February 06, 2019

Automatic context determination for answering user queries

Ruxandra Davies

Scott Davies

Automatic context determination for answering user queries

ABSTRACT

Virtual assistant applications enable users to specify a query. When the virtual assistant receives a query without relevant context, determining the appropriate answer is not possible. In these situations, it is necessary that the user respond to follow-up questions and provide the context. This disclosure describes techniques to automatically obtain the query context, when permitted by the user, and create a seamless user experience. Based on the obtained context, the query is answered which reduces back-and-forth with the user.

KEYWORDS

- virtual assistant
- smart assistant
- query context
- user context
- context detection
- smart speaker

BACKGROUND

Assistive technologies, such as smart speakers, enable users to use voice commands to obtain information and complete tasks such as ordering products. Applications based on these technologies allow users to ask follow-up questions after receiving a response to an earlier query. However, the assistive technology lacks knowledge of relevant context when a query is initiated without a connection to a prior query. Therefore, when initiating a query, users need to generate queries that explicitly include relevant contextual information. Often times, users generate

queries while being engaged with some other application, such as a web browser or mobile app, that has contextual information for the query.

For example, a user may wish to identify the animal in a picture encountered while browsing the web. To that end, queries such as “What’s that?” or “What animal is that?” do not provide sufficient information for the assistive technology to determine an answer to the question since the contextual information needed to determine the user’s reference (e.g., “that”) is unavailable. Further, if the user is multitasking and has multiple open applications, a query such as “What’s that?” does not indicate the application context that led the user to initiate the query to the assistive technology. In such cases, the user must keep refining the original query with contextual details and specifics until the query is answered successfully by the assistive technology. The inability of assistive technologies to detect and interpret the context of the user’s engagement with other applications connected to the query results in a less than seamless user experience.

DESCRIPTION

This disclosure describes techniques to create a seamless user experience for assistive technologies via integration with the context of the user activity. Context is determined with permission from the user. The described techniques are not implemented if the user does not provide such permission. Assistive technologies can include, e.g., a virtual assistant application, available as a standalone application or integrated into a device operating system. Devices that incorporate assistive technologies include, e.g., smart speakers, smartphones, tablets, laptops, wearable devices, smart appliances, etc. The assistive technology can be built into the device and with user permission, can utilize remote servers for query processing.

When the user invokes an assistive technology with a generic query that requires knowledge of the context that led to the query, the relevant context is obtained with the user's permission. The query is then answered based on appropriate contextual information, such as activities, content, etc. If the user permits, determination of the appropriate context related to the query is based on examining multiple candidates of contextual information, such as web browser tabs, applications, user devices, etc., and various user activities and their associated timestamps. With permission from respective users, context determination can also include analyzing the activities of multiple users in order to determine the context connected to the query. If the relevant context cannot be determined with sufficient accuracy, a request for clarification is provided.

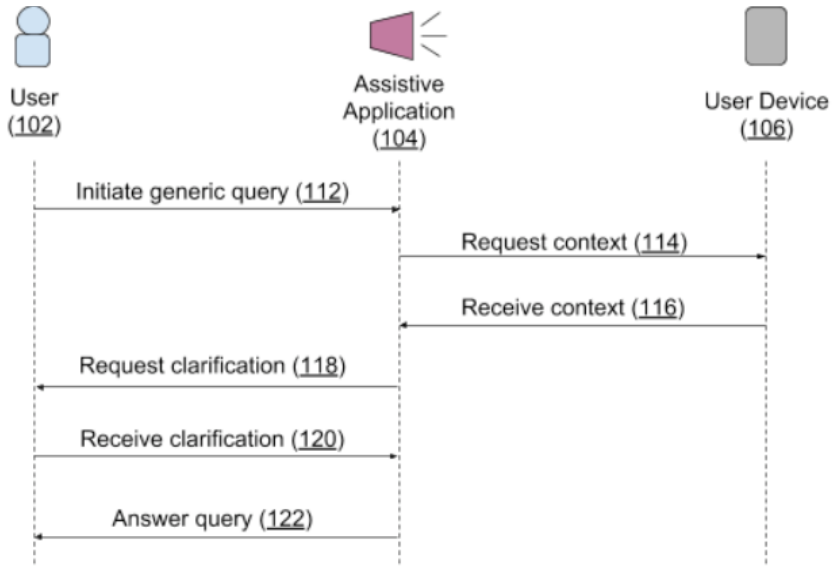

Fig. 1: Determining context connected with a generic query

Fig. 1 shows an example operational flow of the techniques to determine context and answer a query. A user (102) issues a generic query (112) to an application (104) driven by an assistive technology. The assistive application may be provided via a user device (106). Other user devices (106) may also be in use, e.g., if the query is provided to a smart speaker, it can be

detected that other devices such as a smartphone or a computer is also active and may include contextual information.

