulation capabilities.

Published by Key Curriculum

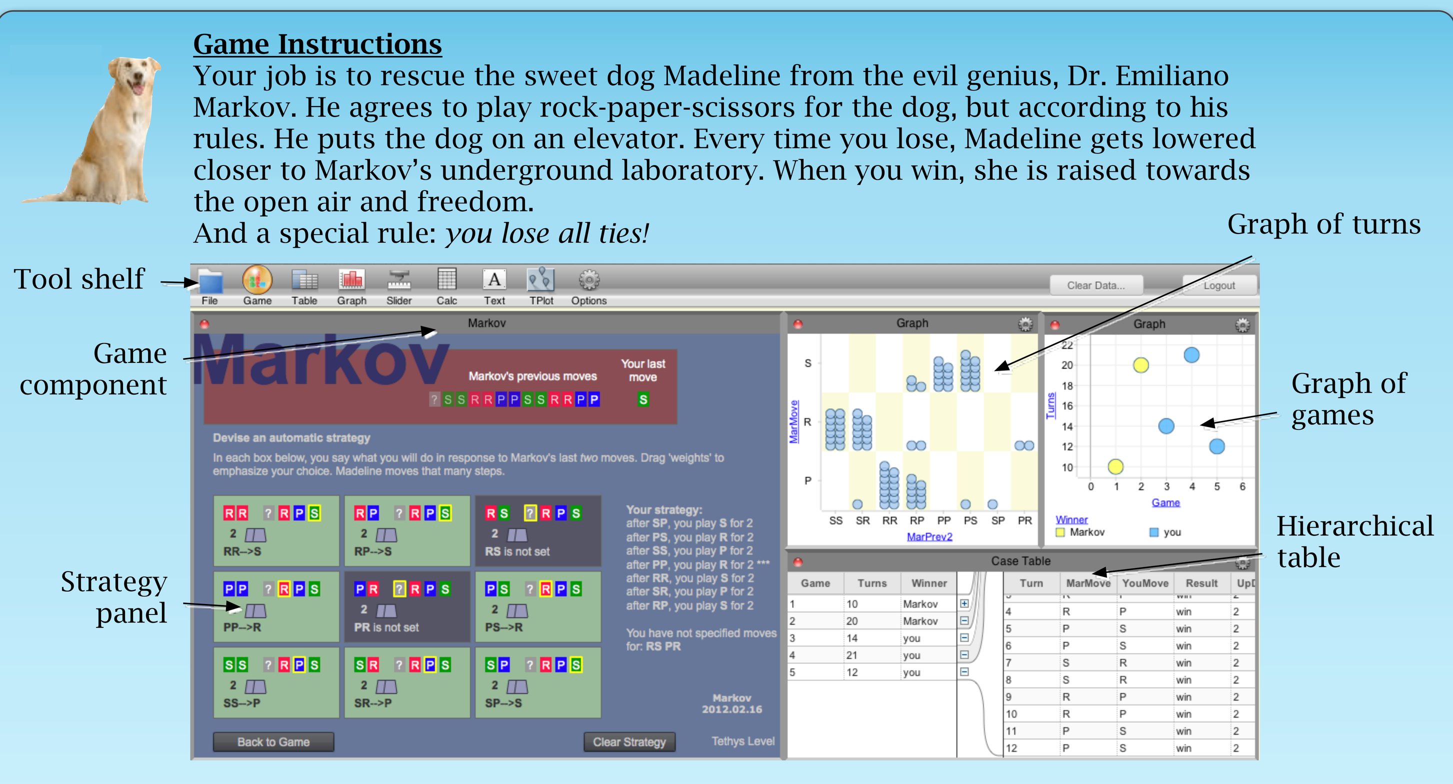
Key Staff Members

William Finzer, co-PI
bfinzer@kcptech.com
Rick Gaston, Research and Project Manager
Kirk Swenson, Software Engineer
KCP Technologies
Emeryville, CA

Cliff Konold, co-PI
konold@srri.umass.edu
Craig Miller, Software Engineer
UMass Amherst
Amherst, MA

