REGULAR PAPER

Wei Li · Mathias Funk · Quan Li · Aarnout Brombacher

Visualizing event sequence game data to understand player's skill growth through behavior complexity

Received: 28 December 2018 / Revised: 24 April 2019 / Accepted: 30 April 2019 / Published online: 25 May 2019
© The Author(s) 2019

Abstract Analysis of game data is used to study player behavior. For puzzle-based games where solutions are usually defined by their action sequences, player behavior can also be studied by their solution complexity. In this paper, we present a visualization system to help learning expert to understand how actions, timing and the resulting strategy change with regard to the solution complexity. To establish a novel perspective into the patterns not only in action choices but also in behavior complexity, we designed an interactive, customized line chart to track how complexity and performance change at each stage of skill acquisition. Specialized glyph system (Strategy Signature) is implemented to find strategy differences easily with simple visual cues. Contextual information can be explored by switching the view modes to see potential links between complexity and raw attributes. Evaluation with expert users shows that the system effectively reduced their time and effort in finding interesting subgroups and gave them unexplored angles of behavior complexity to contemplate player's skill growth. In summary, this paper illustrates a visualization approach to enable analysis into the subtleties of behavior complexity in video games.

Keywords Game visualization · Event sequence · Information theory · Application

1 Introduction

Like many other casual activities, video game is a platform that allows diverse interactions. The inputs during game play are results of individual thinking to a constantly changing situation. This makes game logs valuable assets to understand the player thinking preferences as well as skill learning processes. To obtain a structural understanding of the data, visual analysis is often utilized to leverage the visual channel to enable fast and flexible exploration into players' behavior patterns (Wallner and Kriglstein 2013). The visual approach is also known to lower the barrier of data analysis and thus exploit domain knowledge for informed judgment (Chen et al. 2009).

W. Li (✉) · M. Funk · A. Brombacher
Technische Universiteit Eindhoven, Eindhoven, The Netherlands
E-mail: wei.li@tue.nl

M. Funk
E-mail: m.funk@tue.nl

A. Brombacher
E-mail: a.c.brombacher@tue.nl

Q. Li
Hong Kong University of Science and Technology, Kowloon, Hong Kong
E-mail: qliba@connect.ust.hk

Game data usually take the form of an array of primitive factors in the game (gold, action, kill for instance). But visualization of these primitive values is not always sufficient to illustrate the more underlying behavior pattern. For instance, the complexity of actions, which is tricky to define and usually visually judged, can be used to reflect intrinsic behavior patterns at each state of the play. Based on the context of the game, this complexity can be interpreted as an abstract factor determined by the heterogeneity of players inputs. More interestingly, the complexity usually varies by each attempt a play takes. It is common to see repeated trials and strategy adjustments based on the latest outcomes in many video games. To obtain a more holistic view on how players improve, our understanding of player's and skill and skill growth should consider not only the score obtained, which only measures the attempt's outcome, but also the implicit strategy choice, execution and refinement that are not directly measurable. There are further influences on skill such as reflection ability, intuition and creativity that are out of scope for this article. We will study skill and skill growth by looking at the strategies played over time. Through the changes in measurable scores and player actions, behavior complexity varies accordingly and differs from each individual attempt at executing a particular strategy. Over time, these data show how strategies are chosen, executed and refined, and encoded with appropriate visual signals, the inherent complexity produced by players may lead to a new angle to study their behaviors as they improve through time. The strategy improvement is driven by the desire to obtain higher scores (reliably or in a reproduceable manner), but also by intertwined qualities of a strategy; among others, simplicity, ease of execution, speed, creativity or innovativeness, and reliability. Yet little existing literature has covered a visualization approach to facilitate such perspective, i.e., the integration of behavior complexity and effective visualization design.

In this paper, we present a visualization tool to validate such thinking in the context of puzzle-based game. The domain expert requires unexplored insights of player learning behavior. The implementation focuses on the behavior complexity and its relation to detail player actions, performance and strategies. On-demand multilevel information of game actions can be retrieved responding to user interactions. We introduced a new visual language called *Strategy Signature* to characterize player strategy and signify the intensity of in situ strategy shift. Evaluation with the expert users validates its capability in boosting research productivity and discovering several unseen patterns.

As the main contributions, 1) this paper demonstrated a visualization approach that depicts the implicit action complexity from categorical game log data to study player behavior. 2) A glyph-based visual encoding of event sequence data named *Strategy Signature* is also introduced to extract players' strategy features.

2 Related works

Here we enlisted works that share connections with ours in three categories. Beginning with visualization of game log data, we explained how previous work inspired this research. Then, we cover a few earlier discussions in entropy-based complexity for behavior study. Finally, some techniques in summarizing event sequences outside game field are mentioned.

2.1 Visual analysis for games

For visualization of game data, the design purpose and targeted user group are not always constant. Difference in such aspect significantly steers the design rationales (Wallner and Kriglstein 2013; Medler and Magerko 2011). Two most significant sorts are the entertainment-oriented visualization and the developer-oriented visualization.

The entertainment-oriented visualizations are also noted as “playful visualization” (Caillois and Barash 2001; Medler and Magerko 2011), since they aim at enhancing the game enjoyment itself. Entertainment features (such as online community (Medler 2011) or achievement system (Medler et al. 2011)) are integrated to motivate participation and provoke engagement. The developer-oriented approaches care less about the entertainment concerns. Logging results are primary analyzed to reproduce the gaming process for postmortem study of playing experience (Farooq et al. 2015; Ribeiro et al. 2017; Li et al. 2017). This allows for deliberate tweaking and adjustments of the game design at specific stage/phase of the gaming experience journey. Contrasting to these two categories, the domain requirement centered upon the generalizable behaviors patterns in problem solving and learning progress is not covered. This motivates a less explored approach to speculate the learning pattern and the behavior per se with visualizations.

In this regard, there are also works that share the same interests. Farooq et al. (2015) extract player behavior model in a routine manner, but the dynamic difference between time and individual is less observed. Hernández et al. (2017) utilize process mining technology to form behavior models in serious games (video game with educational purposes), which is effective but lacks advanced visualization and interactions. Wallner (2015) used a transition-focused approach to seek evidence on progressive changes of behavior patterns — specifically with event sequence data. However, the paper primarily emphasis on technical implementation and only basic visualization is implemented.

