Chapter 4
The Dynamical Analysis of Log Data Within Educational Games

Erica L. Snow, Laura K. Allen, and Danielle S. McNamara

Abstract Games and game-based environments frequently provide users multiple trajectories and paths. Thus, users often have to make decisions about how to interact and behave during the learning task. These decisions are often captured through the use of log data, which can provide a wealth of information concerning students' choices, agency, and performance while engaged within a game-based system. However, to analyze these changing data sets, researchers need to use methodologies that focus on quantifying fine-grained patterns as they emerge across time. In this chapter, we will consider how dynamical analysis techniques offer researchers a unique means of visualizing and characterizing nuanced decision and behavior patterns that emerge from students' log data within game-based environments. Specifically, we focus on how three distinct types of dynamical methodologies, Random Walks, Entropy analysis, and Hurst exponents, have been used within the game-based system iSTART-2 as a form of stealth assessment. These dynamical techniques provide researchers a means of unobtrusively assessing how students behave and learn within game-based environments.

Keywords Dynamics • Stealth assessments • Data visualization • Game-based environments

1 Introduction

In this chapter, we discuss how the power of dynamical analyses has the potential to provide researchers with a deeper understanding of students' behaviors within game-based systems and the impact that variations in these behaviors have on learning. The research described in this chapter occurs within the context of iSTART-2 (the Interactive Strategy Training for Active Reading and Thinking-2), an intelligent tutoring system (ITS) designed to support the development of adolescent students' reading comprehension skills (Jackson & McNamara, 2013; Snow,

E.L. Snow • L.K. Allen • D.S. McNamara
Arizona State University, Tempe, AZ 85281, USA
e-mail: erica.l.snow@asu.edu; laurakallen@asu.edu; danielle.mcnamara@asu.edu

© Springer International Publishing Switzerland 2015
C.S. Loh et al. (eds.), Serious Games Analytics: Advances in Game-Based Learning, DOI 10.1007/978-3-319-08854-4_4
Jacovina, Allen, Dai, & McNamara, 2014). We first provide a brief overview of the use of log data and dynamical analyses to assess students’ behaviors within game-based environments. Subsequently, we describe iSTART-2 and discuss how log data and dynamical analyses have been used as a means of stealth assessment within the context of our game-based environment.

2 The Utility of Log Data Within Game-Based Environments

Computer-based learning environments increasingly incorporate games and game-based features as a means to enhance students’ engagement during learning and instruction (Gee, 2003; Johnson et al., 2004; McNamara, Jackson, & Graesser, 2010; Rai & Beck, 2012; Sabourin, Shores, Mott, & Lester, 2012). Although these game-based computer systems vary in their design, structure, and content, a common functionality in many of these environments is the element of user choice. Indeed, many games and game-based environments afford users the opportunity to customize their learning experience by providing them with a variety of choices regarding their potential learning paths. These interactive choices can range from avatar personalization to “choose your own adventure” tasks. Accordingly, users are often required to make decisions about how to interact and behave within the game-based system.

When users are afforded the opportunity to exert agency over their learning path, they will most likely vary in their experiences of the game. Indeed, users’ learning trajectories (interaction patterns) vary considerably when they are afforded the opportunity to exert agency within systems (Sabourin et al., 2012; Snow, Jacovina, et al., 2014; Spires, Rowe, Mott, & Lester, 2011). One problem faced by researchers is the analysis and assessment of these interaction patterns, as it can be difficult to quantify the fine-grained changes in users’ behaviors. Recently, however, researchers have turned to a novel form of assessment through the use of the log data generated by these systems. Log data has the potential to capture multiple facets of users’ decisions within games, ranging from keystrokes and mouse clicks to telemetry data. Researchers often intentionally program their game-based environments to log all of a user’s interactions or choices within the system. When utilized appropriately, this data can provide scientists with a wealth of information concerning students’ choices and performance while engaged within game-based systems (Baker et al., 2008; Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007; Sabourin et al., 2012; Snow, Allen, Russell, & McNamara, 2014).

