

人間が手作り→認識AI→生成AI… 進化する対話システム

中村 雅之

● コンピュータと人を対話させるチャットボット

チャットボットは、人間とコンピュータとの間で短いテキストを通じて会話を行うソフトウェアです。一般的には、人間のテキスト入力に対して、まるで中の人がいるかのような気の利いた応答をすることがチャットボットの特徴だと思われているでしょう。

しかし、チャットボットが実現する価値の重要な要素として、応答他に入力の自由度があります。入力の自由度が高いUI(ユーザ・インターフェース)とは、人間にとって自然かつ豊かな表現でコンピュータへ意思表示できるものです。そのようなUIを、ここでは優しいUIと呼びます。

本稿では前半で、チャットボットの基本的な仕組みについて紹介し、生成AIの登場前後でチャットボットがどう変わったのかを簡単に解説します。そして、生成AI型のチャットボットの技術を使うことで、自然言語による対話以外の優しいUIを実現する可能性について検討します。

● ボタンだけのデバイスでも対話できるようにする

後半では、チャットボット以外の優しいUIへの理解を深めるため、簡単なUIの作成に挑戦します。チャットボットは便利ですが、テキストや音声の入力機器を必要とします。入力方法が限られたデバイスでは、優しいUIを別な方法で提供する必要があります。

本稿では、入力方法が限られたデバイスとして自動販売機シミュレータをウェブ・アプリケーションとして作り、実際に操作できるUIを実装します。自動販売機の購入ボタンやお釣り返却ボタンを使って、人間は自販機と対話できるでしょうか。生成AIを使ったUIにどんな可能性があるか、実験で探ってみます。

チャットボットの基本的な仕組み

チャットボットにおけるAIの役割を説明するために、まずはチャットボットの基本的な仕組みについて説明します。単純に考えると、チャットボットは、テ

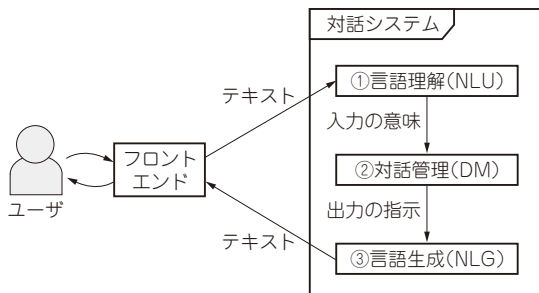


図1 従来の対話システム(チャットボット)のモデル

キストを入力すると、テキストを出力するシステムだと言えます。入出力のUI(フロントエンド部)は、例えばウェブ・アプリケーションだったりLINEアプリケーションだったりします。そういったフロントエンド部分を除いたチャットボットの本質的な部分を表すモデルを図1に示します。このようなモデルに基づき、それぞれの部品を用意して結合させたシステムとして開発されてきています。

▶①言語理解

(NLU : Natural Language Understanding)

ユーザの入力テキストから意味を抽出する部分です。意味とは、ユーザがその発言をした意図や感情などの情報のことです。例えば、「こんにちは!」というテキストからは、意図としてあいさつ、感情としてポジティブなどといった意味を抽出できます。

▶②対話管理 (DM : Dialog Manager)

ユーザとの話の流れを管理して、応答すべき内容を決める部分です。最も単純な例としては、あいさつされたらあいさつを返すというルールで表されます。

▶③言語生成 (NLG : Natural Language Generation)

ユーザに応答する内容を言語化する部分です。例えば、あいさつに応答するとしたら、「はい、こんにちは!」などのテキストを出力します。言語生成といっても、必ずしも生成AIによるものではなく、人間が書いたテキストを元にする場合も含まれます。