To synthesize abstract metrics beyond primitive factors, Moura et al. (2011) summarized the timestamps and performed interactions into a static of the selected session in an A-RPG game. Similarly, Li et al. (2017) used derived metrics to reveal the mechanism behind snowballing effect in MOBA games. However, the abstraction methods used are mostly designed for numeric attributes, which are incompatible with categorical events.

2.2 Informational complexity with entropy

In information theory, entropy is a measurement of the heterogeneity of possible states in a system. The very same method is also used to measure complexity in human behavior with sounding results (Fussell 2005; Chu et al. 2014).

For action events in a video game, measuring the complexity of gaming actions is essentially measuring the randomness or the lack of order in a given set of categorical data. To achieve this, there are several applicable methods. They are complexity index (Gabadinho et al. 2011), the turbulence (Elzinga and Liefbroer 2007) and longitudinal entropy (Gabadinho et al. 2011). However, the first two methods are sensitive on the sequence length, which are prone to duplicate noise of repetitive user inputs despite the same underlying intention. Considering this, the longitudinal entropy which streamlines the essential information quantity suits the best for its consistent performance in both verbose and short sequences.

Given that the p_i is the proportion of positions of the same action i and A is the alphabet size of all the possible actions, the formula of complexity can be defined as:

$$H(p_1, \dots, p_A) = - \sum_{i=1}^A p_i \log(p_i)$$

Provided the complexity can be quantified as scalar values, visualization of behavior complexity can be made possible. But how the behavior complexity varies especially in the context of repetitive gaming process is not touched by existing visualization work.

2.3 Techniques in summarizing event sequence

Making sense of massive event sequence is hard. Some authors insist using mining algorithms to find a simplified subset of events as a representative of similar ones. Their techniques usually vary in the sensitivity on different attributes to categorize sequences based on particular domain requirements. For example, Chen et al. (2018) designed a “soft pattern matching” mechanism to summarize multiple event sequences that tolerates minor inconsistencies of events for less cluttered result. Unger et al. (2018) used both semantic similarity and temporal similarity to form meaningful clusters. Apart from automated pattern extraction, user-defined matching rules before visual inspection or statistical analysis are also possible. Works like Cappers and van Wijk (2018) and Zraggen et al. (2015) adapted query languages and regular expressions to remove noise data and unrelated sequences with graphical user interface. Thus, categories sort by different rules form accordingly with cleaner visual result. However, this approach is more useful when most representative patterns are critical to the research objective. There are also scenarios when identifying frequently occurred sequence types is less important. Techniques in this regard are less explored.

There are also ways to improve the sense making in event sequences with primarily visual design. For example, LifeFlow (Wongsuphasawat et al. 2011) and CoCo (Malik et al. 2015) summarize low carnality patient journey events with a tree-structured view. *MatrixWave* (Zhao et al. 2015) achieved the summarization by mostly layout design, which is appealing because direct visual representation without the loss of information is proven to be highly effective. However, scalability issues may arise when information density increases and visual clutter is unlikely to be avoided easily.

3 Project background

3.1 About puzzle game

Puzzle games are a genre that has a low-entrance barrier for most novice players. Unlike fast-paced action or shooter games, puzzle game players are fairly distributed among both male and female, senior and junior participants (Brown 2017). This quality is essential to include players of more diverse backgrounds. Thus, the insights gained from these player behaviors are not restricted to a particular population segment. Apart from that, we also need a game platform that can quickly generate batches of game logs of multiple retrials, through which not only the solution and score, but also the progressive adjustments can be recorded and analyzed. Finally, we need to make choice between deterministic games (i.e., same input will always produce a fixed result regardless of the player and the time of attempts) and non-deterministic ones. With these considerations in mind, we choose Lix as the data collection platform since it is easily acceptable, quick for data collection and deterministic in mechanism design.

Lix is an open-source variation of Lemmings, which was originally developed in 1991 by DMA Design¹. The game consists of puzzle-like challenges, where new “lixes (the autonomously walking bots)” will be spawned into the two-dimensional level world at a steady pace. Each lix takes a direction until it either falls into water or it is forced to turn back after hitting an impassable obstacle. Players assign actions (such as **jumper**) to a lix, allowing it to interact or change the landscape in a certain way so that following lixes can take a different path toward safety. Available actions are **platformer** **jumper** **climber** **floaters** **batter** **blocker** **nuke** **exploder** and **miner**. With good timing and arrangement of the actions, player try to lead as many as possible lixes to the indicated destination without losing lixes. For example, a player can assign the **miner** to dig a tunnel into a hill or assign the label **platformer** which is useful in modifying the pathway before it becomes too steep for safe passage (cf. Fig. 1).

3.2 Expert background

There are 5 domain experts involved in the project. The leading expert (*E1*) has 4 years of research focus on studying human learning in game scenarios. She also designed and built the data collection infrastructure which provided fundamental support for the success of the visualization project. The second one (*E2*) is a game researcher who has a background in game visualization as well as data mining and published extensively in the field. The third expert (*E3*) has long-term collaboration with the video game industry and a deep understanding of the modern game development process. His knowledge has been widely applied in conducting user research in the game environment. The other two are assistant data analysts (*A1*, *A2*). Their backgrounds are more concentrated on data analysis and informatics. During the research, *A1* and *A2* purpose analysis reports and new methods in analyzing the data. They acquired reasonable knowledge about the game but, unlike *E1*, are not proficient players themselves.

The research is of an exploratory nature, meaning they intent to find as many as possible new ways to speculate the data. The domain experts confirm that learning pattern is essentially tangled with the decisions made during the play. However, how the actions can be understood should not be limited by frequently used analysis tools.

3.3 Material data

The original data set consists of all actions of 15 players in a total of 271 sessions (cf. Table 1 for structure). Every session is an attempt to finish a level successfully. Every time a player assigns an ability, eight attributes are recorded to describe the action. These include: *ID* as the player identifier, the *order number of attempt* (a new player will always start with 0), the assigned *action* (like **jumper**), *elapsed frames*, and identifier of the action *receiver lix*. **result** will be inserted as the ending action of each attempt. The number of *saved lix* will be saved as the final score whenever **result** is triggered. The number of *unique actions* and *seconds of time* elapsed since the beginning of current attempt.

