

〈教育ノート〉

# 生成AIと戯れた日々 2023

藤崎弘士\*

The days with generative AI 2023

Hiroshi FUJISAKI\*

## 1. はじめに

2022年は midjourney, stable diffusion などの画像生成 AI が研究者だけでなく一般人にも利用できるようになり、また、年末には ChatGPT に代表される文章生成 AI も華々しく登場した。それから1年ほどが過ぎ、現在は2023年12月であるが、**生成AI (generative AI)** の進歩は留まるところを知らず、OpenAI 社のサム・アルトマン CEO の解任劇（その後の復職まで1週間ほどだったが）の際に噂されたように、汎用化人工知能 (Artificial General Intelligence, AGI) への道筋までが照準に入ったものと考えられている (AGI については最後に触れる)。

生成AIの登場によって、これまで「AIによって仕事を奪われる」と曖昧に言われていたことが、かなり明瞭になった。また、生成AIを利用することで、仕事の効率が上がった企業も多いという話も聞く。研究者にとっても研究スタイルがかなり変わるほどの衝撃であり（とは言え、研究者の中でもまだ使っていないという人もいるが）、教育においてもその効果や影響は絶大であろう。しかも、これらがたった1年の間の変化であるということが恐ろしい。

本稿では、ここ1年 (2023年1月ごろから12月まで) の生成AIの動向や、個人的な ChatGPT の利用経験、生成AIを用いた教育の現状と今度の動向などについて雑駁に述べていきたい。

---

\* 日本医科大学・物理学教室/数理データサイエンス AI 教育センター, Department of Physics/Center for Mathematics, Data Science, and Artificial Intelligence Education, Nippon Medical School

## 2. 生成 AI とは

### 2.1. 人工知能と統計学の復習

筆者は人工知能と大学（特に医科大学）教育については別のところで既に論じている [1, 2] が、ここでは別の観点から人工知能とは、また、生成 AI とは、ということについて触れよう。

**人工知能 (artificial intelligence, AI)** に関しては、人間の意識や知性のようなものをシミュレートするという壮大な計画も含まれており、それは AGI の議論にも受け継がれているが、現在の大部分の AI は「ある特定のタスク」をデータ駆動式に解くアルゴリズムということであり、これは**機械学習 (machine learning, ML)** と呼ばれる技術のことである。ML の目的は、データの**分類 (categorization)** と**回帰 (regression)** に大別されることが多いが、これに今では**生成 (generation)** が加わっている。そして、これらの背後にある最も基本的な概念は多次元の確率分布であり、これはまさに**統計学** [3]、または物理学では**統計力学** [4] のメインテーマである。つまり、式で表すと、

$$P(x_1, x_2, x_3, \dots, x_N) \quad (1)$$

を求めるということに他ならない<sup>2</sup>。ここで、 $x_1, x_2, x_3, \dots, x_N$  は観測され得るデータ空間が  $N$  次元であるということである。例えば、2次元の画像であれば、これらはピクセルの数を表し、また、文章であれば、トークン（単語のようなもの）の数に相当する。これらは容易に何万から何百万次元になる。 $x_i$  は画像であれば、輝度を表し、それは連続変数であり（とはいえ計算機に入力する場合は離散変数になる）、文字であれば、“dog”, “cat” などの離散値である。

確率分布で表されるということは値が確定していないということであり、例えば  $x_1$  を  $K$  回観測すれば、 $x_1^{(1)}, x_1^{(2)}, \dots, x_1^{(K)}$  という別々のデータが得られる。これは統計学でいうところのサンプルである。これは上の例でいうと、2次元の画像が大量にある（犬や猫、建物、人間の顔など）ということや、様々な文章がある（英語や日本語で書かれた文章やメール、X(旧 Twitter) の文字列など）ということに対応する。

---

<sup>2</sup> $x_1, x_2, \dots, x_N$  が連続変数だとすると、 $x_1, x_2, \dots, x_N$  と  $x_1 + dx_1, x_2 + dx_2, \dots, x_N + dx_N$  で指定される多次元の微小体積の中に  $P(x_1, x_2, x_3, \dots, x_N) dx_1 dx_2 \dots dx_N$  の確率を割り振るということである。

もしこの確率分布を以下のように分割して表すことができれば、話は簡単になる。

$$P_1(x_1) \times P_2(x_2) \times P_3(x_3) \times \cdots \times P_N(x_N) \quad (2)$$

これは分布が各次元のデータごとに**独立 (independent)** であるということである。しかし、画像や文章をピクセルやトークンごとに独立だと考えるのはあまりにも悪い近似だということは容易に想像がつく（ただのノイジーな画像、もしくはランダムな文字列になってしまう）。よって、上の多次元の確率分布を「そのまま」扱わなければならない。しかし、多次元のしかも連続変数の分布を正確に見積もることは絶望的である。

