Frontiers in Educational Research
ISSN 2522-6398 Vol. 4, Issue 10: 90-95, DOI: 10.25236/FER.2021.041018

An exploration of graphical programming learning models based on behavioral sequence analysis

Huang Ruiyan

Faculty of Education, Beijing Normal University, Beijing, China
3072781198@qq.com

Abstract: *Compared with learner engagement, the sequence of learning behaviors during the learning process is a better reflection of the trajectory, willingness and cognitive process of learners' learning behaviors. This study empirically explores learners' behavioral sequences and learning patterns through lagged sequence analysis based on log data on the Snap! graphical programming platform. It was found that learners' learning behaviors exhibited the following characteristics: they showed a clear performance orientation; when encountering difficulties, they often took measures of trial and error, consulting information and reviewing previous learning tasks; and when interacting with classmates or encountering difficulties, their attention was more easily distracted. The findings demonstrate that the sequence of learners' learning behaviors can present teachers with a more comprehensive picture of online learning, help them discover learners' learning habits, preferences, and cognitive processes, and assist them in reflecting on the teaching and learning process.*

Keywords: *graphical programming, learning process, behavioral sequences, lagged sequence analysis*

1. Introduction

Learning process data reflects the state of learners in the learning process. With the continuous development of graphical programming, the learning process data of learners in using graphical programming software has attracted the attention of scholars. Many researchers have used data mining techniques to analyze learning data in the field of graphical programming education to uncover the hidden and potentially valuable information and knowledge. For example, based on Brennan and Resnick's three-dimensional framework, the development of students' computational thinking on three dimensions in graphical programming instruction was explored through learners' learning process data, and the hybrid reorganization of programming code was found to have a significant effect on the development of computational thinking skills (Kafai, Y. B. et al., 2014). Recent studies have found that sequences of learning behaviors during learning, compared to student engagement, better reflect learners' trajectories of learning behaviors, willingness and cognitive processes (Yang, Xianmin et al., 2016). For example, learning behavior sequence analysis is used to study the behavior patterns throughout the activity (Hou et al., 2009), as well as the comparison of behavior patterns in different learning stages and different learning achievement groups (Yang et al., 2015), etc. Inspired by this, this study conducted a process analysis using learners' learning behavior sequences for the two experimental sessions of the "Introduction to Learning Science" course in the fall semester of 2020 at Beijing Normal University based on Snap!

2. Review of the literature

At present, there is no clear definition of the concept of learning behavior, and different scholars have given their own definitions of learning behavior from different levels, different learning environments and different learning objects. For example, some scholars, starting from the network environment, consider learning behavior as the sum of distance independent learning behaviors (Peng Wenhui et al., 2006) or learning activities (Yang Jinlai et al., 2008) carried out by learners in the learning environment created by modern information technology with new communication mechanisms and rich resources. Some scholars have also pointed out that learning behavior is both the positive and negative aspects of the learner's behavior that he or she exhibits in learning, in terms of attention, motivation, learning attitudes and strategy application covered by learning behavior (Yao Chunzhen et al., 2009).Considering various factors such as learning environment and learning objects, the study defines learning behavior as

ISSN 2522-6398 Vol. 4, Issue 10: 90-95, DOI: 10.25236/FER.2021.041018

the sum of learners' two-way interaction activities with their surroundings to obtain certain learning outcomes, guided by certain motives (Yu et al., 2013). Based on this, the sequence of learning behaviors in this study refers to the sequence of behaviors formed according to the time sequence of learning behaviors, which means that the transition from one behavior to another is defined as a sequence of behaviors.

Finding a number of sequences of learned behaviors that are significant can reveal patterns of learned behavior in learners. Lag Sequential Analysis (LSA) is a method proposed by Sackett (1978) for testing the significance of behavioral sequences, aiming to explore human behavioral patterns by analyzing the significance of the probabilities formed by the occurrence of one behavior after another. Among the existing studies, some of them have used lagged series analysis to explore the differences between groups in behavioral patterns. For example, Li Shuang et al. (2017) explored online learning behavioural sequences and engagement patterns using LSA on log data from 2,131 students at the Open University on the Moodle platform. Their study found that different online engagement patterns could be defined based on behavioral sequences, which in turn led to more targeted pedagogical improvements.

The process data generated by graphical programming is easier to obtain comprehensively and clearly, and is a quantitative representation of students' learning behaviors during process activities. And learning behavior can to some extent characterize students' implicit behaviors in the form of extrinsic operations of external behaviors to characterize learners' learning status, which is important for understanding online learning patterns (Ma et al., 2019). In view of this, this study applies lagged sequence analysis to explore the patterns of learners' learning behaviors based on the existing learning process dataset.