Players are recruited from college students who do not have prior experience with the game. The chosen stage is a novice level beyond which three more difficult levels are available. As only 20 lixes will be

¹Par18 <http://www.lixgame.com>.

Fig. 1 Lix game: **jumper** is currently selected and ready to be assigned to any of the lixes. A deep tunnel into the ground are made by **miner**s. Yellow bars to the opening are made by **platformer**s, a type of action to make small horizontal bridges to let others walk through. The destination is can be found in the next screen to the left where the player can navigate to by touching the screen edge with pointer

Table 1 Source data format: data points of the 1st attempt by player No.1

Player	Attempt	Action	Update	Lix	Saved	Ability	Second
1	1	Jumper	76	0			
1	1	Platformer	158	2			
1	1	Platformer	174	3			
—	—	—	—	—	—	—	—
1	1	Result	9205	NA	0	4	614

released per attempt, a player can score between 0 (none saved) and 20 (all saved) depending on skill level. Diversity in solutions is encouraged.

4 User study

We investigate the domain requirements in three phases: 1) understanding their current workflow and examining the pain points, 2) exploring design possibilities with visual mock-ups, and finally 3) the concluding requirements for design guidelines ensuring that the designed solution fits the context and domain needs.

4.1 Phase I: domain investigation

4.1.1 Process

In this phase, we performed semistructured interview with the domain expert. They were asked to demonstrate their current work style and latest discoveries. In the meanwhile, questions regarding how insights were generated by current methods were asked. Experts are free to suggest needed functionalities. From their demonstration, we clarified their usage of apparatus and pipeline for generating raw data from in-game replays. It was also informed that players are selected from non-players of the lix game. Each games session is played in an isolated experimental place to ensure players achieve their performance independently. A1 and A2 shared their existing explorations (mostly with R packages). They also pointed a few possible future works in the incoming agenda.

4.1.2 Pain points

The initiation of such dialog helped us to obtain a holistic overview of their work and identify design opportunities to improve the research. As a result, we identified a few inconveniences in their current data analysis workflow. For instance, the analysis operations are hindered by the cumbersome reconfiguration in the source code to alternate between different research hypothesis. Also, the result produced by the algorithm is efficient and simple but usually entangled with trust and communicability issues. Expert would appreciate a direct way to reconnect the results to its original context and draw insights with clear awareness of related factors. The general-purpose plotting method in R they use frequently is too primitive and lacks levels of detail on demand. The possibility of immediate inquiries for further evidence is not well supported.

4.2 Phase II: iterating alternatives

Designing with mock-up is known to be a viable method to boost productivity in communication and design development (Ferreira et al. 2007). As the experts are not knowledgeable about visualization designs, we employed co-design method with versions of on-screen mock-ups to communicate the requirement in more explicit details. This went through an iterative process by gradually adjusting functions and layouts (cf. Fig. 2).

With a clear understanding of how the visualization tool is going to be used in real world, former pain points are incrementally answered by enriching the design details. Elaborating each next version of visual design is achieved by inputs from the last version. This helped us to identify quite a few insights. For instance, the time between actions can be insensitive to the result of solution pattern. However, as learning behavior study, user may still wonder how the player is eloquently deciding what to do next. Then, seeing the time difference maybe necessary.

4.3 Phase III: summarizing requirements

Based on previous insights, we discussed and enlisted the most important capabilities the visualization should support. To the most extent, the final work should meet the requirements below.

R1: Experts prefer the information to be visually organized by players. Since the research is mostly focused on player behavior, the individual difference and group similarity always play a key role in the analysis of people. The visual design should be centered upon players to support the difference finding. Visual elements of actions also need to be sorted by player as a priority. Comparisons between players, which emphasize on the subtle individual differences in behavior patterns, is strongly valued by the experts.

R2: Explorations into detailed attributes. Analysts do not want to overlook important information. This means quick peek into all available contextual information in the raw data is preferred. The ability to revise the game play situation with every recorded detail for scrutiny can be helpful. Such implementation would eliminate a substantial effort from the time-consuming task of reviewing game replays.

R3: Experts want to study the details of how play strategies improve based on previous attempts. Beyond performance, researchers also care about the patterns that preceded before strategy improvements. As learning is an important concept in our case, the incremental performance improvement may indicate

Fig. 2 Alternative designs and rationale history (color bleached for presentation): *Ver.1* differentiates action types with color. Actions are horizontally placed by time, and each line represents a single attempt. Scores are displayed to its left. *Ver.2* added **a** aggregated player performance, **b** filter to show/hide certain action for better focus, **c** ability to reduce actions to mere sequences by checking “ignore time” box. *Ver.3* reserves both sequence action view and timed action view for different tasks. Switching between score by attempt and score by player is made possible by top controls. Same for action display modes

valuable behavior patterns in this context. Therefore, a view to compare consecutive attempts and reason with the factors is needed.

R4: Experts need a way to compare diverse play styles. According to their experience, players would reach their success differently. But they lack an intuitive way to see or describe the qualitative difference. They look forward to a visual means to reveal such trait, which is not as easy to obtain with statistical computations. Visually identifying the differences in play strategy could be an effective way to convey such characteristics.

R5: Understand how the complexity of player strategies influences their performance. Experts frequently use “complexity” in describing player action combinations. However, such complexity is easily perceived by looking at the heterogeneity and density of actions. But comparing is difficult if two similarly complex sequences of actions are presented side by side, let alone bird viewing the global trend. Visual comparison between more instances of strategies is challenging yet appreciated to the experts.

5 Methods

To satisfy the defined requirements, there are a few used methods that we would employ to sketch out our design. The involved methods also distinguish our work from most of other event sequence visualizations in video games.

5.1 Behavior complexity

Following the action, data, complexity pipeline, we introduce the concept of behavior complexity which refers to the measurable heterogeneity from data as a result of complicated behaviors. The higher the behavior complexity, the more one subject is to perform actions that are diverse in type and unpredictable in composition. Here we exemplify the quantification of behavior complexity and discuss its meaning in the context of puzzle game.

