June 2020

Saving and Recalling Interaction Sequences of In-app Actions

Victor Carbune

Matthew Sharifi

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Saving and Recalling Interaction Sequences of In-app Actions

ABSTRACT

When using apps, users can encounter difficulties in figuring out or remembering how to perform action sequences pertaining to specific tasks, especially tasks that are complicated or rarely performed. After successfully performing such tasks, it is cumbersome for the user to remember the details at a later time when they need to perform the task again. Currently, apps or devices do not provide users with the ability to capture and save these details for later recall and use. This disclosure describes techniques that, with user permission, save sequences of actions performed with an app on their device. The user can later recall the saved sequences as needed to navigate the app and/or to invoke automatic execution of the interaction sequence.

KEYWORDS

- Interaction sequence
- Action flow
- Contextual understanding
- In-app action
- App context
- Virtual assistant
- Screenshot
- Action automation

BACKGROUND

Users utilize a variety of apps on their mobile devices, computers, and other devices. When using apps, users can encounter difficulties in figuring out or remembering how to

perform action sequences pertaining to specific tasks. For examples, users may face such difficulties when performing tasks that are infrequent or for apps that are rarely used.

For instance, initial configuration of a virtual assistant may involve an action flow that includes setting up the virtual assistant to recognize the user's voice, to provide a specification of an output mode for the virtual assistant, to configure whether the virtual assistant responds to a single activation command or can engage in a back-and-forth conversation, etc. Once the virtual assistant is configured, the process may not be repeated frequently, e.g., the user may only use the action flow to modify the configuration, to set up the virtual assistant on a different device, etc. In such a case, it is cumbersome for the user to remember the details of the interaction sequence to set up the virtual assistant. Currently, apps or devices do not provide users with the ability to capture and save these details for later recall and use.

DESCRIPTION

 This disclosure describes techniques that, with user permission, save sequences of actions performed with an app on their device. The user can later recall the saved sequences as needed to navigate the app and/or to invoke automatic execution of the interaction sequence.

With user permission, as the user interacts with apps on their device as usual, the interaction sequences are stored in a cache. At any point during the interaction, the user can request that the preceding sequence of actions within the app be saved. For instance, a user may engage in a series of steps to locate a specific type of content nested within all content included in the app of an online content provider. If the same content is expected to be accessed frequently, the user may wish to save the series of navigation steps to avoid the need to search for the content again in case the steps are forgotten. The user can request saving the interaction sequence via any suitable mechanism, such as a voice command to a virtual assistant. For

instance, the user may issue the voice command: "Remember how I arrived at this content screen." Upon receiving the request from the user, the screen areas and/or buttons tapped from the start of the app till the invocation of the user request that have been cached are stored for later use.

The above operation is enabled with user permission by an on-device cache to store each step of the user's interaction with any app. If the user permits, the cache can include relevant app screenshots. When the user issues the request to save the sequence of past actions, the actions are retrieved from the cache. The cache is purged after the actions are saved upon user request or when the user leaves the app.
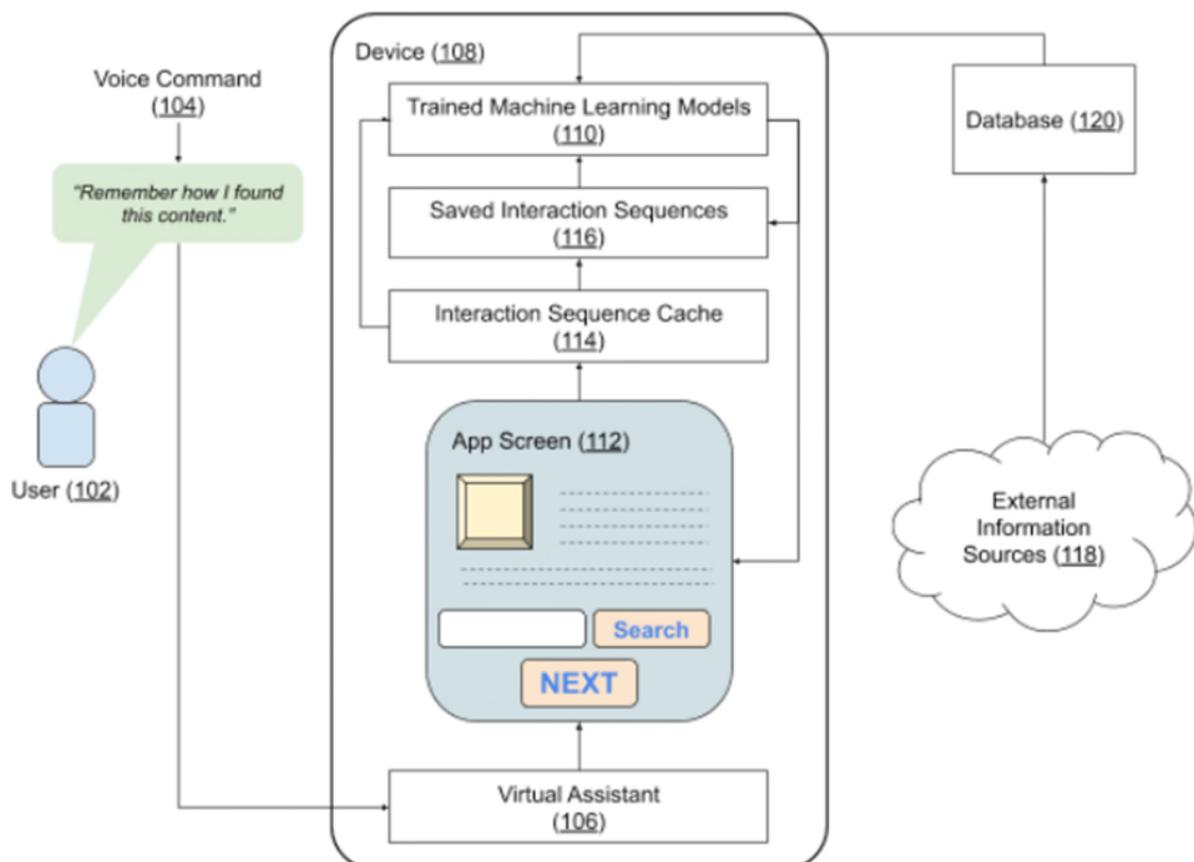
Fig. 1: Saving sequence of actions within an app