3. Research methodology

3.1. Research questions

Based on existing research, this study attempts to apply lagged sequence analysis to explore the following questions: What are the behavioral sequences that have a probability of occurring at a significant level during the learners' learning process? What are the patterns of learners' engagement in learning graphical programming?

3.2. Data processing and coding

In this study, 30 undergraduate students (4 male and 26 female) majoring in educational technology who had complete records of their learning process on the platform in two sessions of the "Introduction to Learning Science" course based on Snap! The student learning records were aggregated to obtain 18,453 data, forming the main behavioral data sample for the study.

Based on the action records generated by students learning in class using the Snap! platform, 54 action behaviors were defined and recorded in the original dataset. When analyzed, it was found that there was a combinatorial relationship between some of the behavior records, i.e., there must be a fixed follow-up behavior after a certain behavior occurs. Therefore, similar learning activities were combined in this study and grouped into 11 learning behaviors as the behavioral objects for the behavioral sequence analysis. Table 1 presents the names, descriptions, source activities, and study codes of the 11 behaviors.

Table 1 Behavioral descriptions and codes

Name of the act	Description of behaviour	activities	encoding
Engineering Setup	Open, export, and save projects	IDE.exportProject, IDE.opened, IDE.toggleAppMode	GS
Engineering Testing	Suspend, stop, start work	IDE.greenFlag, IDE.pause, IDE.stop	GC
Create edit module	Creation, setup and modification	Block.created, IDE.changeCategory, IDE.setSpriteTab	BC
Select Module	option	Block.grabbed	BS
Run module	movement	Block.clickRun	BR
Delete Module	removing	Block.dragDestroy, Block.userDestroy	BD
Withdrawal of operation	Withdraw previous operation	Scripts.undrop	BU
Reproduction module	replicate	Block.duplicateAll	DU
Editing parameters	Parameter modification, input, editing	InputSlot.menuItemSelected, MultiArg.addInput, MultiArg.removeInput, InputSlot.edited	IN
mistakes	An error message appears	Error	ER
standstill	No operation for more than 30s	standstill	Stop

ISSN 2522-6398 Vol. 4, Issue 10: 90-95, DOI: 10.25236/FER.2021.041018

In this study, the transition from one behavior to another is defined as a sequence of behaviors, and the sequence formed by the two behaviors is represented by the combination of the codes of the two behaviors, and the order before and after the behavioral codes in the combination represents the direction of the behavioral transition in the sequence. For example, BS-DU denotes the sequence of behaviors in which the module (BS) is selected first and then the module is copied (DU). Based on the definition of behavior sequences, the 11 behaviors under study form a sequence of 121 behaviors available for study.

3.3. Research tools and methods

The study first used the Python language to write an algorithm to extract the learned behavior data from the original dataset and generate a behavior sequence file in the format required for the study; after that, the lagged sequence analysis tool GSEQ5.1 was used to calculate the frequency and probability values of the sequences formed by the two behaviors and to map the behavior transformation patterns based on the results.

4. Research findings and analysis

The study applies lagged sequence analysis to examine the behavioral sequences of learners reaching significance levels in the learning process, constructs a behavioral sequence transition diagram of learners learning through the Snap! graphical programming platform, presents a panoramic picture of the learning process, and further analyzes the behavioral patterns and learning habits reflected by the behavioral sequence transition diagram.

4.1. Overall behavioural sequence

The study began by counting the frequency of sequences of learning behaviors produced by students during their course of study, resulting in a table of the frequency of transitions between the 11 learning behaviors shown in Table 2. The rows in the table represent the starting behavior, the columns represent the behavior that occurs immediately after the behavior in that row, and the data in the table represent the frequency of occurrence of the sequence of behaviors formed by the two behaviors. For example, the number 607 in column 5 of row 4 indicates that the total frequency of occurrence of the sequence of behaviors that immediately follows the run module behavior (BR) after the selection module behavior (BS) is 607. Also, the general picture of the frequency of the sequence occurrence can be quickly found in Table 2, for example, the behavior sequences with more occurrences are BS-BS (continuous selection behavior for modules, 1077 times), BR-BR (continuous running behavior for modules, 3000 times), and the behavior sequences with less occurrences are BU-IN (withdrawal of operation behavior to jump to template parameter editing behavior, 3 times), etc.