5.1.1 Quantifying complexity

The strategy for this type of puzzle-based game is defined by a set of action combinations, which outcomes are series of event sequences. Therefore, behavior complexity is associated with longitudinal entropy which takes only the order and occurrence of a subset (or all in rare cases) of the available options and calculates the information quantity. Complex actions, which may be attributed to the undetermined thinking during play (Li et al. 2018), will introduce more information when measuring its action combinations. To acquire this result, the implementation follows through three steps (cf. Fig. 3):

- Data points as discrete actions are aggregated into action sequences according to their belonging attempt. The produced action sequence preserves the original temporal order of each action it contains.
- Quantify the entropy of each action sequence. The exact computation is based on Gabadinho et al. (2011)'s longitudinal entropy.
- Appending the entropy value to a new column. The output file was used in combination with the original data set to inform different facet of the issues in game play.

5.1.2 Required complexity or redundant complexity

In this video game, initial experiments with one or two simple moves can barely succeed. As an example, we observed that some players start their game with only a handful of random actions and quickly get a game-over with zero score achieved. Such failure is not incidental because although being a simple game, stages in *lix* usually require a minimum level of complexity in the conception of strategy to pass.

But complexity increase does not always correlate higher performance. If carelessly performed, the correlation can potentially negative. It is the reflective thinking instead of mindless inputs that leads to the learning and growth. Making aimless noisy actions certainly generate high complexities, but redundant and unnecessary movements are apparently not encouraged, since they counter the spirit of learning, i.e., carefully improving by choosing effective approaches.

Fig. 3 All actions in the same attempt are merged into one event sequence to calculate behavior complexity. The output format consists of data points of attempts as event sequences with an entropy-based complexity level. Every returned data point has a unique combination of player number and attempt number

This makes reasoning with only complexity metric difficult because user needs to understand the complexity in the specific context. Otherwise, judgment upon whether the complexity is a result of active learning or blindly exhausting the actions is difficult.

Complexity expansion should only be appreciated when it contributes to performance increase. Here we assume that simpler solutions are more preferable once they are able to achieve full score, which is an indication of refined thinking and planning. To make this discussion easier, we define the complexity that can be reduced without damaging the final score as *redundant complexity*, while the minimum complexity of actions required to pass the stage as *required complexity*. The definition of the two concepts provides two functions. First, it raises the awareness of undesired over-complication, which may not be a reflection of meaningful growth but immaturity of a solution instead. Second, the ratio of redundant complexity and required complexity may be an indicator of stage difficulty on the basis of players which may find some stages difficult to reduce the redundant complexity, suggesting the stage to be a difficult one.

5.2 Strategy Signature

Although some strategies are close in the complexity level, there is a considerable chance that the exact solution patterns are fundamentally different. This means complexity alone is good at telling the subtleties in strategy composition. To summarize the qualitative difference in strategy and play style, we need to have a clear grasp of how strategies differ from or assimilate with each other. To this end, behavior complexity may not be the ideal way to facilitate this goal. To address this predictable problem, we created a new functionality named Strategy Signature.

Strategy Signature is glyph-based model which renders a (long) action sequence into a compact circular polyline glyph. The coordinates of consisting points are determined by their actions and order of appearance, which are mathematically defined as:

Here the coordinate is constructed with a radial angle θ (determined by the c_{th} action type order in the predetermined list of actions) and distance to the center r (determined by the time order t). N ($N = 10$) is constant size of total action set. R is the maximum radius of the outer line.

This model ensures that same actions will always be aligned to the same radial direction, which is useful to tell the usage and frequency of certain actions. As shown in Fig. 4, the produced shape is sensitive to the time order of actions, which allows distinction between similar action statistics with different time orders.

With the new glyph system, we can easily judge the diversity of strategies such as sorting the strategies into a few typical categories. While preserving the subtle difference between similar ones, incremental changes and variation in strategy in consecutive attempts are still observable. The goal of this design is to make both commonality and distinction of strategies stand out without the hassle of scanning each event in detail.

Fig. 4 Example: Strategy Signature of event sequence platformer platformer climber blocker miner batter batter floater batter

Fig. 5 The default view mode layout with 3 views and 2 panels: (1) player and performance shows overall performance of a player; (2) actions and attempts displays attempts and actions within in horizontal lines; (3) complexity and growth enables exploration of complexity change along score; (4) control buttons is used to change view modes and apply filter; (5) bottom bar is dedicated area for Strategy Signatures

6 Design

6.1 Layout

As shown in Fig. 5, the design of the visualization interface is a few cohesively positioned views—visible dividers are replaced with natural spaces to provide better visual clearance. On the top of layout sits the

control bottoms for global effects (see Sect. 6.3 *Interactivity* for details). Three (implicit) columns focusing the *player and performance*, *actions and attempts* and *complexity and growth* take the major body area. Strategy Signatures are aligned at the bar at the bottom of the screen for the ease of comparison.

The first two columns (*player and performance* and *actions and attempts*) are vertically aligned, meaning *actions and attempts* next to the *player* on the right are ones performed by the displayed player (**R1**). These areas enclose the list of all the information from the raw data. User can browse up and down dynamically for searching the interested subset.

6.2 Visual encoding

Analysis inquiries differ in the demand of level and type of information, such as the distinction between individual actions and the complexity of all the actions in an attempt. We designed specialized encoding to treat this need, respectively. Here are the few most important key points to cover.

6.2.1 Action

To easily differentiate the action types is the most basic as well as most frequently performed task. The pre-attentive channel of color is selected for this task (Healey et al. 1996). Nine vibrant colored dots on a dark background are utilized to display them with contrast (cf. left of Fig. 7). Player of the actions is placed in juxtaposition on the left. Each attempt by the player composes a horizontal line of colored dots. Number of attempts as well as action number in each attempt may vary from player to player. Skimming through the action dots gives user a direct grasp of the exact events in detail (**R2**).

6.2.2 Complexity and performance growth

The quantified behavior complexity and performance variation are displayed next to them on the right (cf. Fig. 6). Here the covariance of achieved score and complexity value are depicted in a scatter plot view. In this view, each dot in the area depicts an individual attempt (instead of single actions), with higher score to the right and more complex to the right. Here, the *x*-axis indicates the performance of an attempt in percentage and the *y*-axis represents the complexity level normalized to the range [0, 1]. Each attempt is plotted as a diamond with white border.

