Tuning Perceptions and Inferences to See a Graph in a New Way

Hillary Swanson, hillary.swanson@usu.edu, Utah State University
Bruce Sherin, bsherin@northwestern.edu; Uri Wilensky, uri@northwestern.edu, Northwestern University

Abstract: This work examines the representational and social features of a computational modeling activity to understand the role they play in helping a student make sense of a simulation graph. Specifically, it examines shifts in the student's thinking as she reflects on the output of a computational model of thermal equilibration, which she has built. By analyzing the student's trajectory through the lens of Coordination Class theory, the paper illuminates how she comes to see the simulation graph in a more sophisticated way through her interactions with the technology and the interviewer. The paper shows how these features of the computational modeling activity support her transition from seeing oscillations of temperature about the mean as a visual effect caused by the computer screen, to understanding that oscillation amplitudes are related to microscopic fluctuations in temperature, which decrease as the number of particles in the system increases.

Introduction

Today, many scientists use computational models to represent and make sense of scientific phenomena. Science education research has investigated student engagement in computational modeling. Some research programs have focused on the design of environments that make the construction of computational models more accessible to younger learners and studied how these environments support learning (Bain, Anton, Horn & Wilensky, 2020). The present paper focuses on student learning in a block-based modeling environment, seeking to characterize, at a fine grain-size, how exploring a computational model of thermal equilibration helps one student find new meaning in a graph of temperatures.

Theoretical foundations

This study is informed by the Knowledge in Pieces (KiP) perspective, which views an individual's knowledge as a complex system of many different kinds of elements and structures (diSessa, 1993). One structure discussed in the KiP literature is a *coordination class* (diSessa & Sherin, 1998). Coordination classes are used to obtain measurable information about the world and are composed of both perceptual and inferential components. The perceptual component consists of *readout strategies*, which the individual uses to *extract* data from the world. The inferential component consists of an *inferential ne*t, which the individual uses to decide what data to extract, and to *draw inferences* from the extracted data. For example, suppose an individual wants to determine which of two objects is being pulled toward Earth with greater force: a bowling ball in their left hand, or a marble in their right. In this situation, the individual draws on their readout strategies to perceive the relative weights of each object. They draw on their inferential net to determine that the bowling ball feels heavier, and therefore must be experiencing a greater Earthward gravitational force. The individual's readout strategies are related to their sense of touch and their inferential net includes ideas about heaviness and force. This paper analyzes the development of one student's understanding of a graph through the lens of Coordination Class theory. Though the theory was originally developed to understand how individuals obtain information about measurable quantities (such as force or speed), a number of studies have borrowed the theoretical machinery to make sense of how individuals see other kinds of information in the world. This work uses readout strategies and inferential nets to characterize shifts in the focal student's extractions and inferences, which allow her to understand a graph of thermal equilibration in a new, more sophisticated way.

Methods

This paper analyzes data from a larger study focused on students' engagement in theory building through computational modeling. To lower the threshold for participation, our team designed block-based computational modeling microworlds using NetTango (Horn, Baker, & Wilensky, 2020), an interface to NetLogo (Wilensky, 1999). The present study seeks to characterize, at a fine grain-size, how theory building through computational modeling supports learning and asks the question: *How do the representational and social features of a computational modeling activity help a student make sense of a graph in terms of the law of large numbers?* To address this question, we analyze one student's explorations in a NetTango microworld during a 1.5-hour interview. Data were collected in three forms: video, screen capture, and audio recording. The present report focuses on data from an interview with a student we call Sage (13 years old) who explored the *Thermal Equilibration* modeling microworld, pictured below (Figure 1). The screenshot shows a model that has been built and initialized. The black box is the *world*, which depicts the activity of the agents (green and purple particles) that are programmed to behave according to the rules specified by the model. To the left of the world are parameters the user can adjust to prescribe initial numbers and speeds for green and purple particles. When the user presses the *setup* button, the world populates with particles that match these parameters. When they press *go*, the particles begin to behave according to the *act* procedure, which the user defines with blocks in the field to the far right. Just to the right of the world is a graph, which shows the temperature or average energy for the groups of green and purple particles over time. It is important to note that the graph shows quantities the model can calculate, which could never be measured in reality, as the particles mix over time. The student is asked to imagine that the particles constitute green and purple gasses at different temperatures, which are mixed together in a container, and to model the situation where the two gasses arrive at the same temperature

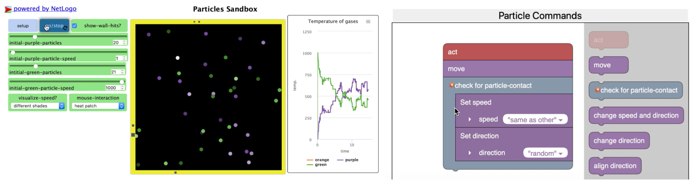

Figure 1. Screenshots of the *Thermal Equilibration* modeling microworld.

