Technical Disclosure Commons

Defensive Publications Series

October 08, 2019

Automatic grammar enhancement for virtual assistant

Tao Feng

Hongjie Chai

Automatic grammar enhancement for virtual assistant

ABSTRACT

Grammar rule generation for virtual assistant applications is difficult to scale due to the need for manual labeling. This disclosure describes a scalable solution to automatically generate enhanced grammar rules in predefined verticals. The techniques enable incremental improvements to interpretations of voice commands by a virtual assistant application. Queries that have not already been processed for a target vertical are identified and extracted from a corpus of user queries, e.g., a time-limited corpus in a particular language. Queries are analyzed to discover arguments and patterns that are specific to the vertical. Grammar rules are generated based on the arguments and patterns.

KEYWORDS

- Natural Language Understanding
- Virtual assistant
- Voice assistant
- Grammar rule
- Machine learning
- User query
- Query argument
- Smart speaker

BACKGROUND

The use of virtual or intelligent personal assistants, e.g., activated by a voice command, are popular for various tasks such as setting reminders, playing media, etc. Such assistants utilize grammar rules to determine the action that is to be performed in response to a user voice query which can be in a natural language form, e.g., "Set alarm for 5 pm," "Play song X," "Turn off lights," etc. For example, the query "Play song X," matches a pattern of a command ("Play") followed by a media title ("Song X") and in response, playback of the requested song is triggered as the target of the command. Grammar rules are advantageous due to stability and low latency in processing.

Development of grammar rules relies on skilled individuals, e.g., linguists, to manually label user queries. Individuals that develop grammar rules is required to possess knowledge of the language of the query as well as of the natural language understanding (NLU) techniques utilized in the virtual assistant. The volume of grammar rules needed to respond to user queries increases when users issue queries in a greater number of languages. Manual labeling to develop grammar enhancements for individual languages is unscalable due to the cost and effort involved.

DESCRIPTION

Techniques are described for scalable generation of grammar rules for NLU of commands received by a virtual assistant or in other contexts. In particular, grammar rules are generated specific to particular verticals or domains. The described techniques can eliminate the need for manual labeling to develop the grammar rules. The grammar rule generating techniques are applicable to the target vertical area in any language. The target vertical area is predefined to include any domain of knowledge, such as media, healthcare, education, etc.

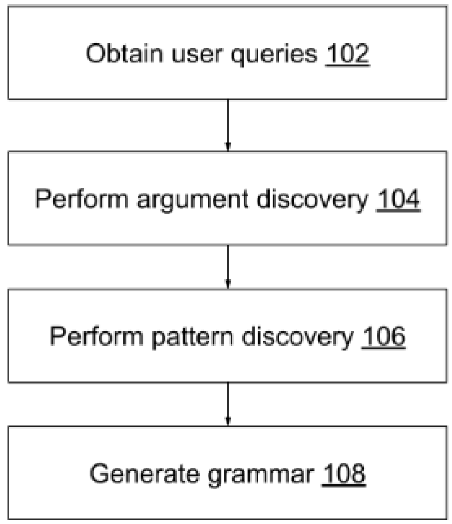

Fig. 1: Example grammar rule generation method

Obtain user queries (102): A data extraction model is utilized to extract user queries, e.g., voice queries received by a virtual assistant application, from a given language that are not processed as vertical-specific queries, e.g., media queries. The queries may be extracted based on a configurable time window, e.g., a one-week window, a two-week window, etc. Queries extracted for grammar generation are limited to users who have provided permission for such use of queries. In the extraction process, machine learning models can be leveraged to exclude user queries which are not vertical-specific queries.

Argument Discovery (104): Vertical-specific query arguments are identified within the extracted queries. For example, if the vertical is media, arguments such as song/album/artist/playlist/video/movie that are present in the user queries are identified. Identification of arguments can be based on vertical-specific signals, such as web annotations, search results, user click distribution, etc. An iterative approach can be used to discover more free text arguments that do not have web annotations. For example, consider that an artist

"Alice" has a web annotation as an "Artist," and a pattern of "Play the song of $artist" is determined from the query "Play the song by Alice." If additional queries of the same form, e.g., "Play the song by Bob," are observed, where "Bob" is not associated with a web annotation, it is determined (with low confidence) that Bob is also an entity of the type "Artist."

Pattern Discovery (106): Based on arguments discovered in the argument discovery phase, the actual arguments in the user queries are replaced with annotators such as $artist, $song, $album, $video, $playlist. After replacement, the patterns are analyzed to determine the frequency with which certain patterns, e.g., "play $song," occur in the extracted queries. If the frequency (or absolute count) of certain patterns exceeds a threshold, the pattern is identified as a valid pattern.

Grammar Generation (108): Two types of signals are utilized to determine query intent. The first signal is the prediction of machine learning models based on the input query. The second signal includes vertical-specific heuristics that takes into account the identified arguments.

The combination of intent and pattern is used to generate grammar enhancement. The described techniques are vertical-centric since unique heuristics are defined for individual verticals such as media, communication, navigation, etc. to discover arguments, patterns, intents and to generate the final grammar enhancement. The techniques are also local-agnostic, since once defined, the heuristics can be used to generate grammar enhancements for the same vertical in any language. The described techniques enable improvements to voice assistant applications, e.g., time and resources needed to process a user query, and enable the applications to support user queries in more languages.

The described techniques can be implemented as part of a virtual assistant or voice assistant application, e.g., that executes on a server or client device. The techniques can improve responses to voice queries received on any device, e.g., a smart speaker, a smart appliance, a computer, a smartphone, etc.

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 a scalable solution to automatically generate enhanced grammar rules in predefined verticals. The techniques enable incremental improvements to interpretations of voice commands by a virtual assistant application. Queries that have not already been processed for a target vertical are identified and extracted from a corpus of user queries, e.g., a time-limited corpus in a particular language. Queries are analyzed to discover arguments and patterns that are specific to the vertical. Grammar rules are generated based on the arguments and patterns.