Since classical scatterplot does not provide the ability to communicate grouping and temporal orders, we implemented joining lines to highlight the attempts belonging to the same player. For the temporal order, a family of visual appearance is design to describe the gradual steps of the one before, previous, current, next and the one after (cf. Fig. 6). In this encoding rule, the blue ones are before the current selection and the red ones are beyond. Temporally adjacent attempts are filled and more distant ones are outlined.

This design is an enhanced version of the scatter plot with the extra ability to delineate the relationship of time order in a subgroup explicitly.

6.2.3 Strategy

User can quickly skim through the colors of the dots to obtain an general view of used actions in the attempts. However, finding strategic similarity and difference in them is not directly supported as to reduce the load of visual memory (Luck and Vogel 2013). Without some degree of summarization, drawing connections between the used strategies can be difficult as the number of attempts and actions are larger in quantity.

Based on the algorithm introduced in Sect. 5.2, *Strategy Signature* we can encode the sequences in to glyphs which take significant less visual information. Figure 7 is an example of how two distinctive strategies by Player 2 are illustrated. Two strategies used very different set of actions, resulting in dramatic contrast in the outcome shapes. Subtle difference or adjustments are also visible even attempts belong to the same general strategy pattern. In *strategy1*, where **jumper** is frequently used as a starter action, the similarity of the attempts is visually reflected by its signatures. Also, the last attempt has some modification at the ending actions by including more **miner**s. Such adjustment is also visible from the signatures as the last signature takes more triangular shape than preceding ones (**R3**).

Fig. 6 Visual encoding of temporal order in growth journey. Time order of consecutive attempts is encoded with different appearances with blue for the past, white for current selection and red for the future. Growth lines are treated separately with solid line and dashed line to distinguish current selection and last/historical selections. Strategy Signatures are collectively displayed at the Bottom Bar to see how strategy evolves by iteration. In this case, Player 2 experienced a dramatic strategy shift after the 5th attempt and scored impressively after the 4th attempt

Fig. 7 Color and strategy signature at identifying strategies difference: an example of two distinctive strategies by Player 2

6.3 Interactivity

In this section, we describe how the questions of the data can be answered by interactions like selection, hovering or inspections.

6.3.1 Action filter

While the user is interested in a few actions and wonders how they influence the final score (**R2**), they can isolate the visual elements into a smaller set. Action filter allows leaving out the irrelevant actions (colors) to only selected types. Thus, the frequency of certain actions among other players is presented with improved clearance.

By clicking action buttons on the top panel, distraction of the irrelevant colors can be turned on/off. Deselected ones will be grayed out after click. As shown in Fig. 8, Player 7 uses much less **climber** in the later stage of his attempts. This is especially useful to see how occurrence of certain actions affects performance over time or comparing the action choice among different players (**R1**, **R3**).

6.3.2 Switching view modes

The visualization system provides two modes for two levels of analysis. The default view of Fig. 5 displays most relevant information to the research goal, such as player overall performance (shown in percentage under the player label), action sequences, strategy character and complexity. Timed series events are depicted in aggregated encoding of strategy and complexity where detailed temporal information of raw data is hidden to create this compact workspace. In this mode, the attribute of interval time between actions is not visible. To enable user with sufficient contextual information (**R2**), user can click the switcher bottom—*Timed Actions* at the screen top to display restore the timing information.

To facilitate comparison between different players, dashed lines are used against solid line to provide enough contrast between two or more players. This is especially useful when comparing players with more attempts than average.

7 Evaluation

To validate the effectiveness of this system, we conducted a semistructured evaluation with each of the domain experts, respectively. The evaluation process consists of two stages: the orientation phase and self-guided phase. The goal is to help us to understand 1) to which extent the tool is capable of facilitating knowledge generation of unknown aspects and 2) how much the design is effective to meet the identified requirements (Sect. 4.3 *Phase III: summarizing requirements*).

To begin with, an introduction that covers the key functionality explained above was given. We demonstrated the how each view and control would work interactively. Domain experts were asked to perform the task again to ensure their familiarization with the design is sufficient. Then, they were encouraged to navigate each part freely. Questions regarding any ambiguity in the usage will be answered.

After the orientation, experts were given full access of the system and freedom to analyze the game log data to their need. They started with some confirmatory analysis of their prior knowledge which had been obtained by the old workflow. For instance, the outstanding learning ability and sharp performance increase of Player 2 is easily confirmed. And the popularity of **miner** among all players is much more quickly noticeable with visual means.

Beyond these discoveries, the experts also identified novel patterns which were not approachable before. These discoveries are usually associated with more implicit information of strategies and complexities.

7.1 Novel discoveries

7.1.1 D1-complexity increase

The idea of quantified behavior complexity caught analysts' attention immediately after our introduction. As an initial exploration, they are curious about more and less complex attempts would perform globally among all the attempts of the 15 players. One expert clicked through all the players' label each at a time until all the player's growth journey is drawn. Here, the global patterns of complexity distribution with performance are shown (cf. Fig. 10).

Fig. 10 Behavior complexity growth: the lines tends to link the bottom left corner to the upper right, suggesting experience gain coexisted with complexity increase

As shown in the view, the growth lines are heavily intertwined. This confirms the prior assumption that player's behavior can be highly heterogeneous in terms of growth pattern. From the visual result, the analyst also feels that the graphic seems to suggest that players generally tend to start from less complex attempts and gradually reach higher score by combining more complex moves. Previously described case of Player 2 in Fig. 6 is a good example of this behavior. By referring to the Strategy Signature and exact actions on the left, we can argue that the player's behavior experiences three stages: (1) Player 2 begins by experimenting with several **jumper** and **nuke**, (2) Player 2 included some extra **miner** moves and found a working solution, and (3) Player 2 continues to refine it by removing the unnecessary **jumper**s and added a few more diverse actions which introduces more complexity in behavior. However, such abrupt success since the fifth attempt is rare. This particular player has scored nothing in between and reached full score from zero. This seems to be an outlier as we find it not as prevalent for the rest of the players.

This observation has confirmed that the user is able to discover the global correlation of complexity increase and player experience on the overall level as well as dig into specific players to validate such assumption (R1, R5).

7.1.2 D2-“Tail optimization” behavior

This time the analysts focus on the successful attempts only, which are positioned at the leftmost vertical area (the highest score percentage) in the player growth line view. They have a particular interest in how players change their solution after a success is already obtained. The user selected Players 2, 7, 9 and 13 because these players have achieved continuous high scores in the last a few attempts.