One particular benefit of log-data analyses is that they can act as a form of stealth assessment (Shute, 2011; Shute, Ventura, Bauer, & Zapata-Rivera, 2009). Stealth assessments covertly measure designated constructs (e.g., engagement, cognitive skills) without disrupting the users with explicit tests. In other words, these measures are virtually invisible to users. Log data has previously been used as a form of stealth assessment to measure a multitude of constructs, such as students’ study habits (Hadwin et al., 2007), self-regulation ability (Sabourin et al., 2012), and
gaming behavior (Baker et al., 2008). For instance, Hadwin and colleagues (2007) examined how students varied in their studying patterns within the g study system (i.e., a platform designed to aid in students studying behaviors) and how these variations ultimately relate to self-regulative behaviors. This work revealed that log data from students' time within the system was not only predictive of self-regulation, but also captured behaviors that would be missed by traditional self-report measures.

3 Applying Dynamical Analyses to Log Data

Log data generated from game-based systems has proven to be an invaluable assessment tool for researchers. However, researchers have often struggled with ways to quantify patterns that emerge within this type of system data. Indeed, an important goal going forward is for scientists to devise methods for evaluating and quantifying the variations that manifest within log data. These quantification methods will allow researchers to assess the extent to which behavior patterns can shed light upon students' experiences within game-based environments and how variations in those experiences influence learning outcomes.

Dynamic systems theory and its associated analysis techniques afford researchers a nuanced and fine-grained way to characterize patterns that emerge across time. Dynamic analyses do not treat behaviors or actions as static (i.e., unchanging), as is customary in many statistical approaches, but instead focus on complex and sometimes fluid changes that occur across time. Recently, we have proposed that dynamical systems theory and its associated analysis techniques may be useful for examining behavioral patterns and variations within game-based log data (Snow, Allen, Russell, & McNamara, 2014; Snow, Jacovina, et al., 2014). Current work in this area supports this notion, as dynamical analyses have been successfully applied to log data from adaptive environments to capture the fine-grained behavior patterns enacted by students during various learning tasks (Allen et al., 2014; Hadwin et al., 2007; Snow, Allen, Russell, & McNamara 2014; Snow, Likens, Jackson, & McNamara, 2013; Zhou, 2013). For instance, we have previously used dynamical analyses to classify fluctuations in students' choice patterns within the game-based system iSTART-ME (interactive Strategy Training for Active Reading and Thinking—Motivationally Enhanced; Jackson & McNamara, 2013; Snow, Allen, Russell, & McNamara 2014). These analyses revealed that some students acted in a controlled and decisive manner within the system, whereas others acted more randomly. These behavior classifications would have otherwise been missed without the combination of log data and dynamical analyses.

There are many forms of analysis techniques and methodologies used within dynamical systems theory (Granic & Hollenstein, 2003). The current chapter discusses three of these methodologies (Random Walks, Entropy, and Hurst exponents), which we have used to develop stealth assessments within iSTART-2. First, Random Walks are mathematical tools that generate a spatial representation of a path or pattern that forms within categorical data across time (Benhamou & Bovet,
1989; Lobry, 1996; Snow et al., 2013). This technique has been used in economics (Nelson & Plosser, 1982), ecology (Benhamou & Bovet, 1989), psychology (Allen et al., 2014), and genetics (Lobry, 1996) as a way to visualize changes in patterns over time. Geneticists in particular have used this technique to investigate pairings of genes within gene sequences (Arneodo et al., 1995; Lobry, 1996). Within the context of educational games, this technique can provide a visualization of various learning trajectories or paths within the game. Thus, if students can "choose their own adventure," these tools can provide researchers with a means to track and trace these choices as they manifest across time.

Although Random Walks afford researchers a way to visualize patterns in their data, they do not provide a quantifiable measure of change or fluctuations in those patterns. Thus, other dynamical methodologies, such as entropy and Hurst analyses, can be used in conjunction with Random Walks to quantify these fluctuations and changes across time. Entropy is a dynamical methodology that originated in the field of thermodynamics (Clausius, 1865) and is used to measure the amount of predictability that exists in a system across time (Grossman, 1953). Specifically, Entropy analyses provide a measure of random (unpredictable) and ordered (predictable) processes by calculating how many pieces of information are contained within a system or time series (Grossman, 1953). Thus, the more information that is present within a time series, the more unpredictable or random the entire series is considered. Similar to Random Walks, Entropy has been used across a variety of domains, from thermodynamics (Clausius, 1865) to linguistics (Berger, Pietra, & Pietra, 1996). Within the context of educational games, this methodology provides a quantifiable measure of the changes in students’ behaviors. For instance, if a student makes a variety of different choices within a game, they will produce an Entropy score that contains numerous pieces of information and therefore is indicative of a more unpredictable or random time series. Entropy calculations afford researchers the opportunity to examine the predictability of users’ movements and choices within game-based environments.

