A “User-Flow Description” Method for Usability Investigation

Akira Takahashi(✉)

Faculty of Informatics, Shizuoka University, Hamamatsu, Japan
akirtaka@inf.shizuoka.ac.jp

Abstract. This study proposes user-flow (UF) description, a new method for investigating usability. UF is defined as a series of user operations. A delay in flow in one part of a task procedure can reveal the interface problem in that part of the task. A UF description is a method for investigating usability by measuring and describing a user’s operation flow. The method has three parts: (1) a time-series description consisting of a table that lists user’s operation along the time series; (2) a personal UF description summarizing the time line for a flow diagram of the operation and describing each individual’s UF; and (3) a population UF description integrating the individual UF description for one figure to measure the usability of the all user’s operation. This study describes the method’s purpose and explains the procedure of the UF description method. Finally, we discuss a problem and its solution using the method.

Keywords: Usability measurement · User-flow · Qualitative and quantitative analysis

1 Introduction

Using a tool or system for achieving a goal requires the user to learn how to operate the tool or system. Typically, the user’s purpose is not to learn the system, but to use it to achieve some goal. However, sometimes fatigue or lack of resources (attention, motivation, or time) for using complex and difficult to learn tools, such as computers or smart phones, can prevent the user from achieving the goal. This results in the low usability of complex systems.

Usability is defined in ISO9241-11 as “The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.” We are able to measure these three elements of usability both quantitatively and qualitatively. Qualitative measurement data might be obtained through a user test, but such data are difficult to summarize. In contrast, a quantitative measurement has the advantage of being able to summarize multi-user data in a single index. However, despite this advantage, quantitative data are difficult to obtain from large subject samples.

Qualitative and quantitative measurement each has its own advantage. We have developed a new measurement method, the user-flow (UF) description, to combine the advantage of each type of measurement. The goal of the UF description method is to measure effectiveness and efficiency of a usability definition both quantitative and qualitative terms.

C. Stephanidis (Ed.): HCII 2016 Posters, Part I, CCIS 617, pp. 155–160, 2016.
DOI: 10.1007/978-3-319-40548-3_26

2 User-Flow (UF)

We consider a series of procedures that a user is follows to achieve a task as a continuous stream, because the series of actions involved in such procedures are not independent from each other. In general, actions taken to achieve a goal are dynamic and continuous. When a succession of acts has stalled for some reason, then we notice a discrepancy in the task performance process. Accordingly, we regard a user's performance as a flow-like structure, UF.

UF is a series of user operations taken towards achieving a purpose, the sequential steps of a procedure that a user actually performs. It has following features:

1. A user aims to achieve the goal through a minimum number of steps, but this might not succeed in some situations. UF follows a straight path from the user's point of view, but the actual path might deviate in practice.
2. UF is restricted by cognitive resources such as attention, motivation, and memory, and by such physical resources as time. Once these resources are exhausted, the UF interrupts the task sequence, preventing the user from achieving the goal.
3. The structure of UF evolve from the skill level toward the rule revel before finally reaching the objective level. For example, in the event that the user is learning to use a computer system to write a report, s/he must first learn how to use the physical devices (e.g., keyboard, mouse, and printer). Once these skills are acquired, the user's attention would shift to how to use the application. When the user has learned the rules of the application (e.g., for a word processing software), then s/he can concentrate on the content of the purpose (e.g., report). The user must perform the task only on this objective level, not in the skill or rule level.
4. UF develops toward maximum efficiency with minimum cost as the user users the system repeatedly. When the learning of the UF is completed, the UF can proceed automatically without consuming further user resources.
5. UF becomes nested. On the way to the goal, sometimes a user must solve another "lower-level" problem of the task. The new lower level's operation is called "sub-flow." The user can not return to the primary UF until the sub-flow is resolved (see Fig. 1). If the nested sub-flows become sufficiently deep, the UF could not return to the primary level because the user's resource become exhausted along the way.
6. UF emerges from the interaction between user interface and the user's previous experience. Hence, UF is created by the user, not by a designer or engineer.

In contrast, work-flow (WF) is defined as the operational steps set up by a system or user interface designer. For example, the procedure represented in user operation manuals comprise WF. WF is the ideal operation procedure that is considered to be the most effective and efficient. WF provides the foundation for analyzing UF.

3 How to Describe UF

The UF description method involves four stages. First, a user's operations are recorded using video or operation capture software (when the target to be measured is a computer system).

The second stage involves observing the user's behaviors and describing them step by step in a table, along with a time series, in accordance with the rules (Table 1). Each column of the table contains nine slots, each of which is filled by analysts (and if possible) with the cooperation of the participant. The detailed information in the table is the basis for the further analysis. The detailed information in the table is the basis for further analysis.