By comparing the growth line of these players, the user finds that, among the high score attempts in the later stage, any one with a complexity over 0.63 or below 0.58 will eventually gravitate toward 0.6, which is also where most successful attempts resides (cf. Fig. 11). This seems to suggest that the solution to the level requires this amount of complexity to win. However, based on the discussions in Sect. 5.1.2 **Required Complexity or Redundant Complexity**, it is easy to that players will optimize for simpler actions as they improve because it is a less expensive way to get the same score. But counterintuitively, players more often “over-complicate” their solutions after they win which seems to run the opposite of optimization.

After switching the view mode to Timed Actions, the reason became clear: Player's later successful attempts always use less time to finish and the exceptions are few. This suggests that player likes to trade behavior complexity for shorter time consumption.

This confirms that user can always seek evidence to certain hypothesis by players (R1) as well as retrieve additional contextual information (time consumption in this case) (R2).

7.1.3 D3-strategy categories

One engaging quality of this game is its wide possibilities of solutions. This means given the same stage that there are usually more than one way to pass. Despite the minor difference in exact execution, we found that users can spot the categories of successful attempts confidently with their visual similarities.

By checking the shapes of signatures of various players, user can tell the similarity among solutions. For instance, the best attempts by Players 2, 3, 4, 6, 13 share very similar shape in their Strategy Signature (cf. *i* in Fig. 12). Similar solution is taken by Player 9 and Player 7 with minor deviations. Players 14, 15, 5 choose a fundamentally different composition of actions. This can be reflected by the highly discernible contrast in the glyph appearance.

Predicting the success rate of an attempt is therefore made possible by looking at its signatures and judging the likelihood of fitting into a successful category. The *failed* cluster in Fig. 12 exemplified the best

Fig. 11 Tail optimization: final adjustments to improve previous success

Fig. 12 Best strategies by players: using Strategy Signature to sort different solutions visually

attempts by players who never managed to get a full score. This can be inferred by acknowledging its odd visual shape.

The discovery of strategy categories with Strategy Signature helped experts to not only confirm the game stage can be played in versatile ways but also unveiled the exact possibilities of successfully routines from different players (**R4**).

7.2 Expert feedback

The enlisted cases above have presented explicit examples of insights that are otherwise difficult to obtain without the visualization tool. After experimenting with the tool, we asked the analysts to give detailed feedback on the design. We specifically prompted the analysts by revising the initial requirements in Sect. 4.3 *Phase III: Summarizing Requirements*. Then, we collected their views concerning to which extent the design is effective regarding the requirements.

All the experts consider our design to be useful in identifying interesting players or attempts efficiently while preserving the subtle differences. According to their domain knowledge in player behavior, players that exhibit higher persistence in trying can be regarded as a long-term performer, meaning the player is likely to do well in long-term accumulation of experience. For example, *E3* has found that Player 5 has made gradual improvements in the strategy. “Especially that you can see that they approached simpler and simpler solutions using **climber** and **miner**.” This is regarded as a positive trait in learning, according to *E1*. *E1*, *E2*, *E3* and *A1* reported that design has facilitated the understanding of player growth journey and see how strategy transform and refinement would impact the score in each original temporal order (**R3**).

Both player growth journey and action dots jointly synthesize a clear depiction of multiple attributes (inherent actions and derived behavior complexity) of selected player. The interactive interface allows them to dig into any aspect of the attributes rather efficiently (**R1**, **R2**). “I like that fact that you can quickly gain an overview about which are the most dominant ability used by players, e.g., P2 focus on **jumper**s, and across all players you find a lot of red, pink, light, blue, and orange, indicating that **platformer**, **miner**, **climber**, and **jumper** are abilities that players gravitate towards,” said by *E3*. The interactivity also gives the analysts freedom to scrutinize how much effort was put into the repeated attempts and how the player learns (**R3**).

“The global view of complexity trend is very inspiring (**R5**),” stated *A2*. Although the data points are limited in quantity to validate if such pattern is a reliable insight, he believes this extra dimension to study the behavior is interesting and worth continued discussion.

The use of Strategy Signature has rediscovered a solution that was not considered seriously. Before using the tool, *A2* believes that the most suitable solution should be a modest deviation from the solution by Player 2 (cf. *i* Fig. 12). And the solution by Player 5 is a very risky and less recommended approach. From the signatures, *E1* rediscovered the success on Players 14 and 15. This leads to a perspective change on the diversity of the play (**R4**). “Yes, maybe the other (strategy iii in Fig. 12) is a viable solution,” said *E1*.

In general, the analysts think the visualization tool is useful in terms of generating new knowledge from the game log data and effective in terms of boosting their research productivity. They are especially surprised by the ease of visually detect behavioral patterns. The interface is well organized and polished, according to *E2*. Interactions and transitions are coherent and intuitive to understand. Despite the visualization approach is not capable of all the operations of their conventional data analysis, it leads to novel ideas and hypothesis which are otherwise difficult to find. The beauty and enjoyment of graphics are truly a plus.

On the other hand, they also shared their opinions of some inconveniences from our work. There are particularly two instances - 1) The Timed Action view is a bit difficult to find (by *E3*). User may only navigate in the default view mode, assuming that temporal information is simply discarded for ease of

visualization (by *AI*). 2) They also suggested a functionality to cherry-pick some action sequences of the dots and squeeze them together to eliminate vertical scrolling (by *EI*, *AI*).

8 Discussion

8.1 Applicability and caveat

There are two novel methods we employed in our design—the entropy-based approach to quantify behavior complexity in the repeated gaming and the glyph-based illustration to extract characteristics of a sequence. The former gives us an objective perspective on how much diversity/heterogeneity exists in an event sequence, which, in this context, is useful to detect how much player are doing complex action than previous ones. A precondition of this usage is the actions are devised facing an unvarying challenge. Thus, comparison of solution simplicity is possible. The later uses radial display of event types with a randomly defined order, which could introduce unnecessary bias when the order of event types does not inherently exist. Also, the color choice may signal similarities between some actions when there isn't. Or, on the other hand, similarities of event types may exist but color choice failed to convey this information. This problem can be moderated with using automatic color picking tools like ColorBrewer (Harrower and Brewer 2003). However, conventional models of color seem inadequate to eliminate this issue completely (Szafir 2018). We advise manual selection by trained graphic expert to cope with exact situations. The characteristic extraction method is also sensitive to the cardinality of sequences if the number of event types increases. For optimum result, cardinality should be no bigger than 12 to produce clean and effective distinctions.