Similar to Entropy, Hurst exponents (Hurst, 1951) quantify tendencies of a time series. Hurst exponents act as long-term correlations that characterize statistical fluctuations across time as persistent, random, or antipersistent (Mandelbrot, 1982). Persistent patterns are similar to positive correlations, where fluctuations in patterns are positively correlated from one moment to the next. These patterns reflect self-organized and controlled processes (Van Orden, Holden, & Turvey, 2003). In the context of a game, Hurst exponents may be indicative of a student choosing to do the same action or a set of actions repetitively. By contrast, random patterns are said to be independent, where each moment in the pattern does not influence what comes before or after it. These patterns represent a breakdown in control (e.g., Peng et al., 1995). Random patterns within a game could be indicative of a student exploring the interface in an impetuous manner. Thus, the student does not demonstrate a strategy or plan of action. Finally, antipersistent patterns are similar to negative correlations, where the time series demonstrates a corrective process (Collins & De Luca, 1994). These patterns can manifest if a student demonstrates
reactive behavior, where their next action within a game is in opposition to what they just experienced. Within the context of educational games, Hurst exponents can provide a fine-grained measure of the relationship between behavior changes. Thus, Hurst affords researchers the opportunity to examine the overall tendency of users’ choices within game-based environments. It is important to note the difference between Hurst exponents and Entropy calculations. Hurst exponents capture how each time point (or action) is related to what happens before and after, where correlated actions are considered to be persistent or controlled. Conversely, Entropy provides a quantification of the degree to which the entire time series is predictable versus random.

4 iSTART-2

iSTART (Interactive Strategy Training for Active Reading and Thinking) provides high school students with instruction and practice to use self-explanation and comprehension strategies to understand challenging texts (McNamara, Levinstein, & Boonthum, 2004). It focuses on strategies such as making bridging inferences that link different parts of a text and using prior knowledge to connect the ideas in the text to what the student already knows. When students are provided with instruction to use these strategies, the quality of their explanations improves and their ability to understand challenging texts, such as science texts, is enhanced (McNamara, O’Reilly, Rowe, Boonthum, & Levinstein, 2007; O’Reilly, Sinclair, & McNamara, 2004; Taylor, O’Reilly, Rowe, & McNamara, 2006). iSTART-ME (Jackson & McNamara, 2013) and iSTART-2 (Snow, Allen, Jacovina, & McNamara, 2015; Snow, Jacovina, et al., 2014) are more recent versions of iSTART that provide students with the same comprehension strategy instruction within game-based platforms. These game-based systems were designed to provide adaptive instruction and at the same time enhance students’ motivation and engagement through the inclusion of games and game-based features (Jackson & McNamara, 2013).

Within iSTART-2 (see Fig. 4.1), there are two phases: training and practice. Students first engage in training, where they are introduced to a pedagogical agent (Mr. Evans) who defines and explains self-explanation and comprehension strategies and demonstrates how they can be applied to complex science texts. Students are introduced to five comprehension strategies: comprehension monitoring, predicting, paraphrasing, elaborating, and bridging. Each strategy is first introduced and explained in a video narrated by Mr. Evans. At the end of each video, students are transitioned to a checkpoint, where they are quizzed on their understanding of the strategy they just learned. After students watch the five lesson videos, they watch a final summary video. In this video, Mr. Evans summarizes the five strategies that the students just learned. Once these videos are completed, students watch as Mr. Evans provides demonstrations on how to combine multiple strategies to better understand complex science texts.
After training, students transition to the practice phase of iSTART-2. During this phase, students engage with an interactive game-based interface, where they can freely choose to self-explain science texts, personalize different aspects of the interface, practice identifying self-explanations within the context of mini-games, or view their personal accomplishments in the system (see Fig. 4.2). Within iSTART-2, there are four different types of game-based features: generative practice, identification mini-games, personalizable features, and achievement screens. Generative practice requires students to write their own self-explanations. Within iSTART-2, there are three generative practice environments: Coached Practice, Showdown, and Map Conquest. Coached Practice is a non-game-based method of practice, where students generate self-explanations and then receive feedback from Mr. Evans. Conversely, Showdown and Map Conquest are game-based forms of generative practice. In these games, students generate self-explanations for complex science texts within the context of a game. For example, in Map Conquest, students are asked to generate self-explanations for numerous target sentences. Higher quality self-explanations earn more dice. These dice are then used to conquer neighboring territories (see Fig. 4.3). Students win the game by conquering the most territories; to do this, they must earn a sufficient number of dice by generating high quality self-explanations. Within all three generative practice environments, the quality of students’ self-explanations is assessed through an algorithm that relies on both Latent Semantic Analysis (LSA; Landauer, McNamara, Dennis, & Kintsch, 2007) and word-based measures (McNamara, Boonthum, Levinstein, & Millis, 2007).
Fig. 4.2 A screenshot of the iSTART-2 game-based practice menu