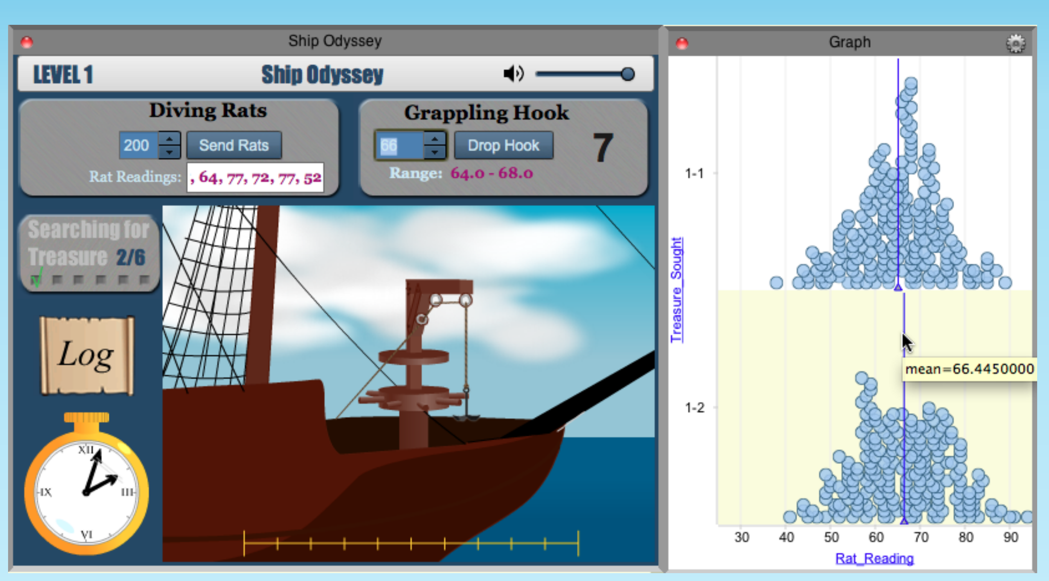
Tim Erickson, Game and Activity Developer,
eepsmedia@gmail.com
Epistemological Engineering
Oakland, CA

Three Example Data Games



After playing the game a few times (and losing), students learn to read the graph of turns, with which they can predict Markov's next move. Then they transfer their understanding to the automated strategy panel shown above. The gray weights sub-panels determine how many steps up or down Madeline makes on a win or loss, and these weights can be moved around to improve the strategy.