An audio recording of the interview with Sage was transcribed. Both screen capture (with audio recording) and transcript were analyzed to look for evidence of learning. A shift in Sage's thinking was identified during an activity toward the end of the session, where she attempted to make sense of the graph produced by her thermal equilibration model. A Knowledge Analysis (diSessa, Sherin, & Levin, 2016) was conducted to understand the smaller steps through which her thinking was refined. The data were analyzed using constructs from Coordination Class theory (readout strategies and inferential net) to understand how her extractions and inferences were refined, leading her to see a graph of thermal equilibration in a more sophisticated way. Her trajectory was divided into five episodes, which were analyzed to identify how her extractions and inferences, which determined what she could see in the graph, changed in relation to her interaction with the technology and the interviewer.

Findings and Discussion

Episode 1: Noticing the behavior of the graph

Sage had been tinkering with parameter settings and her model produced a plot (Figure 2, right) that resembled a graph she had drawn earlier to illustrate hot and cold liquids equilibrating over time (Figure 2, left). The episode opened with Sage commenting on the similarity between the NetTango plot and her graph.

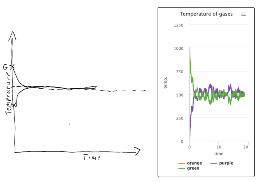

Figure 2. The plot on the left was drawn by Sage, the one on the right was produced by the microworld.

Sage: It's sort of looking like that graph [points at her drawing]. Well now it's like inverted. Well it's like correcting itself. So, I think eventually it will get to align or, well no, maybe.
Int.: What do you think is causing it to go...
Sage: Because I, I don't know. Every time I'm like, "Oh look, it's going smooth." They just like, just like straight up. It's like, um, it, they're like swapping each time.

In this episode, Sage extracts information from the graph. She first notices that it has the same global shape as the graph she sketched. As time goes on, she notices that the curves overshoot the mean and appear to swap places. Her extractions, which move from global shape to a finer-grained detail, are dictated by her readout strategies as well as her inferential net, which dictates what of the available information appears to her to be relevant to making sense of the graph. The particular readout strategies and inferential net she brings to her initial observation arise spontaneously in response to the activity. Sage's extraction of the fluctuation about the mean was afforded by the modeling environment, which plotted the separate temperatures of the two gasses, despite their having mixed. In reality, such a graph, and therefore such an extraction, would be impossible.

Episode 2: Focusing on the aggregate-level behavior

Hoping to capitalize on the problem Sage has stumbled upon, the interviewer asks her to explain the phenomenon.

Sage: Maybe it's being set to random and then they're going the inverse of each other. So, if one is slowing down, then the other one is speeding up. Then the average will always be the same.
Int.: What is the average…? Temperature is going to be the measure of the average speeds…
Sage: So, the temperature. For green, the speed, it looks like it's centering like right around 500, which is good 'cause that's the middle. So, it should be just a little bit of both. So, it's like if, if the 500 was zero then it would be 500 and negative 500, then it would be like the inverse.

In this episode, Sage extracts more detailed information about the plot, observing that the oscillation of the curves is symmetric. She also draws an inference from the symmetric oscillations, determining that they represent the particle speeds going faster and slower in a mirrored way. The interviewer asks her about the temperature, or average speed of the particles. In response, Sage extracts quantitative information of the mean value about which the two curves oscillate. She infers that this value is sensible, given the values of the two particle groups she started with. She further extracts that the curves are oscillating about the mean value by the same amount. Importantly, Sage is extracting information and drawing inferences in response to the interviewer's questions.

Episode 3: Focusing on an agent-level mechanism

The interviewer asks Sage to explain the "jagged" nature of the curves.

Sage: Well, at the beginning I thought it was because this just happens over a long time. It was just pixels and if you zoomed out, it wouldn't be jagged, and it was just 'cause of the computer screen. But now that it's going up and down, I think that it's just because, if one bounces into a wall and gets a little bit slower, then the other one just has to get a little bit faster.

The interviewer's question prompts Sage to share two inferences. The first concerns her initial thinking, which was that when a graph exhibits oscillation it is an effect of the computer visualization. The second concerns her new thinking, which is that the phenomenon is real, and is caused by particles speeding up and slowing down. It appears the interviewer's question has activated these two inferential nets, which Sage uses to make sense of the phenomenon. Sage's extractions and inferences appear to occur in response to the interviewer's questions.

Episode 4: Comparing systems with different particle numbers

The interviewer is interested in helping Sage see that the amplitude of the oscillations decreases as the number of particles in the system increases. Towards this, she asks Sage to compare different numbers of particles.