Table 1. Time-series description

Categories	Explanation	Example of description
Time	Elapsed time for user's operation	1"58
Handled object	An object that the user has actually used	Floating menu
Object to operate	An object that the user aims to use	Graph
Real action	An observed action	D (drag)
Purpose	Purpose of an act in the task	To focus on graph
Result (success, S or failure, F)	Whether the act is a success or failure	S
Operation level (1, 2, 3 …)	Depth of flow: primary flow is 1, sub-flow is 2, and so on.	2
Contribution to the purpose in principal level (Y or N)	Whether the act contributes to achieving the task or not	N
Comment	Noteworthy relevant information (if necessary)	Moving the menu to operate the graph under the menu

1. Time: The start time of the described event relative to the begging of the task.
2. Handled object: An object that the user actually touched. In this example, the user moved the floating menu of the software.
3. Object to operate: The object that the user actually wants to work on. In this example, although the user has moved the floating menu, the real purpose of that behavior was to focus on the graph behind the menu. In this case, the object to work is not the menu but the graph. The contents of this slot is determined by reasoning from the next movement or from an introspective report from the participant.
4. Real action: The observed specific action that a user has performed. Abbreviations are used in this slot, for example, "d" standing for "drag", "dc" standing for "double click."

5. Purpose: The meaning of an act in the task described using a phrase that indicates the purpose of the behavior. This slot is filled based on the contents of slot 3.
6. Result: A record of whether the act ended in success ("S") or in failure ("F"). If the user failed to click the object, the result is "F."
7. Operation level: A number here indicates the depth of the flow level. The primary flow level is "1" and the numbers increase incrementally with sub-flow levels.
8. Contribution to the purpose in main level ("Y" or "N"): In a simple case, the user failed to click the object, and it is recorded as "N" (marked as "F" in slot 6). In a more complex case, when the entire sub-flow was incorrect, all the steps in the sub-flow are judged as "N", even if not every operational procedure failed. This is because the judgement "Y" in this slot would indicate the progress on the UF to the goal, whereas the incorrect sub-flow did not contribute to reaching the goal at all. In the same sense, the steps involved in reading the manual are necessary but are not considered to be contributing to the task and thus are judged as "N."
9. Comment: Describe any notable information meriting further analysis.

The third stage involves drawing the UF diagram for each participant based on the time-series table created at the second stage. Figure 1 is a conceptual drawing that explains the UF description. To draw a diagram, the analyst takes the WF as the foundation. The WF consists of only primary-level flow. The information in the table is further broken down into some units relating to each operation object in the task. In this stage, the order information in the operating was discarded and the all steps are rearranged in order of WF for comparison of each UF diagram.

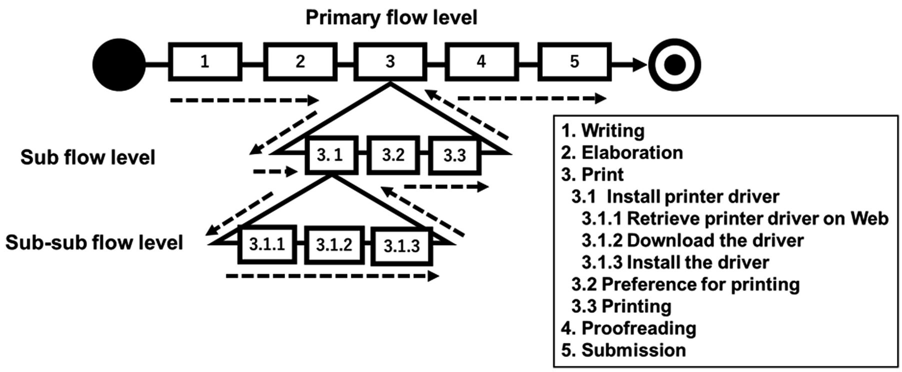

Fig. 1. Conceptual diagram of UF for submitting a report

For the sample in Fig. 1, the user wants to submit a report. Although the first and second steps on the primary level (1 and 2) were proceeding smoothly, a problem with printing the paper has occurred in the third step. To print the paper, the user had to install a printer driver (down to the sub flow level); furthermore, the printer driver is distributed on the Internet, requiring the user to find the driver by searching the network (down to the sub-sub flow level). It is important to note that we designate these sub-flow and sub-sub flow actions as not contributing to achieving the task, as noted previously, since the user's purpose is to submit the report, not to install or search for the printer driver. Certainly, installing the printer driver was a necessary step in the printing procedure, but it is not an essential work directly related to submitting the report. Ideally, these sub-flows and sub-sub flows should be performed automatically by the computer or the printer itself. Users would prefer working towards task goals by proceeding only on the primary level of the UF without needing to dive into sub-flow level. Consequently, the tools must assist the user in achieving the task's true purpose. Procedures not contributing to the purpose itself are essentially useless. This analysis illustrates the qualitative aspects of a user's operations.