Fig. 4.3 A screenshot of the iSTART-2 generative practice game strategy match
This algorithm scores self-explanations on a scale ranging from 0 to 3, with scores of “0” indicating that the self-explanation is irrelevant and scores of “3” indicating that the self-explanation is relevant, uses prior knowledge, and incorporates information from outside of the text.

Within identification mini-games, students are provided the opportunity to practice identifying the five self-explanation strategies. For instance, in Bridge Builder, students are asked to help a man cross a bridge by building the bridge “brick by brick.” Each brick represents one of the five self-explanation strategies they have learned. Students are first shown a text and a self-explanation; they must then identify the strategy that was used to generate the self-explanation by placing the corresponding brick on the bridge (see Fig. 4.4). This process repeats until students have helped the man cross the bridge. In total, there are five identification mini-games (see Jackson & McNamara, 2013, for a complete description).

Within iSTART-2, students can earn system points by interacting with texts, either within the context of generative games or identification mini-games. As students collect more points within the system, they subsequently progress through a series of 25 achievement levels (ranging from Bookworm to Ultimate Allen Intelligence). For students to progress to a new level, they must earn more points than required for the previous level. This mechanic was designed to ensure that students exert more effort as they progress through higher levels in the system. Students also have the opportunity to win trophies in the generative and identification games. These trophies range from bronze to gold and are awarded based on gameplay performance.

iSTART-2 also builds in non-practice game-based features as a way to engage students’ interest. These elements include personalizable features and achievement
screens. Personalizable features are elements designed to enhance students' feelings of personal investment; they include an editable avatar and changeable background colors. Students can use these elements to customize the system interface. Finally, achievement screens were built into the system to allow students to monitor their progress. Students can use these screens to view their last ten self-explanation scores or any trophies they have won throughout their time in the system.

Overall, the iSTART program has been effective at improving students' use of self-explanations and reading comprehension ability (Jackson & McNamara, 2013). When game-based features are embedded within the iSTART program, students have expressed increased motivation and enjoyment across multiple training sessions (Jackson & McNamara, 2013). Combined, these results suggest that the game-based iSTART system effectively captures users' engagement across multiple training sessions and subsequently improves target skill acquisition.

4.1 iSTART-2 Log Data

Recently, log data from the iSTART programs have been used to develop stealth assessments (Snow, Allen, Russell, & McNamara 2014; Snow, Jacovina, et al., 2014; Snow et al., 2013). This system, like many game-based environments, provides users with agency over their learning paths. Thus, the log data generated from this environment contains a wealth of information regarding variations in students' choices and their influence on learning outcomes.

The log data generated from iSTART-2 contains information about how students interact within the system (choices, time stamps, and language input). For instance, iSTART-2 collects data on every choice a student makes while engaged with the game-based interface. This data provides a detailed list of actions as well as the duration of each action. Table 4.1 provides an example of what this log data looks like. In Table 4.1, there are only five columns (Student ID, Start Time, Stop Time, Action, Complete.