Int.: So, it's really jagged around 500. Let's see if you raise the number of particles, if it's the same. Do you think you'll get the same jagged line around 500, but up 250 or down 250?
Sage: Um, uh, maybe, I think. Yeah. [increases each group of particles to 50].
Int.: Well it's definitely going fast towards 500. And then it's a little bit slow and then it's doing the thing you saw, which is a little bit of this dance around the 500.
Sage: Yeah. But it's a lot closer now. Like, like it's not going up.
Int.: Okay. What about if you made the particles even more -
Sage: Um, go to like a hundred [slides each group of particles to 100]. Um, I think maybe it will be, it might get all the way flat, but it might just be very close.

The interviewer's suggestion to compare the model with different numbers of particles leads Sage to run the two different models and observe the difference in amplitude heights. She extracts the amplitude information. She then uses her inferential net to compare the two different heights and establish that the one with more particles has the smaller amplitude. This prompts her to predict that 100 should flatten the amplitudes to zero, or close to zero. Here, as in the previous episodes, Sage's extractions and inferences appear to be influenced in important ways by her interactions with the interviewer. They are also afforded by the modeling environment, which allows her to run the simulation with different numbers of particles and compare the amplitudes of the temperature fluctuations. Sage appears to have recognized a relationship relevant to understanding the law of large numbers: the larger the number of particles in the system, the less any single particle will influence the average.

Episode 5: Connecting agent and aggregate levels

The interviewer asks Sage what happens to the curve as she adds particles.

Sage: Um, it's still not like quite a curve. As the number of particles increase, as the sample size increases, then the actions of one particle don't affect the big total as much.
Int.: That is such a cool thing to have said. What do you mean by that?
Sage: So, if you're doing a study like this, if you just had me, it wouldn't really be like a full and complete study because I am not the same as every person and if you have like a hundred people it will be way more accurate because a hundred people is closer to the diversity of knowledge that our population has than one person. And if you had every seventh and eighth grader do it, it would be a hundred percent accurate because everyone was represented.

Sage uses her readout strategies to extract the shape of the curve and describes it as not perfectly smooth. She activates her inferential net and offers an explanation for why increasing the number of particles smooths out the curve. Her inferential net includes information about sample size and how a sample more closely characterizes a population as it gets closer to the size of the population. Much of her thinking is left unarticulated, however, it appears that she has activated prior knowledge related to the law of large numbers, in that she understands the distance of deviations from the mean increases for smaller samples, since individuals within the population have a greater weight within the average. Her activation of this knowledge appears to be in response to her inference about the relationship between a large number of particles and a lower amplitude of fluctuation, as compared to a small number of particles and a higher amplitude of fluctuation about the mean. She drew this inference as a result of comparing simulation runs with different particle numbers, which she did at the request of the interviewer.

Summary and Conclusion

The interviewer asked Sage why the curve was jagged. Sage shifted from thinking it was an effect that would resolve itself (Ep. 1), to describing what was happening at the aggregate level (Ep. 2), to describing what was happening at the agent level (Ep. 3). The interviewer then asked Sage to compare different numbers of particles (Ep. 4). This caused Sage to infer the relationship: the more particles, the lower the amplitude of fluctuation. The interviewer then asked Sage what she noticed about the relationship between particle number and curve (Ep.5). Sage offered an analogy to explain her understanding, drawing on ideas that suggest intuitive resources for understanding the law of large numbers. This sequence illustrates Sage's trajectory from having an incorrect explanation for the graph, to a somewhat scientifically sophisticated explanation. The mediating influence of the technology and interviewer appear significant. The technology allowed her to see invisible fluctuations in average speed and the average temperatures of groups of particles that mixed together as soon as the simulation began. Though Sage initiated the discussion, the interviewer continued to focus her attention on the details of the curve, asking her what might be causing the fluctuations about the mean, and finally, how increasing the number of particles influenced the size of the fluctuations. Some of her questions activated Sage's readout strategies and allowed her to extract new information, while others activated Sage's inferential net, which she used to explain what might be causing the fluctuations and why increasing the number of particles decreased the size of the fluctuations. The interviewer's questions played a role in changing the composition of Sage's readout strategies and inferential net, which she used to make sense of the graph. Shifts in activation of the perceptual and inferential components of her knowledge system helped her make sense of the graph in a new and more sophisticated way.

References

Bain, C., Anton, G., Horn, M., Wilensky, U. (2020). Using blocks-based agent-based modeling for computational activities in STEM classrooms. Paper presented at the 2020 Blocks and Beyond Conference.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and instruction*, *10*(2-3), 105-225.

disessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International journal of science education*, *20*(10), 1155-1191.

diSessa, A. A., Sherin, B., & Levin, M. (2016). Knowledge analysis: An introduction. *A. diSessa, M. Levin, & N. Brown (Eds.), Knowledge and interaction: A synthetic agenda for the learning sciences*, 30-71.

Horn, M.S., Baker, J. & Wilensky, U. (2020). NetTango Web [Computer Software]. Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Wilensky, U. (1999). NetLogo. [Computer Software]. Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Acknowledgements

This work was supported by the National Science Foundation (1842375).