

コミュニケーションロボットの研究
—API を用いた言語処理—
A study of communication robot
-Language processing using the API-

○小林文俊¹, 入江寿弘²
 Fumitoshi Kobayashi¹, Toshihiro Irie²

Abstract: The current high efficiency computer is said to exceed far human brain, but it is the present conditions that processing capacity for the uncertain information such as a sound or the image is inferior to the teachability that a human being learned from experience. The computer which approached for human intelligence was not readily done, but it was with the level that human intelligence was near by the longtime study by the engineer of the whole world. Because the spread of robots which a considerable labor shortage was concerned about in society in the future, and were equipped with AI as a substitute for one of a human being was anticipated, I have begun to want to study the language processing in that.

1. はじめに

今現在の高性能コンピュータは人間の脳をはるかに超えるといわれているが、音声や画像などのあいまいで不確かな情報に対する処理能力が、人間が経験から学んだ学習能力には及ばないのが現状である。人間の知能に迫るようなコンピュータはなかなかできなかったが全世界の技術者による長年の研究により人間の知能に近いレベルとなってきた。また、最近ではインターネット通信インフラの普及と共に、多くの人工知能はクラウドサービスの形としても提供されている。

本報告では、クラウドサービスを使った会話応答システムの実験結果を報告する。

2. 人工知能自然会話

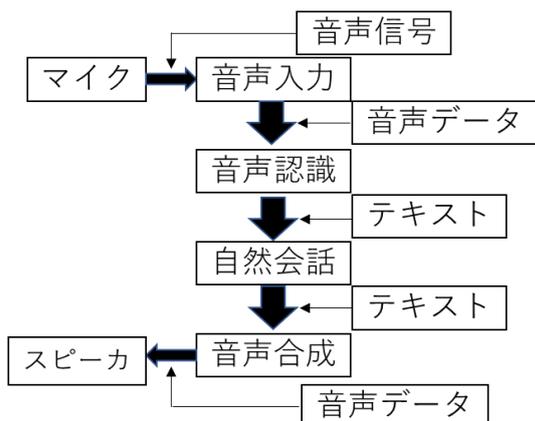


Figure 1. 人工知能会話

コンピュータとの自然会話は人類の長年の夢であり、それを現実にすべくコンピュータに人間の知能を持た

せる人工知能技術が開発されている。コンピュータが人間と会話するためには4つの機能が必要といえる。

(1) 音声入力

人間が話す音声信号を音声データとして記録することであり、マイクで音声入力信号を電気信号へ変換しコンピュータに決められた音声ファイル形式で入力された信号を音声ファイルとして記録して使用する。

(2) 音声認識

音声認識(Speech Recognition/Voice Recognition)技術は、音声データを文字テキストに変換すること。STT(Speech-to-Text)と呼ばれて、音声認識技術は人工知能技術の1つであり、最近の人工知能技術は雑音の多い環境や曖昧な発音でも音声をほぼ正しく認識できるようになってきている。

(3) 自然会話

人間との自然会話の成立には、聞いた言葉の意味の理解と返事の作成が必要であり、コンピュータが話しかけた言葉の意味が理解できれば返事の作成は難しい。しかし、音声認識の完成度の高さに比べ、人工知能技術の完成度は低く人間の言葉の理解はいまだに発展途上である。

(4) 音声合成

音声合成(Speech Synthesis/Voice Synthesis)とは、テキストを音声で読み上げる機能であり、TTS(Text-to-