<table>
<thead>
<tr>
<th>Student ID</th>
<th>Start time</th>
<th>Stop time</th>
<th>Action</th>
<th>Complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>004</td>
<td>8:45 am</td>
<td>9:00 am</td>
<td>Bridge Builder</td>
<td>Y</td>
</tr>
<tr>
<td>004</td>
<td>9:01 am</td>
<td>9:12 am</td>
<td>Map Conquest</td>
<td>Y</td>
</tr>
<tr>
<td>004</td>
<td>9:13 am</td>
<td>9:14 am</td>
<td>Avatar Edit</td>
<td>Y</td>
</tr>
<tr>
<td>004</td>
<td>9:14 am</td>
<td>9:16 am</td>
<td>Bridge Builder</td>
<td>N</td>
</tr>
<tr>
<td>004</td>
<td>9:17 am</td>
<td>9:18 am</td>
<td>Achievement Screen</td>
<td>Y</td>
</tr>
<tr>
<td>007</td>
<td>3:00 pm</td>
<td>3:02 pm</td>
<td>Avatar Edits</td>
<td>Y</td>
</tr>
<tr>
<td>007</td>
<td>3:03 pm</td>
<td>3:05 pm</td>
<td>Background Edits</td>
<td>Y</td>
</tr>
<tr>
<td>007</td>
<td>4:25 pm</td>
<td>4:35 pm</td>
<td>Map Conquest</td>
<td>N</td>
</tr>
<tr>
<td>007</td>
<td>4:37 pm</td>
<td>4:45 pm</td>
<td>Showdown</td>
<td>Y</td>
</tr>
<tr>
<td>007</td>
<td>4:47 pm</td>
<td>5:01 pm</td>
<td>Balloon Bust</td>
<td>Y</td>
</tr>
</tbody>
</table>
Action, and Complete); however, log data can be much more detailed, as the researcher often dictates the detail of the log data generated from the system. In this simplified example, there are two students (004 and 007) who have each made five choices within the system. The log data presented here reveals the start and stop time of each choice and whether or not it has been completed. This detailed report affords researchers the opportunity to trace each user’s learning path within the system. It is important to note that these learning paths constantly vary as iSTART-2 affords users with high levels of agency over their learning path (Snow, Jacovina, et al., 2014).

Although iSTART-2 provides detailed descriptions of each student’s interaction path within the system, the log data on its own cannot quantify the variations and fluctuations in behavior patterns that manifest in these data sets. Thus, dynamical analysis techniques are needed to characterize patterns that emerge in this system log data. Because dynamical systems theory treats time as a critical variable, the log data must be first organized chronologically. It is important to note that in order for these methodologies to provide accurate quantifications of users’ behaviors, there needs to be some form of time-based classification for each behavior, along with its association with the other behaviors within the system (i.e., chronological or temporal).

4.2 Dynamical Methodologies and Log Data Within iSTART-2

In the following sections, we describe how log data and dynamical analyses can be combined to better understand students’ system behaviors. We describe how the three dynamical methodologies discussed earlier (Random Walks, Entropy, and Hurst exponents) have been utilized to covertly assess students’ behaviors and the impact of variations in those behaviors on target skill acquisition within iSTART-2. These three techniques provide a novel means of visualizing and categorizing nuances in students’ behavior patterns that emerge within log data across time.

4.2.1 Random Walks

Random Walks can provide researchers with a visualization of how students choose to play or interact within game-based environments. These tools are quite flexible, as the researcher can set the parameters and dimensions represented within the walk. For instance, Random Walks have been created that incorporate multiple vectors (Snow et al., 2013; Snow, Allen, Jackson, & McNamara, 2014) and dimensions (Berg, 1993). Indeed, the number of dimensions that can be included when using random walk analyses is, in theory, unlimited. The Random Walks that have been generated for the log data in iSTART-2 have four orthogonal vectors that lie on an X, Y scatter plot (see Fig. 4.5). Each of these vectors corresponds to one of the four
types of game-based features embedded within the system: generative practice, identification mini-games, personalizable features, and achievement screens.

In general, Random Walks follow a set of basic rules that trace movements across categorical data. These rules are predetermined and must stay consistent throughout the entire Walk analysis. Within iSTART-2, these rules dictate how an imaginary particle moves along the X, Y scatterplot and traces students’ movements (i.e., their choice of interactions) between the four orthogonal vectors (i.e., the game-based features). The rules for the Random Walks generated within iSTART-2 are listed in Table 4.2.

