November 2021

Implicit Invocation of Device-specific Custom Actions Using a Virtual Assistant

Amit Singhal

Yao Lin

Implicit Invocation of Device-specific Custom Actions Using a Virtual Assistant

ABSTRACT

This disclosure describes techniques to augment the built-in grammar of a general-purpose voice-activated virtual assistant with custom commands and actions that are specific to a particular make and model of a device such as an in-vehicle system. The techniques enable the virtual assistant to service implicit commands that make no reference to vehicular make by providing custom, make-specific responses. A single, consistent grammar enables the general-purpose virtual assistant to handle device-specific commands without the user having to provide a query that includes a device-specific invocation phrase.

KEYWORDS

- Virtual assistant (VA)
- Voice command
- Spoken command
- Grammar augmentation
- Implicit invocation
- Query pattern
- Automobile assistant
- In-vehicle system
- Vehicle infotainment system
- Hardware abstraction layer (HAL)

BACKGROUND

Many devices can be controlled by issuing voice commands to a general-purpose virtual assistant. Such devices include, e.g., smart speakers, smart televisions, smart appliances, in-vehicle systems, etc. Many voice commands, such as “set the temperature to 70 °F” are common across many devices (e.g., home thermostats, in-vehicle systems, etc.) and are independent of all makes and models of the device or vehicle.

However, vehicles (or other voice-controllable devices or equipment) from various manufacturers can have certain commands and actions associated uniquely with their makes. For

example, a first manufacturer might specify custom voice commands to enable control of air quality; a second might specify custom voice commands to test tail-light (or other module) functionality; a third might specify custom voice commands to activate and configure a seat massager; a fourth might specify custom voice commands to activate and configure one-pedal driving; a fifth might specify custom voice commands to query car status or enquire about the date of next service; etc. The commands for any particular task can be expressed in several ways. For example, a command to test a module might take the form, “test the taillights,” “are the taillights working?”, “are the taillights OK?”, “check the rear passenger-side taillight,” etc. Therefore, a custom voice command effectively entails augmenting the grammar or recognizable query patterns of the virtual assistant.

Vehicle (or other voice-controllable device or equipment) manufacturers can specify unique responses, also known as fulfillments or actions, to commands. For example, a query relating to the date of the next service can be responded to by one manufacturer by looking up the odometer and by another by looking up the oil and brake-fluid levels.

As mentioned earlier, vehicle (or other voice-controllable device or equipment) manufacturers often want to enable *unique* voice commands and actions for their makes, e.g., voice commands that are understood and acted upon only by their models, not by those from other manufacturers. While general-purpose voice assistants can provide built-in common actions that work across makes and also the framework for diverse manufacturers to create unique actions (with custom grammar and fulfillment logic), grammar added by one device manufacturer needs to be invoked only for queries from vehicles by that manufacturer. However, it is important that the user be able to use a single, consistent voice interface, grammar, and wake-word across vehicle makes without having to explicitly ask to connect to make-specific

custom actions. For example, for a car of make A that uniquely supports seat-massager voice-commands, a user should be able to use the wake-word associated with the virtual assistant followed by a command such as “start lumbar massage,” rather than “talk to Make A to start lumbar massage.”

Although there exist techniques that enable third-party developers to augment general-purpose virtual assistants for purposes such as enabling users to access third-party conversational actions, web content, apps, smart home devices, etc., third-party voice commands need to be explicitly invoked with the name of the third-party app or device. For example, a user has to say “book a taxi on taxi-company A.” These techniques don’t offer *implicit* command invocation regardless of how the manufacturer chooses to fulfill the command, e.g., app-based, vehicle hardware abstraction layer (VHAL) based, etc.

DESCRIPTION

This disclosure describes techniques to augment the built-in grammar of a general-purpose, voice-based virtual assistant with custom commands and actions that are specific to a given vehicular (or other voice-controllable device or equipment) make or model. The techniques enable a user to issue *implicit* (e.g., with no reference to vehicular make or model) commands to the virtual assistant using a single, consistent grammar that evokes custom, make-specific responses. The techniques are applicable generally to vehicles, to devices (e.g., smartphones, smartwatches, fitness trackers, smart home devices, etc.), to equipment (e.g., smart appliances such as treadmills, washing machines, printers, thermostats, etc.), etc. that are equipped to respond to voice commands.