Table 2: Frequency matrix of behavioral sequences

n	GS	GC	BC	BS	BR	BD	BU	DU	IN	ER	Stop
GS	24	1	74	5	8	0	0	0	1	2	24
GC	0	46	2	5	24	0	0	0	5	1	5
BC	30	3	441	128	151	19	2	7	88	0	69
BS	9	1	106	1077	607	111	62	74	195	5	92
BR	28	26	212	332	3000	197	40	286	574	4	275
BD	7	4	23	259	63	32	14	10	14	0	23
BU	0	0	5	54	24	43	130	4	3	4	5
DU	0	0	2	157	92	22	16	40	114	1	32
IN	5	3	53	215	698	5	4	41	997	2	111
ER	3	1	0	5	6	1	0	0	3	40	2

4.2. Analysis of behavioural change patterns

According to the theory of lagged series analysis, Table 3 shows the residual table adjusted on the basis of the behavior frequency matrix, and the data in the table are the residual values (Z-score) adjusted for the frequency of occurrence of the behavior sequences formed by the two behaviors. Where, Z-score > 1.96 indicates that the frequency of occurrence of the behavior sequence is statistically significant ($p < 0.05$). From Table 3, it can be found that 30 behavioral sequences such as GS-GS (jumping from setting up a project to setting up another project), GS-BS (jumping from setting up a project to creating an edit

ISSN 2522-6398 Vol. 4, Issue 10: 90-95, DOI: 10.25236/FER.2021.041018

module), and IN-BR (jumping from editing parameter behavior to running a module) occurred with a significant level of frequency. In order to present a more visual representation of the jumps between behavior sequences during the learning process, a behavior sequence transition diagram was drawn based on the 30 behavior sequences that were significant, as shown in Figure 1. The nodes in the diagram represent various behaviors in the learning process, the connecting lines between the nodes represent the meaningful behavioral sequences, and the arrows represent the direction of behavioral transitions.

Table 3: Table of adjusted residuals (z-scores) based on the behavioral frequency matrix

n	GS	GC	BC	BS	BR	BD	BU	DU	IN	ER	Stop
GS	19.208*	-0.006	20.177*	-4.465	-7.987	-2.245	-1.733	-2.312	-5.05	1.654	5.664*
GC	-0.981	55.66*	-2.025	-3.199	-2.534	-1.833	-1.415	-1.887	-2.911	0.84	-0.237
BC	6.474*	-1.657	44.596*	-4.613	-15.972	-2.844	-4.37	-5.216	-6.937	-2.287	1.258
BS	-3.376	-4.3	-6.461	38.623*	-13.953	3.706*	2.05*	-1.454	-11.875	-2.051	-4.606
BR	-4.112	-2.176	-11.853	-26.898	39.603*	2.322*	-8.164	9.838*	-12.371	-5.204	-1.777
BD	1.162	0.421	-2.126	22.123*	-10.891	4.307*	1.535	-1.656	-7.809	-1.495	-0.772
BU	-1.689	-1.423	-3.682	0.758	-10.18	11.253*	53.207*	-1.946	-6.929	2.403*	-2.9
DU	-2.252	-1.898	-6.09	8.591*	-8.78	1.408	1.969*	5.608*	4.422*	-0.862	0.719
IN	-3.967	-3.499	-10.028	-10.69	-6.365	-8.971	-6.765	-4.75	40.701*	-2.831	-1.709
ER	3.009*	0.828	-2.287	-2.058	-4.682	-0.799	-1.154	-1.539	-2.488	73.346*	-0.911
Stop	3.747*	0.175	-0.545	-3.572	-0.791	-1.38	-3.159	-3.03	0.286	-0.911	9.692*

The numerous cross-linkages between the different nodes in Figure 1 indicate that students' behavioral transitions on the Snap! platform are richer. As can be seen from the figure, among the 11 learning behaviors, selecting modules (BS), editing parameters (IN), and running modules (BR) have more sequences with significant meaning compared to other behaviors. In addition, it can be found that these sequences include both one-way jump sequences and two-way sequences of interaction with each other. Frequent learning behavior sequence analysis, i.e., identifying the learner-preferred sequence patterns among all the learning behavior sequences generated in online learning, can to some extent reflect the characteristics of learners' online learning activities.