Fig. 1 shows an operational implementation of the techniques described in this disclosure. A user (102) is interacting with an app (112) on a device (108). At some point during the interaction, the user wishes to save the sequence of actions that led to the current app screen. The user issues a voice command (104) asking a virtual assistant (106) to save the interaction sequence. The sequence is then added to the set of saved interaction sequences (116) by retrieving the in-app actions of the user from the on-device cache (114) that captures each in-app action with user permission.

If the user permits, suitable trained machine learning models (110) are employed to select cached interaction sequences for automatic saving without explicit user request and/or to recommend contextually appropriate saved interaction sequences based on the user's app interactions. To that end, the models can use information from a database (120) that includes pre-processed relevant information from appropriate external sources (118), such as help pages and online videos that detail interaction sequences of various apps.

The user can later recall the saved interaction sequences at opportune times via voice commands or other appropriate mechanism. For instance, the user may ask a virtual assistant to "show me how I found the content last time." The saved interaction sequence is then shown to the user in a suitable form. For example, the user can be shown a video recording of the sequence of past actions, with the play possibly synchronized to the user's current actions in order to help the user mirror the sequence. The user can also request that the sequence of steps be performed automatically on the user's behalf. In the latter case, the actions are executed in an action-specific manner that is robust to changes in the app interface.

With permission from the user, the saved interaction sequences can be surfaced at contextually appropriate times without explicit user request. For instance, the user can be

prompted to use a particular saved interaction sequence if the user's current app interactions indicate that the user's inferred intent matches any of the currently saved interaction sequences. Suitable saved interaction sequences are identified based on considering a variety of user-permitted pieces of contextual information, such as the current app screen, previous app actions, typical app use, long pauses in interaction or repeated circular interaction sequences that indicate confusion, time of app use, location, etc. In case multiple suitable candidate sequences are identified, the user is suggested the most suitable sequence as determined by a ranking algorithm.

Instead of requesting post hoc saving a set of app actions, the user can proactively trigger app interaction recording prior to beginning the use of an app. For instance, the user can issue a voice command to a virtual assistant to "remember how I use this app." In addition, if the user permits, certain app interaction sequences can be remembered automatically if it is deemed that the user is likely to need them later. Such automatically remembered sequences can include interactions with a newly installed app, new interaction sequences within a previously used app, etc.

As shown in Fig. 1, with user permission, automatically remembered interaction sequences can be based on relevant external information sources such as app help pages, online instructional videos, etc. If the user permits, standardized geometry-based or neural network based algorithms can be employed to perform a comparison of the app screen and images extracted from external information sources and to surface cues for automatically triggering interaction sequence capture. The extraction of images from external information sources, such as videos, can be performed offline to generate a database to be used when matching with app screens is sought.

The external information sources can additionally be applied to mark up interactive elements within the currently visible app screen and/or the app screens shown when previously saved interaction sequences are shown to the user. For instance, such markup can be used to explain the function of the corresponding user interface element by delivering information extracted from sources such as the operating system.

Implementation of the techniques described in this disclosure can employ any suitably trained machine learning models. The threshold values used by the various models and algorithms used in the implementation can be set by the developers and/or specified by the user and/or determined dynamically at run time. The operation can be implemented within any platform, device, or apps. The techniques make it easier and more effective to use difficult or infrequently accessed app functionality. The described techniques can be implemented as part of an operating system that includes virtual assistant functionality, as part of a virtual assistant application, or within a particular app.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a user's social network, social actions or activities, profession, a user's preferences, or a user's current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of

a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes techniques that, with user permission, save sequences of actions performed with an app on their device. The user can later recall the saved sequences as needed to navigate the app and/or to invoke automatic execution of the interaction sequence. With user permission, as the user interacts with apps on their device as usual, the interaction sequences are stored in an on-device cache to record various steps of the user's interaction with any app. Further, app interaction sequences can be remembered automatically if it is deemed that the user is likely to need them later. The saved interaction sequences can be surfaced at contextually appropriate times without explicit user request. The operation can be implemented within any platform, device, or app and can make it easier and more effective for a user to utilize difficult or infrequently accessed app functionality.

REFERENCES

1. Carbune, Victor, Daniel Keysers, and Thomas Deselaers. "Smart assist for repeated actions." U.S. Patent Application 15/383,966, filed December 19, 2016.

2. Sharifi, Matthew, and David Petrou. "Replaying user input actions using screen capture images." U.S. Patent 9,811,352, issued November 7, 2017.