Abstract

This research focuses on natural, robot-led, human-robot interaction that enables a robot to discover what drinks a bartender can prepare through continuous application of speech recognition, understanding and generation. Speech was recognised using Google Cloud's speech to text API, understood by matching either the object or main verb of a sentence against a list of key words and, finally, generated using templates with variable parts. The difficulty lies in the large quantity of key words, as they are based on the properties of the ordered drinks. The results show that having the aforementioned interaction works well to some extent, i.e. the naturalness of the interaction was ranked 5.5 on average. Furthermore, the obtained precision when identifying the unavailable drinks was 0.625 and the obtained recall was 1, resulting in an F1 measure of 0.769.