Every Walk begins at the origin point (0, 0). An imaginary particle is placed at the origin and only moves after a student has interacted with one of the four game-based features. Every movement of the particle corresponds to the directional assignment established by the researcher. Figure 4.5 demonstrates how the rules described in Table 4.2 would be applied to a student who has made four interaction choices within iSTART-2. This student’s sequence of choices is as follows: (1) identification mini-game (move up), (2) generative practice game (move right), (3) second identification mini-game (move up), and (4) personalizable feature (move left).
Random Walks have been applied to over 300 students (across multiple studies) within the iSTART-2 system as a way to visualize various learning paths within the game-based interface. Figure 4.6 reveals what an actual Random Walk looks like for a college student who spent approximately 2 h interacting with the iSTART-2 interface and made 38 total interaction choices. This student’s Random Walk provides a visualization of those interactions. From Fig. 4.6, we can see that this student’s Walk moved in an upward direction along the Y-axis. This indicates that the majority of this student’s interactions were with identification mini-games. Indeed, the raw log data reveals that of the 38 total interactions, 22 were with an identification mini-game. Hence, this student’s Random Walk provides a means of visualizing fluctuations in these choice patterns as they manifest across time.

Figure 4.6 shows a Random Walk for one student; however, these tools can also be used to visualize differences in interaction patterns (or choices) comparing groups of individuals (Snow, Allen, Jackson, et al., 2014; Snow et al., 2013). For instance, Snow et al. (2013) used aggregated Random Walks to visualize differences in how high reading ability and low reading ability students engaged with game-based features within the iSTART program (see Fig. 4.7). Using this visualization technique, they took the slope of each student’s random walk (n = 40) and plotted it along the XY axis. A median split on pretest reading ability was used to separate students into groups of high reading ability (green slopes) and low reading ability (blue slopes). Results from this visualization revealed that high ability students tended to gravitate more towards identification mini-games whereas low ability students interacted most frequently with the generative practice games (Fig. 4.7). It is important to note that within this random walk, directionality is used only to visualize students’ interaction preferences. Thus, Fig. 4.7 reveals that high ability students (green lines) are more likely to select identification mini-games compared
Fig. 4.7 Aggregated random walk for high \((n=18)\) and low reading ability \((n=20)\) students (Figure adapted from Snow et al., 2013)

to the low ability students (blue lines). Overall, Random Walks can be used to trace students’ choice patterns within game-based systems. These techniques can be used to track a single student’s progress throughout the game or they can be aggregated to provide a visualization of differences in choice patterns comparing two or more groups of individuals.

4.2.2 Entropy

- While Random Walks offer researchers compelling visualizations of students’ trajectories within game-based systems, these tools cannot, on their own, quantify variations in choice patterns that emerge across time. Entropy can be used in conjunction with Random Walk analyses to provide an overall quantification of students’ interaction patterns. There are many different variations of the Entropy calculation (Bandt & Pompe, 2002; Costa, Goldberger, & Peng, 2002; Shannon, 1951); however, the current chapter focuses on the most widely used Entropy calculation, Shannon Entropy (Shannon, 1951). Equation 4.1 shows the equation for Shannon Entropy. In this equation, \(P(x)\) represents the probability of a given state (or interaction). In the context of iSTART-2, this formula could be used to analyze log data to calculate how ordered students’ choices are across time. Specifically, Entropy for a given student would be calculated by taking the
additive inverse of the sum of products calculated by multiplying the probability of each interaction by the natural log of the probability of that interaction. Thus, Entropy scores reflect the degree to which students' interactions within iSTART-2 are ordered (or random) across time. In general, low Entropy scores are indicative of ordered processes, whereas high Entropy scores suggest disorganized or random processes. Thus, if a student's choice pattern is highly organized, they are likely to produce a low Entropy score. Conversely, when a student's choice pattern is disorganized (i.e., interactions within the system are not systematic), the Entropy score will likely be high. Entropy scores are guided by the bits of information presented within a time series. For instance, let's say we flip a fair coin (even probability of heads and tails) twice. If the coin lands on heads both times, Entropy will be zero. Thus, the flip of the coin resulted in uniformed bits of information. However, if we flip the coin and get one heads and one tails, the Entropy of the flips would be 1.0. This is because the maximum possible Entropy increases as the number of possible outcomes (or choices) increases.

\[ H(x) = -\sum_{i=0}^{N} P(x_i) \left( \log_p P(x_i) \right) \quad (4.1) \]

Within iSTART-2, we have conducted post hoc analyses using log data in combination with Shannon Entropy to assess how much control students exerted over their learning paths. In one study, it was hypothesized that when students demonstrated higher levels of agency, they would also act in a more controlled and organized manner (Snow, Jacovina, et al., 2014). To test this hypothesis, we conducted a single session study where college students \((n=75)\) freely interacted with the iSTART-2 system for two hours. Every choice that the students made was then categorized into one of the four previously mentioned game-based categories (generative practice identification mini-games, personalizable features, and achievement screens). Entropy analyses were conducted at the end of the study on each student's categorized log data to examine the extent to which the interaction patterns reflected ordered or disordered behavior patterns.