Since the domain experts are mainly focused on learning more than gaming, we also consulted game experts to discuss the generalizability of our approaches in other games. We found that there are plenty of other game genres where our system is useful. 1) Tower defense games like plants vs. zombies (Electronic-Arts 2016), where choices of plant defenders are encoded as events and key to the play. In these games, the construction of defenders is likely to cost a certain amount of resources, which makes streamlining the token use (like action use in *lix*) equally relevant toward skillful gameplay. Likewise, the complexity of construction and building choices still remains an indicator of skillfulness as lower resource consumption means more late game elasticity and snowballing potential. Considering the huge genre (Unkown 2019), applicable games can be many. 2) We also find our approach applicable to fighting games like Street Fighters (Capcom 2018), in which players can repeatedly perform combos (a series of controlled attacks all at once) and combat routines can be analyzed. In this genre, the damage dealt can be regarded as the score achieved. Players with better attack accuracy are likely to produce more damage impact within a smaller set of moves (e.g., heavy punch, jump backward). However, our tool needs to be adjusted to also consider other factors such as the character choice and time consumption which are especially sensitive factors necessary for fair assessments in these type of games. Considering the data type, the tool can be presumably used to investigate the complexity of hospital operations in the IT system (Spaulding et al. 2013), or to understand diversity of visitors travel routine for different activities in an amusement park (Liu et al. 2017).

In general, given by the raw data being in event sequence format, we found the application area of the two methods can be extended to other fields where reasoning with the diversity and complexity of repetitive solution is meaningful.

8.2 Future work

We planned two branches in our future works. On one hand, we would like to continue improving the system with more functionality we discovered from the evaluation such as improving functionalities based on the feedback. There also unspoken requirements such as the automatic sorting of similar Strategy Signatures. The integration of data mining techniques may be helpful as the visual signature and large group clustering can each solve the communicable problem and the efficiency problem. The complement of both would create much potential to be explored. On the other hand, we would seek application scenarios outside game domain, such as EHR data in medical fields where patients' journey are encoded as event sequences (Craig 2011; Malik et al. 2015).

8.3 Conclusion

The paper is motivated by facilitating the analysis of learning in player behavior. It novelly introduces two methods in enabling domain experts with advanced analytical abilities in reasoning with the complexity of behavior and drawing connections and distinctions between strategies and play styles. As a result, the interactive visualization tool facilitates player behavior analysis in four effective ways: 1) by studying the effect of certain actions on final score, 2) by joining behavior complexity and performance in the context of play style difference, 3) by exploring how players' strategy evolves in repetitive attempts, and 4) by categorizing provided solutions of a given stage. Use cases of the experts indicate that the design enables swift explorations to the discovery the behavior patterns that are difficult to find with expert user's existing tools.

The design has received plausible feedback from the expert users in terms of usefulness, usability and efficiency. The evaluation validated the Strategy Signature's usefulness in characterizing player's attempts and successful strategies. The utilization of the entropy-based measurement of behavior complexity allows for quick visual deductions of how players are moving toward complicated decisions or simplification of actions. Systematically, it is proven to be a meaningful contribution to the analysis of game log data in *liX* game. And the application area of the utilized methods can be extended to other games and other domains.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Caillois R, Barash M (2001) *Man, play, and games*. University of Illinois Press, Urbana
- Capcom (2018) *Street fighter V | rise up!* <https://streetfighter.com/>
- Cappers BCM, van Wijk JJ (2018) Exploring multivariate event sequences using rules, aggregations, and selections. *IEEE Trans Vis Comput Graph* 24(1):532–541. <https://doi.org/10.1109/TVCG.2017.2745278>
- Chen M, Ebert D, Hagen H, Laramée RS, van Liere R, Ma KL, Ribarsky W, Scheuermann G, Silver D (2009) Data, information, and knowledge in visualization. *IEEE Comput Graph Appl* 29(1):12–19. <https://doi.org/10.1109/MCG.2009.6>
- Chen Y, Xu P, Ren L (2018) Sequence synopsis: optimize visual summary of temporal event data. *IEEE Trans Vis Comput Graph* 24(1):45–55. <https://doi.org/10.1109/TVCG.2017.2745083>
- Chu D, Sheets DA, Zhao Y, Wu Y, Yang J, Zheng M, Chen G (2014) Visualizing hidden themes of taxi movement with semantic transformation. In: 2014 IEEE Pacific visualization symposium, IEEE, Yokohama, Japan, pp 137–144. <https://doi.org/10.1109/PacificVis.2014.50>
- Craig D (2011) An EHR interface for viewing and accessing patient health events from collaborative sources. In: 2011 International conference on collaboration technologies and systems (CTS), pp 319–324. <https://doi.org/10.1109/CTS.2011.5928704>
- Electronic-Arts (2016) *Plants vs zombies* video games—PopCap Studios—Official EA Site. <https://www.ea.com/studios/popcap/plants-vs-zombies>. Accessed 28 Dec 2018
- Elzinga CH, Liefbroer AC (2007) De-standardization of family-life trajectories of young adults: a cross-national comparison using sequence analysis. *Eur J Popul/Rev Eur Démograph* 23(3–4):225–250. <https://doi.org/10.1007/s10680-007-9133-7>
- Farooq SS, Baek J, Kim K (2015) Interpreting behaviors of mobile game players from in-game data and context logs. In: 2015 IEEE conference on computational intelligence and games (CIG), IEEE, Tainan, Taiwan, pp 548–549. <https://doi.org/10.1109/CIG.2015.7317895>
- Ferreira J, Noble J, Biddle R (2007) Agile development iterations and UI design. In: Agile 2007 (AGILE 2007), IEEE, Washington, DC, USA, pp 50–58. <https://doi.org/10.1109/AGILE.2007.8>
- Fussell E (2005) Measuring the early adult life course in Mexico: an application of the entropy index. *Adv Life Course Res* 9:91–122. [https://doi.org/10.1016/S1040-2608\(04\)09004-5](https://doi.org/10.1016/S1040-2608(04)09004-5)
- Gabadoh A, Ritschard G, Mueller NS, Studer M (2011) Analyzing and visualizing state sequences in R with TraMineR. *J Stat Softw* 40(4):1–37
- Harrower M, Brewer CA (2003) ColorBrewer.org: an online tool for selecting colour schemes for maps. *Cartogr J* 40(1):27–37. <https://doi.org/10.1179/000870403235002042>
- Healey CG, Booth KS, Enns JT (1996) High-speed visual estimation using preattentive processing. *ACM Trans Comput-Hum Interact* 3(2):107–135. <https://doi.org/10.1145/230562.230563>
- Hernández JAC, Duarte MP, Doderio JM (2017) An architecture for skill assessment in serious games based on event sequence analysis. In: Proceedings of the 5th international conference on technological ecosystems for enhancing multicultural (TEEM 2017), ACM, New York, NY, USA, pp 50:1–50:9. <https://doi.org/10.1145/3144826.3145400>
- Li Q, Xu P, Chan YY, Wang Y, Wang Z, Qu H, Ma X (2017) A visual analytics approach for understanding reasons behind snowballing and comeback in MOBA games. *IEEE Trans Vis Comput Graph* 23(1):211–220. <https://doi.org/10.1109/TVCG.2016.2598415>
- Li W, Funk M, Brombacher AC (2018) Toward visualizing subjective uncertainty: a conceptual framework addressing perceived uncertainty through action redundancy. In: Lawonn K, Smit N, Linsen L, Kosara R (eds) EuroVis workshop on

