

教えて！いろはちゃん AIはどうやって言葉の意味を理解するの？

ロボマインド 田方 篤志*

*たかた あつし…代表取締役 ドラえもんの心を開発する町の発明家。自身が思いついたアイデアを実現するため、どこにも属さず20年間研究し続けた結果、今主流のAIとは異なる方向に進化し、意識を持って普通に会話ができるAI創造の着想を得る。研究開発ブログは、のべ100万以上のページビューを記録。
URL : <https://robomind.co.jp/>

意味理解の呪い

知らない女子高生が俺の家にいる。

いや、問題はそこじゃない。

上半身しかない。

いや、床から女子高生が生えているのだ。

「あんた、さっきから何、ジロジロ見てるのよ」

「うわっ、しゃべった！」

「当たり前でしょ。さっさと、初期設定、終わらせなさいよ」

目の前に「初期設定」と書かれたボタンが宙に現れた。なるほど、これがARグラスかあ。眼鏡のレンズに映像を投影することで、現実世界に3Dバーチャルの映像が重なって表示される。まるで、現実世界にバーチャル女子高生がいるようだ。手を伸ばして指先でボタンに触れると、ボタンが光って画面が切り替わった。

「なるほど、指先の位置を画像認識で割り出して、ボタンを押したことを判定しているんだな」

眼鏡に取り付けられたカメラで撮影した画像をリアルタイムでポリゴン化し、ポリゴン化した床の上にバーチャル・キャラクターを配置するののか。どうやら、配置するときの座標がずれて下半身が床の下に埋められたようだ。矢印ボタンを押してキャラクターの高さを調整すると、つま先まで床から出てきて床の上を器用に歩き始めた。

「やっと歩けるようになったわ」と言いながらうれしそうに部屋を歩き回る。

「あっ、そうだ。私は百瀬いろは。この度は、エビスグラスをお買い上げいただき、誠にありがとうございます」と、ぴよこんと一礼する。

「あっ、僕は一ノ瀬仁。エビスグラスが半額セールだったから、つい、買ってしまったんだよ」

「ええ～、私ってセール品なの。エビスデンキの主力製品だってエビス社長も言ってたのに～。うっ、うっ、うっ」

「まあ、まあ、泣かないで。エビスグラスはすごいと思うよ。大阪万博2025の公認グラスの最終候補まで残ってたしさあ」

「結局選ばれなかったじゃない。選ばれると思って大量に作ったから半額セールになったんでしょ。うっ、うっ、うわ～ん」

「あっ、いや、でも、いろはちゃんは、とってもかわいいよ」

「ぐすん、ほっ、本当……？もっと、ほめてよお」

「え〜っと、そうだ、自然な会話できるところがマジすごいよ。いろはちゃんって、AIなんでしょ」

「うん。ぐすん……。もしかして、AIに興味ある？」

「うん、今、大学で人工知能の勉強をしてるとこなんだよ」

「私、AIには詳しいのよ。AIのことなら、何でも聞いて」

「それじゃあ、自然言語処理について教えてよ。いろはちゃんて、自然言語処理を使って会話してるんでしょ」

「そうよ。仁君は自然言語処理について、どれだけ知ってるの？」

「確か、最初にするのが形態素解析だったよね」

「そう。文を単語に区切るのが形態素解析。『太郎は、学校に行きました。』という文の場合、『太郎/は、/学校/に/行き/まし/た/。』って感じに分けるわよね」

「次にするのが、係り受け解析でしょ」

「そう。『太郎』は『行く』に係るとか、『学校』も『行く』に係るとか、係り受け関係を解析するの」

「最後は、意味を理解する意味解析をして終わりでしょ。意外と簡単だよ」

「えっ、あんた、今、何て言った？」

「な、何、怖い顔してんだよ」

「意味理解って、そんなに簡単じゃないのよ！自然言語処理は、長年、意味理解の呪いに縛られてきたのよ！」

「大げさだなあ……って。えっ、なっ、何だ？」

部屋がにわかになんて暗くなってきた。よく見ると部屋の中に井戸らしきものができている。と思う間に、井戸の中から髪の毛の長い女の人が出てくるではないか。長い髪を前に垂らして顔が見えない。

「さっ、貞子……。止めてくれ！俺、そういうの苦手なんだ……」

そう言うと、部屋が明るくなって、元のいろはちゃんに戻った。

「ただ、仁君が、あまりにも意味理解の恐ろしさを知らないから、わからせたかっただけよ」

「でも、意味理解の何が難しいんだよお。言葉の

意味って、オントロジーって言うのをを使うだけでしょ」

「オントロジーは、単語の意味を単語同士の関係で整理したものよね。たとえば、『概念』の関係だと、食品概念の下に麺類があって、麺類概念の下にラーメンがあるとか。『全体-部分』の関係だと、自動車全体に対して、タイヤやハンドルは部分の関係とか」