We can also use this diagram for measurement of the quantitative side. Figure 2 is an actual analysis example of the UF diagram for a participant. The triangles comprise sub-flow (and sub-sub flow) levels. The numbers separated by slashes in each triangle are the total number of actual steps, the number of the effective or contributed steps, and the number of ideal steps (WF), in order. These values represent the efficiency of the procedure that the particular participant performed.

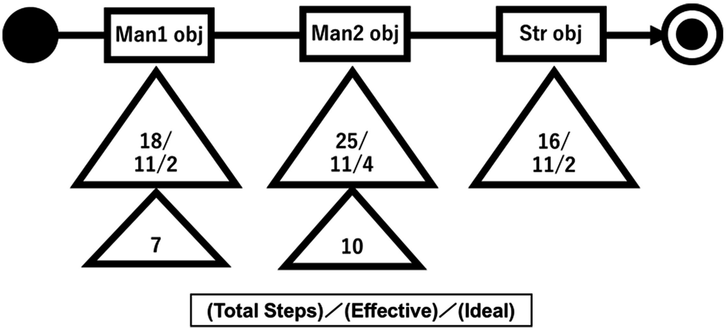

Fig. 2. Example of a participant's UF diagram for a poster construction task on graphic software

In the following procedure, the all users' full diagram is summarized by individual diagrams in Fig. 3. The shape of the full UF diagram is almost consistent with the individual diagrams but differs in two points. First, the meaning of each user operating in sub-flow levels is abstracted, showing only the number of procedures, as we focus only on the number of steps in sub-flows to measure unnecessary steps, regardless of their cause. Second, for each actual step, the minimum and maximum number among all of the user operations are described in parentheses. This indicates the magnitude and the range of the number of actual steps from all users. The description is also adopted for effective steps (but the WF steps are constant). As a result, the full users' combined diagram contains the distribution of the users' operational steps for specific conditions. These values are used, for example, for estimating the quality of a user's manual. When the range is narrow, the procedure may be represented clearly in the manual and much of the users will follow the procedure in the same way. Conversely, in the case where the range is wide, that may suggest that the manual represents ambiguous expressions or multiple procedures, causing different users to follow the procedure in different manners.

If an operation problem is found in total users' diagram, an analyst is able to refer backs to each of the user's diagram for detailed analysis and if necessary return to the time-series table of each participant to dissect the problem.

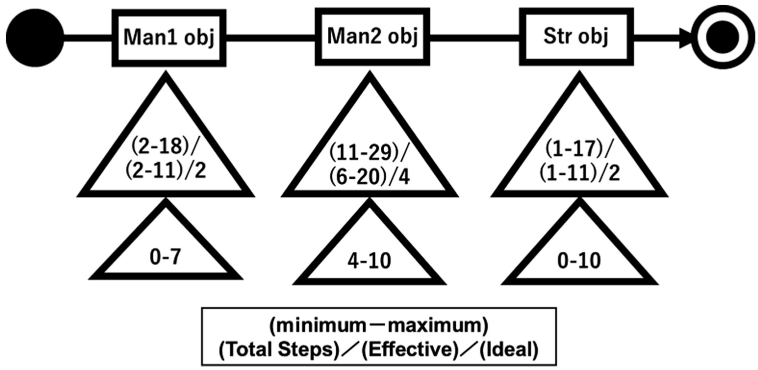

Fig. 3. Example of all users' full UF diagram for a poster construction task on graphic software

4 Conclusion

The principal concept of this methodology is the creation of a detailed UF table from an operational video (or screen capture data) of the user. The most important aim is to infer the intent of a user's action to enable an analyst to describe the user's behavior more exactly. The analyst can also access or select any necessary data from the table and can summarize the diagram in different manner toward other analytical purposes. The UF diagram explained in this study is one of the analysis tools that can derived from the data summarized in the UF table.

Since this methodology is still under development, the details of the technique are likely to change as work progress. In addition, we have found some inherent problems in this method.

This UF description method has three problems. First, it requires substantial effort. For example, transcribing the user's operational video to a time-series table requires more than an hour, even if the video itself is only a few minutes long. Second, the meanings of a user's behaviors are difficult to infer, even if we can obtain information from the participant during the analysis stage. Because many user behaviors are unconscious, users themselves might not able to explain their behavior. Third, the descriptions tend to become ambiguous. Analyses by multiple analysts sometimes do not quite match.

We plan to (semi)automate the description process to solve a portion of these problems, especially in transcription from the recorded data to UF table. Although this tool can only support describing the time-series table from a user's operational log data, we have checked that using it reduced the required effort to some extent. We would like to improve this methodology by introducing more automation and process that require less effort.