- reproducibility, verification, and validation in visualization (EuroRV3), The Eurographics Association. <https://doi.org/10.2312/eurovr3.20181144>
- Liu Z, Wang Y, Dontcheva M, Hoffman M, Walker S, Wilson A (2017) Patterns and Sequences: interactive exploration of clickstreams to understand common visitor paths. *IEEE Trans Vis Comput Graph* 23(1):321–330. <https://doi.org/10.1109/TVCG.2016.2598797>
- Luck SJ, Vogel EK (2013) Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cognit Sci* 17(8):391–400. <https://doi.org/10.1016/j.tics.2013.06.006>
- Malik S, Du F, Monroe M, Onukwugha E, Plaisant C, Shneiderman B (2015) Cohort comparison of event sequences with balanced integration of visual analytics and statistics. In: *Proceedings of the 20th international conference on intelligent user interfaces (IUI '15)*, ACM Press, pp 38–49. <https://doi.org/10.1145/2678025.2701407>
- Medler B (2011) Player dossiers: analyzing gameplay data as a reward. *Game Stud* 11(1). <http://gamestudies.org/1101/articles/medler>. Accessed 19 Mar 2017
- Medler B, Magerko B (2011) Analytics of play: using information visualization and gameplay practices for visualizing video game data. *Parsons J Inf Mapp* 3(1):1–12
- Medler B, John M, Lane J (2011) Data cracker: developing a visual game analytic tool for analyzing online Gameplay. In: *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '11)*, ACM, New York, NY, USA, pp 2365–2374. <https://doi.org/10.1145/1978942.1979288>
- Moura D, el-Nasr MS, Shaw CD (2011) Visualizing and understanding players' behavior in video games: discovering patterns and supporting aggregation and comparison. In: *Proceedings of the 2011 ACM SIGGRAPH symposium on video games*, ACM, New York, NY, USA, Sandbox '11, pp 11–15. <https://doi.org/10.1145/2018556.2018559>
- Brown A (2017) Who plays video games? Younger men, but many others too. In: *Pew Research Center*. <https://www.pewresearch.org/facttank/2017/09/11/younger-men-play-video-games-but-so-do-a-diverse-group-of-other-americans/>. Accessed 16 Apr 2019
- Ribeiro PC, Biles ML, Lang C, Silva C, Plass JL (2017) Visualizing log-file data from a game using timed word trees. *Inf Vis* 17(3):183–195. <https://doi.org/10.1177/1473871617720810>
- Spaulding TJ, Furukawa MF, Raghu TS, Vinze A (2013) Event sequence modeling of IT adoption in healthcare. *Decis. Support Syst* 55(2):428–437. <https://doi.org/10.1016/j.dss.2012.10.002>
- Szafir DA (2018) Modeling Color Difference for Visualization Design. *IEEE Trans. Vis. Comput. Graph.* 24(1):392–401. <https://doi.org/10.1109/TVCG.2017.2744359>
- Unger A, Dräger N, Sips M, Lehmann DJ (2018) Understanding a sequence of sequences: visual exploration of categorical states in lake sediment cores. *IEEE Trans. Vis. Comput. Graph.* 24(1):66–76. <https://doi.org/10.1109/TVCG.2017.2744686>
- Unkown (2019) Tower Defense. Wikipedia. https://en.wikipedia.org/w/index.php?title=Tower_defense&oldid=886569060. Accessed 5 Apr 2019
- Wallner G (2015) Sequential analysis of player behavior. In: *Proceedings of the 2015 annual symposium on computer-human interaction in play (CHI PLAY '15)*, ACM, New York, NY, USA, pp 349–358. <https://doi.org/10.1145/2793107.2793112>
- Wallner G, Kriglstein S (2013) Visualization-based analysis of gameplay data—a review of literature. *Entertain Comput* 4(3):143–155. <https://doi.org/10.1016/j.entcom.2013.02.002>
- Wongsuphasawat K, Guerra Gómez JA, Plaisant C, Wang TD, Taieb-Maimon M, Shneiderman B (2011) LifeFlow: visualizing an overview of event sequences. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, ACM, New York, NY, USA, pp 1747–1756. <https://doi.org/10.1145/1978942.1979196>
- Zraggen E, Drucker SM, Fisher D, DeLine R (2015) (s, Qu)Eries: visual regular expressions for querying and exploring event sequences. In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems (CHI '15)*, ACM, New York, NY, USA, pp 2683–2692. <https://doi.org/10.1145/2702123.2702262>
- Zhao J, Liu Z, Dontcheva M, Hertzmann A, Wilson A (2015) MatrixWave: visual comparison of event sequence data. In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems (CHI '15)*, ACM, New York, NY, USA, pp 259–268. <https://doi.org/10.1145/2702123.2702419>