To determine the relevant context that led the user to issue the query, information regarding the user's context is requested (114) from the user device(s) (106). If the query can be connected to the appropriate context based on the received contextual information, the context is added (116) to the query and an answer to the query is provided to the user (122). In case the appropriate context for the query cannot be determined with sufficient certainty, the user is asked for clarification (118), e.g., by providing possible context options that are likely to be connected to the query. The clarification received from the user (120) is then used to add the relevant context to the generic query to determine the answer.

Example of use

Consider a user who comes across an unrecognized animal while browsing for information about rainforests and asks an assistive application, "What's that?" With the user's permission, the assistive technology obtains relevant contextual information from the user's device. For example, it is detected that the user is reading information about rainforests on a browser application on their laptop. Contents of the browser tab are examined to infer that the user's query is likely regarding the image of an animal. An answer is provided by accessing and analyzing the image (with user permission) to determine the name of the animal.

Similarly, with the user's permission, the user's context from apps other than the browser can be obtained using mechanisms such as application programming interfaces (APIs) provided by the app developer. When mechanisms are unavailable, screenshots of the most recently used apps can be analyzed to determine context. When the context includes a video, screenshots can

be timed to coincide with the user's query such that query results can be delivered without requiring the user to pause the video.

It is possible that the user is involved in multiple contexts when a generic initial query is issued. For example, the user may be listening to music while browsing the web, may have multiple tabs open in the web browser, may be using multiple applications, etc. In such cases, the context most likely to be connected to the query is identified based on appropriate heuristics, such as prioritizing video and image information over audio, giving higher importance to recently active tabs, examining the application with the most screen real estate before others, etc. The specifics of the query can additionally be used to determine the appropriate context. For example, if the user asks, "What animal is that?" contextual information that does not contain animal-related materials are discarded from consideration.

If the query and relevant heuristics are insufficient to connect the query to specific contextual information, the user can be asked for clarification, e.g., to determine the specific meaning of the generic query. Such clarifications can include appropriate options as possible choices. For example, if the user asks, "Which animal is that?" with pictures of different animals open in multiple browser windows that had been active within the last minute, the user may be asked to identify the browser window that led to the query. The user can provide the clarification by clicking or tapping on the relevant picture, circling the relevant portion or the screen, indicating the correct window via voice, etc.

If the user query is connected to a device or application that cannot be accessed by the assistive application, the user is offered the option of manually sharing a screenshot of the relevant context or, if applicable, is guided to capture the relevant context via the device camera.

The techniques described in this disclosure can be implemented on any user device that has the capability to obtain and share contextual information. Moreover, if the user permits, relevant contextual information can be gathered from multiple user devices. When contextual information is obtained from multiple devices, the processing of the contextual information to select the appropriate contextual information connected to the user's query can be done on a particular device, e.g., the device that received the query, or can be performed by a server. Determination of the appropriate contextual information connected to the user's query can be refined by the application of machine learning techniques.

The techniques of the disclosure can be extended by obtaining relevant contextual information from devices connected to the same network, with permission from users of the respective devices. If the users permit, such operation can gather relevant context across multiple users signed-in on devices connected to the same network, such as a home WiFi network shared by the members of a household. Such context can be particularly useful when the device that receives the query is a shared device, e.g., a smart speaker in the household.

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

Virtual assistant applications enable users to specify a query. When the virtual assistant receives a query without relevant context, determining the appropriate answer is not possible. In these situations, it is necessary that the user respond to follow-up questions and provide the context. This disclosure describes techniques to automatically obtain the query context, when permitted by the user, and create a seamless user experience. Based on the obtained context, the query is answered which reduces back-and-forth with the user.