Overall, students varied considerably in their ability to act in a controlled and organized fashion (range = 1.32–2.32, \(M=1.83, \text{SD}=0.24\)). Interestingly, results from this study revealed no significant correlation between Entropy scores and the frequency of interactions with any specific feature (i.e., generative practice identification mini-games, personalizable features, and achievement screens). Thus, students' ability to exert controlled interaction patterns was not related to any specific game-based feature. A final analysis examined how variations in students' choice patterns influenced the quality of their self-explanations produced in the generative practice games. A hierarchical regression analysis revealed a significant relation between Entropy and self-explanation quality. Specifically, the students who engaged in more controlled and systematic interaction patterns within iSTART-2 generated higher quality self-explanations than those students who demonstrated disordered behavior patterns.
Entropy analysis has been applied to over 300 high school and college age students' log data (across multiple studies) generated from their time within the iSTART program. This analysis has proven to be a relatively simple way for researchers to examine the overall state of students' choice patterns. Without the use of Entropy, these fine-grained behavior patterns would most likely have been missed. Further, this dynamical methodology can serve as an important stealth assessment when researchers are interested in examining the degree that students' overall behavior patterns are ordered across time.

4.2.3 Hurst

Although Entropy analyses can provide a measure of how ordered students' choices were within a game-based system, this measure does not capture how each choice within the pattern relates to the other choices proceeding and succeeding it. The Hurst exponent has the ability to capture these nuanced fluctuations as they manifest across time (Hurst, 1951). In our recent work, we have calculated Hurst exponents using a Detrend Fluctuations Analysis (DFA). A DFA estimates Hurst exponents by first normalizing the time series (or interaction pattern). Once this data is normalized, it is divided into equal time windows of length, n (which may vary for each student). Every window is then fit with a least squares line and the resulting time series is detrended by subtracting the local trend of the respective window. This is then repeated as the windows increase exponentially by the power of 2. For each window, a characteristic fluctuation $F(n)$ is calculated; this is the average fluctuation as a function of window size. Finally, $\log_2 F(n)$ is regressed onto $\log_2(n)$, the slope of which produces the Hurst exponent, $H$. The resulting Hurst exponent ranges from 0 to 1 and can be interpreted as follows: $0.5 < H \leq 1$ indicates deterministic behavior trends, $H=0.5$ indicates random behavior trends, and $0 \leq H < 0.5$ indicates antipersistent behavior trends.

Within iSTART-2, Hurst exponents have been used in conjunction with log data to examine how fluctuations in students' learning paths influence self-explanations quality (Snow, Allen, Russell, & McNamara 2014). Using this methodology, we were interested in examining how deterministic (and random) patterns of interactions within the game-based environment influenced self-explanations quality (similar to the results from the Entropy analyses). Hurst exponents were calculated for over 80 students (across multiple studies) within the iSTART program. Each of these students spent at least 8 h within the game-based environment and engaged in approximately 275 interactions (i.e., game-based feature choices). Similar to the Entropy analysis, every choice made by students during their time within the system was categorized into one of the four previously mentioned game-based categories and DFA analyses were then calculated using this categorized log data. After the DFA was conducted, each student was assigned a Hurst exponent that quantified the extent to which students' interaction patterns fluctuated in a random or controlled manner.