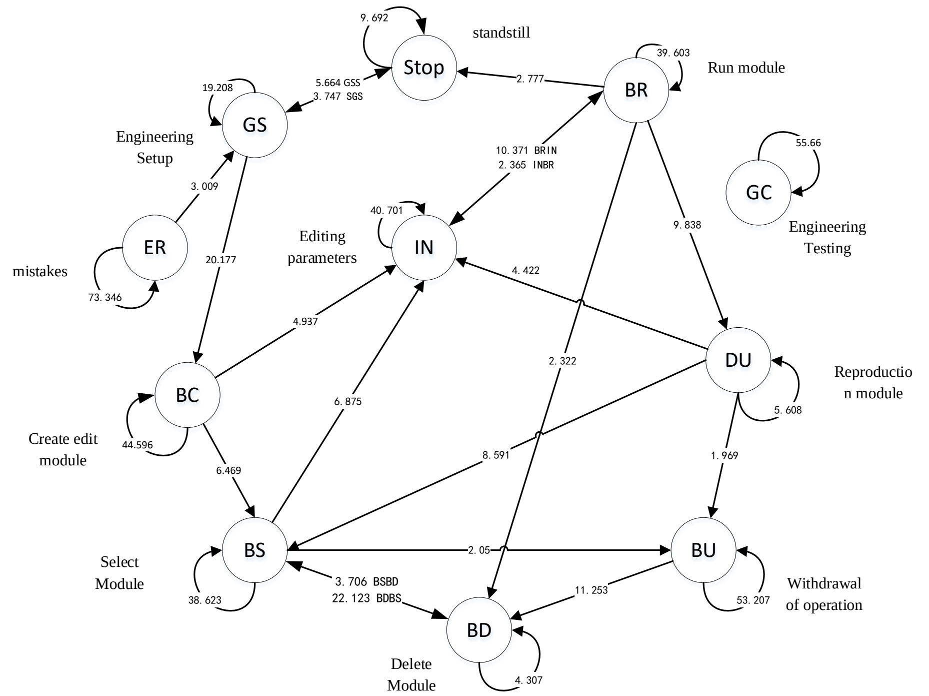

Figure 1: Behavioural sequence transformation diagram

The analysis of the behavioral sequence transition diagram shown in Figure 1 reveals some potential behavioral patterns, learning habits, and cognitive processes of learners: (1) Since the project provided by the teacher to the learner contains descriptions of all subtasks and code blocks for the subtasks, the learner usually follows a step-by-step approach in the activity and tends to copy the existing code blocks of prompts, which is somewhat "performance-oriented". After entering the learning system, they first focus on the assessment criteria in the learning evaluation, then modify the parameters according to the task requirements (DU-IN) and view the results of the run (IN-BR). (2) When learners encounter problems in the learning process, the following solutions are often adopted: one: trial and error. For example, IN-BR, BR-IN indicates the behavior of constantly adjusting the parameters to keep running the module. Turkle et al. believe that learners' cognitive patterns during programming learning can be divided into two types: tinkerers and planners. Overall the students in this study tended to be tinkerer learners perform, i.e., they exhibited small incremental iterative modifications, constant trial and error, and debugging of code during the programming process, with a smaller amount of code modified each time. Second, jumping to a stop (BR-Stop) after running the module tended to be indicated by students looking for solutions by consulting paper learning materials distributed by the instructor, discussing with classmates, or asking the course AI assistant. Third, by reviewing previous learning tasks, such as copying completed blocks of code for parameter modification, to gain inspiration to solve the problem currently encountered (DU-DU, DU-IN). (3) Compared to other learning behaviors, when learners encounter difficulties or interact with classmates, they are more likely to be distracted into a constant cycle of stopping behaviors (Stop-Stop) and may engage in activities that are not related to learning. The reasons why students were more easily distracted after both GS and BR behaviors: on the one hand, they may have encountered difficulties in learning the basics of Natural Language Processing (NLP), resulting in their inability to stimulate their learning interest when carefully understanding the learning tasks for graphical programming operations and directly abandoning the lesson at the first task, while constantly switching between the two activities in order to cope with the teacher's rounds during classroom learning (GS-Stop , Stop-GS); on the other hand, they may give up trying and engage in learning activities that are not related to learning (BR-Stop) when they have problems running the module, are unable to understand the learning material when consulting it or are still unable to find a solution after talking to their classmates. Also, it is often more difficult for learners to exercise self-regulation when they may be engaging in non-learning related behaviors, making it difficult to return even to the current learning process, which coincides with what often occurs in graphical programming learning environments (Stop-Stop).

5. Discussion and summary

Learning process data reflects the state of learners in the learning process. By analyzing learning process data, it is important for teachers to discover teaching phenomena, grasp teaching rules, adjust teaching strategies, optimize teaching paths, improve teaching effectiveness and teachers' professional development capabilities.