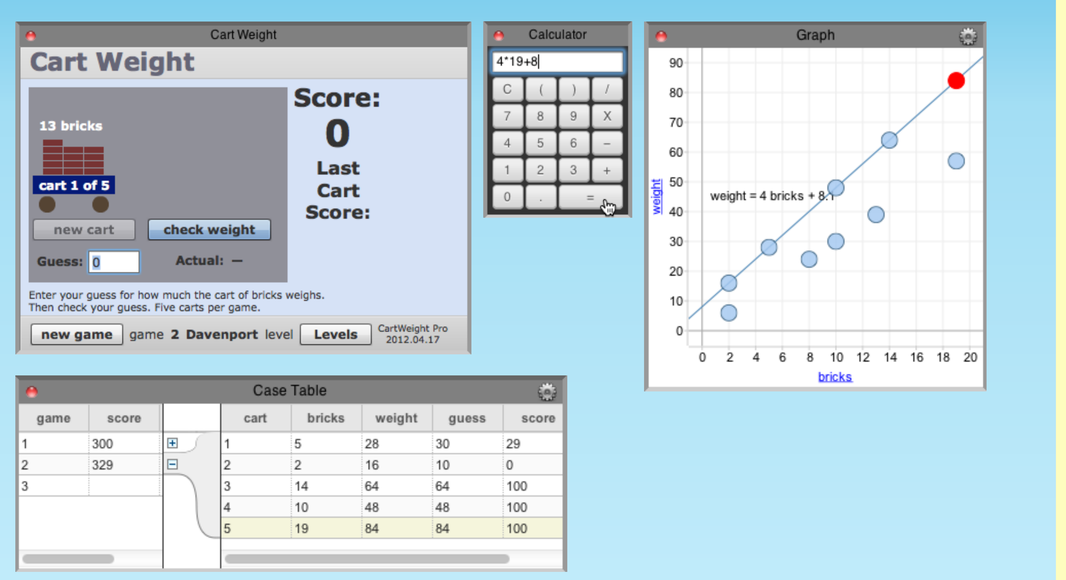
Ship Odyssey

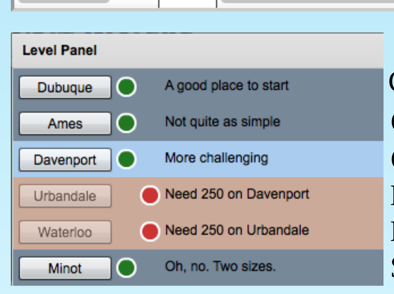


In Ship Odyssey, the goal is to locate sunken treasure. To help locate it, players send down rats to the ocean floor, but the estimates the rats return with are "noisy."

To get high scores, players must first figure out that averages of the noisy data (e.g., the mean) give them more reliable estimates than informal eyeballing of the distribution (e.g., the "modal clump"). Secondly, to minimize the number of rats sent down (rats cost money), they need to develop a sense of the relation between a sample's size and the sample-to-sample variability of its mean (the Law of Large Numbers).

Cart Weight





Cart weighs 0. Bricks weigh the same each game.
Cart non-0 weight. Same brick weight each game.
Cart and brick weight change each game.
Non-integer cart and brick weights.
Noise!
Some bricks are small, some are large.

Cart Weight is the simplest data game, basically an in-out machine with a little bit of context.

To move from one level to the next higher level you need 250 points. This is easy enough on the first two level where the rule does not change from one game to the next. The challenge begins on Davenport, the third level where you have to use the first two points to score well on the remaining three.

Perhaps the most important outcome of this activity is a strengthened belief in the usefulness of graphical tools.

Observations and Conjectures

Rough Stages of Student Involvement with a Data Game

1. Exploration—How does this game work?
2. Intuitive play and practice (no need for data)
3. Keeping track
4. Searching for an algorithm
5. Making sense of graphs
6. Model building—moving beyond algorithm
7. Comparing how well different strategies work

Data Games Design Principles

- Data analysis has to be **useful** in the game, so that analysis really improves your performance.
- Furthermore, data analysis has to be **easy**. If it's hard, it won't be worth the effort just to do well in the game.
- Practice has to be **not useful**. Otherwise some students will avoid thinking.
- The games themselves have to be **quick**. We're in a classroom, after all; whatever we do—the play and the data analysis—has to happen in minutes, not hours of play.
- The games have to **integrate easily** into classroom life.

Conjectures

Motivation—Even very simple games, without fancy graphics or interfaces, can dramatically improve students' engagement with mathematical problem solving. (One student exclaimed, "I *like* math! It helps me *win*!")

Integrating data analysis—Significant opportunities exist for combining data analysis objectives with mathematics objectives using data-games-style activities.

Hierarchical data structures—The intrinsically hierarchical nature of data from playing a series of games provides a meaningful opportunity for working with other-than-flat data structures without sacrificing other learning objectives.

Strategy automation—The requirement that students automate their game play can help them to articulate their mathematical understanding.

Efficacy—Data games activities are effective vehicles for increasing student understanding of math and data concepts.

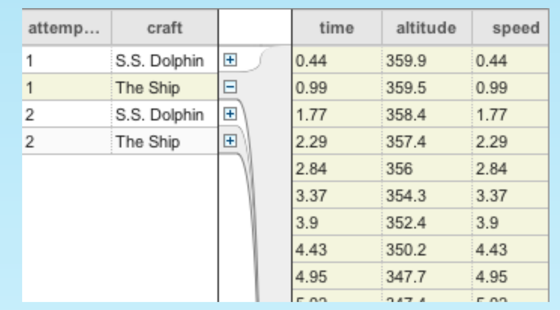
Some Software Development Highlights

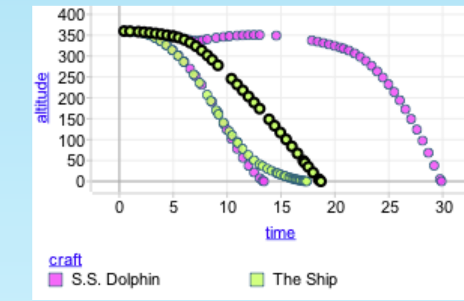
Design of an Interface for Working with Hierarchically Structured Data

Problem—Games create hierarchically structured data: a collection of games, each of which has a collection of moves. How can users easily visualize and interact with the these data?