Results from these analyses revealed that when students engaged in controlled and deterministic patterns of interactions within the game-based system iSTART-2,
Table 4.3 Summary of the benefits and limitations of each methodology

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Benefits</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random walks</td>
<td>Provides visualization of changes in categorical data across time</td>
<td>Cannot quantify variations in choice patterns that emerge across time</td>
</tr>
<tr>
<td>Entropy analysis</td>
<td>Provides a statistical measure of the amount of predictability present within a time series or set of interactions</td>
<td>Does not capture how each choice within a pattern relates to the other choices proceeding and succeeding it</td>
</tr>
<tr>
<td>Hurst exponents</td>
<td>Provides a long-term correlation of how each choice within a pattern relates to the other choices proceeding and succeeding it</td>
<td>In order to perform a reliable calculation, a large data set with multiple data points is needed</td>
</tr>
</tbody>
</table>

they also demonstrated higher target skill acquisition (Snow, Allen, Russell, & McNamara 2014). The use of Hurst exponents provides researchers with a novel way to look at dynamic movements as they occur over time. Within game-based systems, students are often afforded the opportunity to “choose their own adventure” or personalize their learning path. The Hurst exponents afford researchers a way to examine pattern fluctuations that manifest in students’ decisions as they exert agency over their learning path. One limitation of the Hurst exponent analysis is that in order to perform a reliable calculation, a large data set with multiple data points is needed. Although there is no hard-and-fast rule for the exact number of data points needed, in our work, each student completed an average of 275 choices. Understandably, this amount of data may not be readily available for most games. However, one way to combat this issue is to use the Entropy calculation, which requires fewer temporal data points. Although Entropy and Hurst do not measure the same constructs, they are both designed to calculate the relative order of a system or series. The difference, as discussed earlier, is that Hurst focuses on movements between choices (more fine-grained), whereas Entropy measures the overall state of the system. Thus, if a researcher wants to examine patterns of choices or behaviors within a game-based system but they have a smaller data set, Entropy can be calculated to glean an overall measure of a behavior pattern. However, when using Entropy, some fine-grained information will be lost that would otherwise be captured with the Hurst. All three of the methodologies presented here are potentially useful to researchers interested in examining how students interact within game-based environments; however, each has their own benefits and limitations. Table 4.3 provides a summary of the benefits and limitations of each method.

5 Conclusion

Game-based systems often provide students with high levels of agency by allowing them to engage in multiple types of interactions and develop an individualized learning path (Sabourin et al., 2012; Snow et al., 2013). Thus, log data from these
systems afford researchers with a unique means of tracing variations in students' choice patterns that may emerge across time. On their own, game-based log data can be complex and provide little insight into students' learning processes and cognitive states. However, the work described in this chapter demonstrates that dynamic techniques can shed light upon variations in students' behaviors within game-based systems and the impact of these variations on learning outcomes.

Dynamic systems analysis treats time as a critical variable which affords researchers the opportunity to not only look at aggregated information regarding students' interactions in game-based systems, but to also examine the fine-grained behaviors patterns that emerge across time. While the current chapter focused on how dynamic methodologies have been applied to log data from the iSTART-2 system, these techniques are generalizable to a variety of systems. For instance, Allen et al. (2014) have utilized Random Walks to visualize how high school students demonstrated flexibility in their use of various linguistic properties across 16 prompt-based essays (Allen et al., 2014). Similarly, Random Walks and Entropy analyses have been applied as a way to visualize variations in students' interactions within the game-based writing tutor, Writing Pal (Snow, Allen, Jackson, et al., 2014). Indeed, the tools and methods presented here can be used on any temporal log data.

Future work should focus on the practical use of these techniques within game-based environments to capture the emergence of these complex online behaviors. For instance, dynamical methodologies may inform student models in various adaptive game-based environments. Thus, if a student is engaging in a random interaction loop, dynamical methodologies can potentially "flag" this student and the system can then prompt the student to engage in more controlled patterns. Therefore, these analyses serve to inform and provide game-based systems with information about optimal and non-optimal learning patterns.

In conclusion, this chapter describes preliminary work that serves as a starting point for understanding how dynamical techniques can provide a means to trace and classify students' interactions within game-based environments, as well as other environments that offer multiple choices and pathways. All three of the analysis techniques described here (Random Walks, Entropy, and Hurst exponents) have revealed promising results as to how they can inform researchers about the various ways in which students engage with computer-based systems across time. We conjecture that tracing and modeling choice patterns across time will emerge as a key ingredient in better understanding learning processes.

Acknowledgments This research was supported in part by the Institute for Educational Sciences (IES R305G020018-02; R305G040046, R305A080589) and National Science Foundation (NSF REC0241144; IIS-0735682). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the IES or NSF. We would like to thank all of the members of the SoleilLab for their assistance with data collection.
References