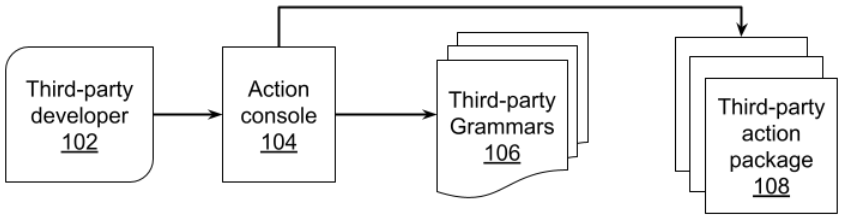

Fig. 1: Creation of actions (which include grammars and fulfillments) by third-party developers

As illustrated in Fig. 1, a third-party developer (102), e.g., a vehicle original equipment manufacturer (OEM), is provided with an interface, referred to as a developer or action console (104), that enables them to input and develop third-party query patterns or grammars (106) and a corresponding third-party action package (108). Sample query patterns provided by the OEM can be used to build natural language understanding (NLU) models. The built-in grammar of the virtual assistant is augmented by a collection of third-party grammars and action packages submitted by different third parties.

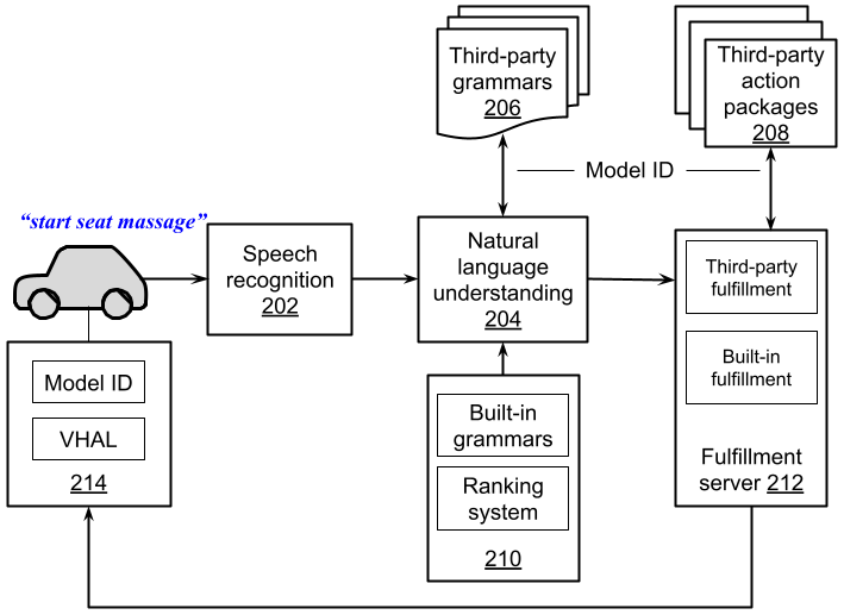

Fig. 2: Implicit Invocation of Custom Actions in a Vehicle using a Virtual Assistant

Fig. 2 illustrates implicit invocation of custom actions in a vehicle (or other voice-controllable device or equipment) using a voice-activated virtual assistant. A voice command

issued by a user is recognized by an automatic speech recognizer that implements speech recognition (202) techniques. The command intent is determined by a natural language understanding (NLU) processor (204). In determining intent, the processor draws upon both the built-in grammars (210) and the store of third-party grammars or NLU models (206) indexed by the ID (e.g., make, model) of the vehicle (or other voice-controllable device or equipment).

For example, only third-party grammars or query patterns associated with the current vehicle's make are surfaced. Based on both built-in and make-specific NLU models, the processor ranks interpretations of user intent and generates a winning interpretation, which is sent to a fulfillment server (212). The fulfillment server (212) draws upon third-party action packages (208) specific to the vehicle model as well as built-in fulfillment packages (that are applicable across makes) to construct a fulfillment request, e.g., a software command that corresponds to the user's intent. The fulfillment server communicates the command to the vehicle hardware abstraction layer (VHAL, 214) to execute the user's command. Alternative to fulfillment via the VHAL, fulfillment can also be app-based or cloud-based. Thus, to trigger a seat massage (an example make-specific action) users can simply say "start seat massage" without referring to the make, rather than having to utter "talk to Make 'A' to start my seat massager."

OEMs can define the fulfillment of a custom vehicle action via direct VHAL application programming interfaces (APIs). Arguments from the NLU processor can be translated to VHAL property metadata for execution. The OEM can also define the voice responses (e.g., in text format) along with the action to be performed. When the responses are defined in text format, text-to-speech technology can be used to generate a spoken response from that text. In this manner, the disclosed techniques enable automobile manufacturers to develop custom voice

commands and actions that are directly invocable and uniquely executable on vehicles of their make using general-purpose virtual assistants.

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 the collection of user information (e.g., information about a user's spoken commands, vehicle (or other voice-controllable device or equipment) make/model, 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 to augment the built-in grammar of a general-purpose voice-activated virtual assistant with custom commands and actions that are specific to a particular make and model of a device such as an in-vehicle system. The techniques enable the virtual assistant to service implicit commands that make no reference to vehicular make by providing custom, make-specific responses. A single, consistent grammar enables the general-purpose virtual assistant to handle device-specific commands without the user having to provide a query that includes a device-specific invocation phrase.