



数理の窓

「言語ゲーム」を高次元空間に実現する

『2人がチェス盤に向かっている。チェスをして
いるように見える。しかし、実際は、悲鳴や足踏み
に対応して、駒を動かしているのかも知れない。・
(中略)・そんなはずがないと思うだろうが、私た
ちは常識や経験をもとにしているだけで、2人が別
の「ルール」に従っている可能性を否定できない。』

これは、ウィトゲンシュタインの後期哲学「哲学
探究」の一節を元にしてしている。彼は「他者への命
令」「しりとり」「演劇」など言語活動全般を「言語
ゲーム」と呼び、その本質は、自己や他者への働き
かけとした。そして、言語の意味は、人々の間に共
有された慣習内の文脈や使い方によって決まるもの
と考えた。つまり、他者と“会話が成立”するの
は、言葉の指す事実が両者で一致しているからでは
なく、語の用法が共有されているからだ。

これは、彼の前半の哲学「世界は事実からなり、
言語はその写像」、「文の真偽は、実世界の対応して
いる事象で決まる」を否定している。そして「言語
は世界の写像」VS「言語ゲーム」の違いは、自然
言語処理の手法にも表れる。

後者に対応した“パラグラフベクトル”と呼ばれ
る手法がある。これは、単語同士や文同士の近さ
を測るのに、2つが同時に出現する文や文書の頻度
データを使う。たくさんの文書をAIに学習させると、
300次元のベクトル空間¹⁾ができる。各単語

はベクトル空間上の1点になり、2つの単語間の距
離を計算することで、意味の近さを測ることができ
る。例えば「スマホとPCの類似度は0.7」や「砂
糖＝醤油－しょっぱい＋あまい」という関係が成立
つ。つまり、単語の意味は、その点の絶対位置でな
く、他の単語や文との相対関係においてのみ説明で
きる。学習文書によってAIのベクトル空間が変わ
るように、人間一人一人も脳に各々のベクトル空間
を持っており、言語ゲームを共有する2人は、それ
が似ていると言うことができる。

一方、前者の「言語は世界の写像」の手法は、単
語や文を論理式に変換し、論理形式の集合として言
語を扱う。単語や文を、論理演算という代数構造を
持つ空間にマップするともいえる。この中で言語の
意味は“確定”している。多くの知識を空間に入れ
ていけば、推論や文の真偽が決定できるため、シス
テム処理に向いている。2つの手法は相補的だが、
現時点での融合は難しく、大きな研究課題だ。

ちなみに我々は、会社やプライベート、趣味、
ネット内など様々な“言語ゲーム”を移動しながら
生きている。それぞれに“私”がいるのだが、果た
して、それらが同一の“私”であるのは自明なのだ
ろうか？

(外園 康智)

1) 300次元である必然はなく、モデルの作り方に依存す
る。学習モデル作りには深層学習を使う。