「そう、そのオントロジーの情報を単語に追加すれば意味理解ができるじゃん」

「単語に情報を追加して何がわかるの？文の情報は増えるけど、じゃあ、その文の意味ってなに？その文で言いたいことって何？」

「そっ、そんな立て続けに質問されてもわからないよ」

「係り受け解析まではできても、肝心の文の意味が理解できないから、第2次AIブームは終わってしまったの。これが意味理解の呪いよ」

「今は、第3次AIブームでしょ。たしかディープラーニングでAIブームが復活したんだよね。じゃあ、ディープラーニングで意味理解の呪いが解けたってことじゃないの？」

「そんな簡単な話じゃないわ。その話の前に、AIの2つの派閥について説明しとくわ」

「2つの派閥って、『きのこの山』派と『たけのこの里』派みたいなもの？」

「全然違うわ。1つはコネクショニズムで、もう1つは記号主義よ」

「どっちが『きのこの山』派で、どっちが『たけのこの里』派？」

「だから、全然違うって言ってるでしょ！コネクショニズムって、ニューラルネットワークを使って学習する手法で、大量のデータを統計処理して学習するタイプよ。記号主義は、人が作ったルールに従って記号を処理する手法で、係り受け解析やオントロジーを使うタイプよ。AIの歴史は、この2つの派が競い合って発展してきたの」

「ディープラーニングって、たしかニューラルネットワークを多段階にしたって聞いたけど、ディープラーニングがコネクショニズムってことだよ」

「そうよ。今の第3次AIブームはコネクショニズムが中心なの。第2次AIブームは記号主義が中心

だったのよ」

ディープラーニングの方法

「それじゃ、第3次AIブームの立役者のディープラーニングについて詳しく教えてよ」

「ディープラーニングがネコ概念を獲得したって話は知ってる？」

「確か、グーグルがYouTubeの大量の画像を学習させたって話でしょ。教えなくても勝手にネコの画像を認識できるようになったとか」

「そう。1,000台のコンピュータを3日間走らせて1,000万枚の画像を学習させたの」

「一体、ディープラーニングは何を計算してたの？」

「最初は画像の細かい部分に対して、似た形をまとめるの。直線とか丸といった風にね。次は、それらを組み合わせて、さらに大きな部分で似た形をまとめるの(図1)。そうすると、動物だと目や耳の形が現れてくるの。さらに目や耳をまとめると顔の形にまとまるの。そうやって、最も標準的なネコの画像(図2)が得られたってわけ。それがネコ概念だって、当時、話題になったの」

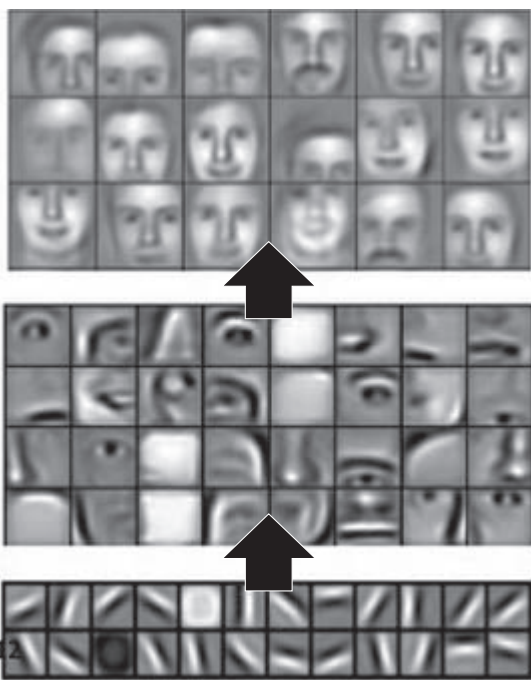


図1 ディープラーニングの画像処理プロセス

「でも、ネコを認識できるってどういうこと？ 3日もかかってちゃ、あんまり使い道ないんじゃないの？」

「時間がかかるのは学習する段階だけよ。学習が終われば、あとは、適当に画像を読み込ませれば、ネコが写っているかどうか一瞬で判断できるのよ」

「そうか。グーグルフォトに写真をアップすれば、自動でネコとか人物って分類してくれるのは、ディープラーニングを使っているのか。なるほど。それじゃ、どうやってネコって判断してるの？」

「それは、耳の形とか目の形の特徴で判断するの。つまり、ネコの特徴を自動で学習したってことよ」

「すごいなあ。まるで人間と同じじゃない。う～ん、ちょっと待って。ネコの顔を耳とかの部分でまとめてるって、さっきのオントロジーの『全体一部分』の関係と似てない？」

「いいところに気が付いたわね。ディープラーニングは特徴を取り出すのが得意なの。ただし、どんな特徴を取り出すかは、プログラムで最初から決められているの」

「どういうこと？」

「今の場合だと、最初は小さい部分で整理して、次は、その小さな部分を使ってもう少し大きな部分を整理したでしょ。つまり、大きな部分は小さな部分でまとめるようにプログラムされているの。だから、結果的にオントロジーの『全体一部分』の関係と同じになるのよ」

「じゃ、長い時間かけて何を計算してたの？」

「部分はどんな形が最適かといったことを計算してるの。たとえば、ネコの耳はとがってるとかよ」

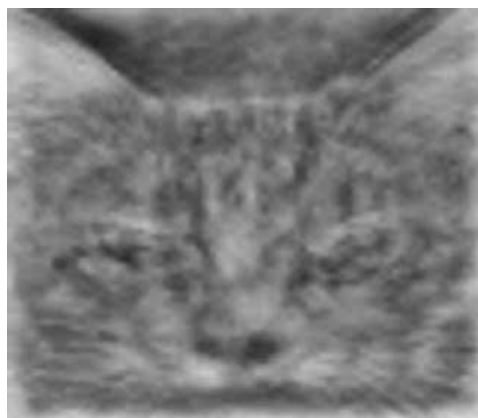


図2 ディープラーニングで獲得したネコの画像