5.1. Identifying effective behavioral sequences and patterns can facilitate understanding of teaching and learning processes

The analysis of learning behavior sequences can facilitate teachers' or teaching researchers' understanding of the teaching-learning process, thus helping researchers and pedagogues, especially young teachers who lack teaching experience or are new to teaching, to discover students' learning paths, grasp students' behavioral preferences, and grasp students' underlying behavioral patterns (Jiang Bo et al., 2018). For example, in the lag sequence analysis, the IN-BR, BR-IN, BR-Stop, DU-DU, and DU-IN behavioral sequences with significance indicate that the problem solving approach often adopted by learners is constant trial and error, access to information, and review of previously mastered content.

At the same time, comparing the behavioral sequence transformation diagrams drawn in the lag sequence analysis with the diagrams of the pedagogues' predetermined learning patterns can help the pedagogues identify the differences that exist between the predetermined and actual learning patterns and support the pedagogues to reflect on their teaching and develop practical knowledge.

5.2. Learner learning models and their implications for instructional design

(1) Supportive resources and scaffolding are important in the process of learning with graphical programming languages. From the learning behavior sequence, it was found that learners often encountered problems and produced repeated runs or pauses in the learning process, indicating that they

ISSN 2522-6398 Vol. 4, Issue 10: 90-95, DOI: 10.25236/FER.2021.041018

had questions about the task and wanted more guidance and support from materials, peers, or teachers. Good scaffolding and the construction of a basic knowledge system is important for the acquisition and transfer of knowledge in the learning process of learners, and teachers should focus on the construction and application of supporting teaching resources, such as paper-based instructions, task descriptions and example codes, and not simply provide them, but further consider exactly what support should be provided when.

(2) Constant trial and error is an important way for learners to learn in online graphical programming platforms. The large number of two-way behavioral sequences in which all learners repeatedly debugged and ran the code indicates that it is often difficult for learners to complete the programming task in one go and requires constant modification and repeated debugging during the programming process. Teachers should inform learners of this situation in advance to help learners improve their self-efficacy. At the same time, teachers can provide good models of learning behavior to students to support their self-regulation of learning, reflecting the support and facilitation of technology for teaching and learning.

(3) More consideration needs to be given to how to motivate learners intrinsically in the course of building a curriculum using graphical programming for NLP learning. There is less overall thinking, less independent exploration, creativity, and authoring, and more repetitive dependent behaviors among students. The curriculum design should consider designing more open inquiry activities to get students interested in the content itself, motivate them to think proactively, engage more fully in various learning activities, develop higher-order thinking skills development, and promote overall development, while ensuring that the various learning activities required for the learning priorities are fully implemented.

References

[1] Li Shuang,Zhong Yao,Yu Chen,Cheng Gang,Wei Shunping(2017) Exploration of online learning engagement model based on behavioral sequence analysis. Chinese e-learning, 3, 88-95.

[2] Ma C,Wang I-Chi,Yue Yun-Zhu,Du Hong-Yu(2019) Research on online learning behavior model under the threshold of learning analytics. Modern distance education, 6, 35-44.

[3] Peng, Wenhui, Yang, Zongkai, Huang, Kebin (2006) Analysis and modeling of online learning behavior. China e-learning, 10, 31-35.

[4] Yang, J. L., Hong, W. L., Zhang, Y. X. (2008) Research and practice of real-time monitoring of online learning behavior. Open Education Research, 4, 87-92.

[5] Yang Xianmin,Wang Huaibo,Li Jihong(2016) The application of lagged sequence analysis method in learning behavior analysis. China Electrochemical Education, 2, 17-23+32.

[6] Yao, Chunzhen, Mi, Jianrong, Wang, Hongcheng (2009) An overview of learning behavior research at home and abroad. Teaching and Management, 3, 48-50.

[7] Yu, Xiaohua, Gu, Xiaoqing (2013) Learning activity flow:A behavioral model for learning analysis. Journal of Distance Education, 4, 20-28.

[8] Hou, H.-T., Sung, Y.-T., & Chang, K.-E. (2009) Exploring the Behavioral Patterns of an Online Knowledge-Sharing Discussion Activity Among Teachers With Problem-Solving Strategy. Teaching and Teacher Education, 25, 101-108.

[9] KafaiB.Y., LeeE., & SearleK. (2014) A Crafts-oriented Approach to Computing in High School: introducing Computational Concepts, Practices, and Perspectives with Electronic Textiles. ACM Transactions on Computing Education, 14, 1-20.

[10] SackettP.(Ed.).G. (1978) Observing Behavior: Theory and applications in mental retardation (Vol. 1). Baltimore: University Park Press.

[11] YangX., LiJ., GuoX., & LiX.(2015) Group Interactive Network and Behavioral Patterns in Online English-To-Chinese Cooperative Translation Activity. The Internet and Higher Education, 25, 28-36.