Steps toward a solution

Nested tables—Make the structure transparent in tables that allow selective disclosure of nesting.

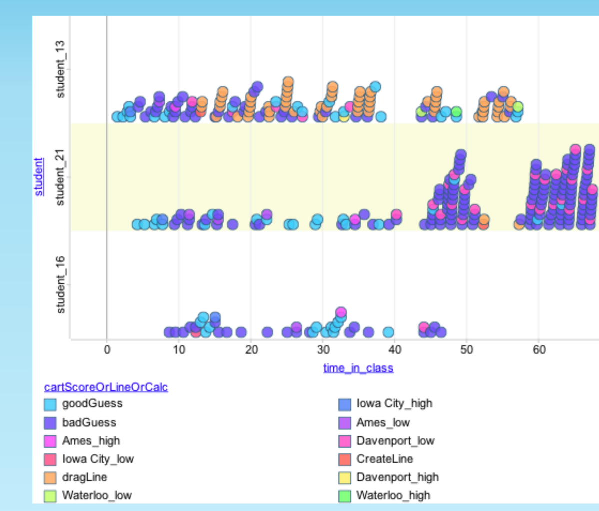




Linked selection—Select a "parent" case and all its "children" get selected too.

Mixing levels in graphs—The number of points changes accordingly.

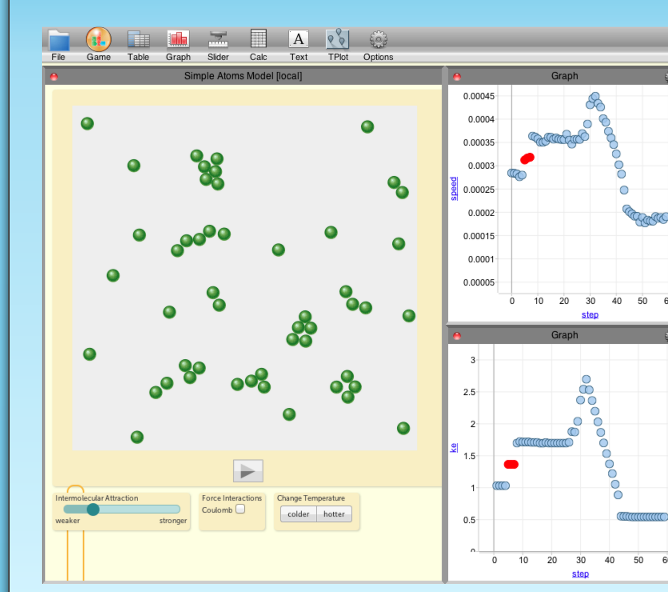
Learning Analytics



The Data Games environment logs student actions to a server. We are developing techniques for determining student progress in an activity and for assessing performance-based understanding.

The graph shows three students playing Cart Weight. Student_13 succeeded in getting a high score on all levels of the game and used a movable line as an algebraic tool. In contrast student_21, though persistent, got stuck on level 3, and, along with student_16, never made good use of the movable line.

Collaboration with InquirySpace



We have begun a collaboration with the Concord Consortium in which games are replaced with science components to become part of an InquirySpace environment.

Acknowledgments

The Data Games project wishes to thank Doug Page and his students at Galileo Academy of Science and Technology, Josh Zucker, Lick–Wilmerding High school and project staff at KCP Technologies and UMass Amherst, for their valuable support of this research work.

This material is based upon work supported by the National Science Foundation under:
KCP Technologies Award ID: 0918735
UMass Amherst Award ID: 0918653
Grant period: September 1, 2009 through August 31, 2012

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.