そこで統計学では有限のサンプルから、もとの多次元の確率分布を「推定」するという手順を踏む。例えば  $x$  という 1 確率変数の平均  $\langle x \rangle$  や**分散**  $\langle (\Delta x)^2 \rangle$  を求める際は

$$\langle x \rangle = \int x P(x) dx \quad (3)$$

$$\langle (\Delta x)^2 \rangle = \int (x - \langle x \rangle)^2 P(x) dx \quad (4)$$

という計算をしなければならない（つまり、 $P(x)$  を求めて、 $x$  について全空間で積分する）が、これを有限のサンプルから計算される平均  $\bar{x}$  と分散  $\overline{(\Delta x)^2}$

$$\bar{x} = \frac{1}{K} \sum_{i=1}^K x_i \quad (5)$$

$$\overline{(\Delta x)^2} = \frac{1}{K} \sum_{i=1}^K (x_i - \bar{x})^2 \quad (6)$$

に置き換える<sup>3</sup>。これがサンプル数  $K$  が無限大のときに両者が一致すると考えるわけである。このように平均や分散くらいしか考えない場合は、背後にある分布として**正規分布 (normal distribution, Gaussian distribution)** を仮定している。正規分布は逆に平均と分散だけで分布のすべての性質が決まるので、分布を正確に見積もっていることにもなる。しかし、現実にはこれでうまくいくとは限らない。もっと裾野が広がったり狭かったり、複数の山や谷をもつ分布も存在する。また、多次元の変数の間の「非線形の」相関が存在する場合もある。

<sup>3</sup>分散に関しては  $K$  より  $K-1$  で割ったほうがよいが、 $K$  が大きければ数値的に影響はない。

もちろん有限のデータから厳密にもとの分布を求めることは不可能なので、どのように近似するかという話になる。1変数であれば正規分布を考えるのは**中心極限定理**からも適切である。平均  $\mu$ 、分散  $\sigma^2$  をもつ正規分布が従う確率密度関数<sup>4</sup>は

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (7)$$

で与えられる。ただし、多次元の場合にどのように確率密度関数を決めたらよいか、一般的な処方箋はない。

## 2.2. 分類の問題

簡単のため2変数  $(x, y)$  で考えると、分類の問題はその属性を  $C$  とすると、

$$P(C|x, y) \quad (8)$$

という**条件付き確率密度**<sup>5</sup>を考えることになる。例えば、図1には乳がん患者のデータが示されている（これは教育的に非常によく用いられる Wisconsin 大学のデータ [5] を規格化したものである）。縦軸と横軸  $(x, y)$  は患者の属性であり、機械学習の場合は**説明変数**と呼ばれる。それに対して、この患者ががんかどうか、ということ为例えば0と1で表現することとすると、これは**目的変数**である。この図の場合は三角（黄色）が0であり、丸（青色）が1となる。分類タスクの目的は適切な説明変数の関数を見つけて、それを用いて目的変数を説明するということである。

例えば、**ロジスティック回帰 (Logistic regression)** と呼ばれる手法<sup>6</sup>では2変数の場合、離散的な点  $(x_i, y_i)$  に対して  $C = 1$  である確率は次式で定義される。

$$P(C = 1|x, y) = \frac{1}{1 + e^{-\eta}} \quad (9)$$

$$\eta = w_1x + w_2y + b \quad (10)$$

<sup>4</sup>確率を計算する際の元になる関数のこと。

<sup>5</sup>確率や統計の分野では条件付き確率と呼ばれるもの考えることが多く、例えば B が起こった条件のもとで A が起こる条件付き確率を

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

と定義する。ここで  $P(A, B)$  は A と B が同時に起こる確率、 $P(B)$  は B が起こる確率である。

<sup>6</sup>回帰という名前がついているが、ここでは分類の手法として扱う。

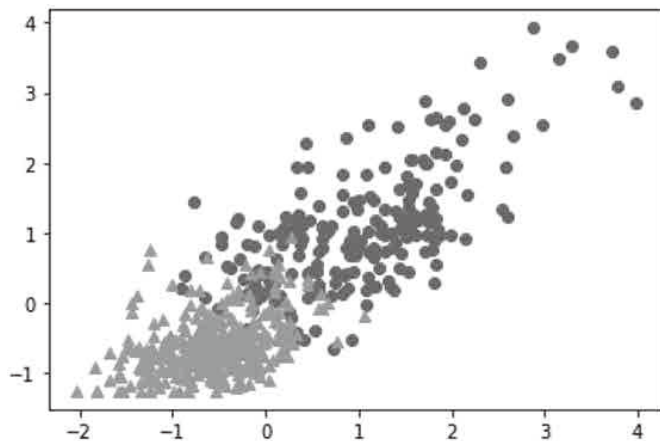


図 1: 乳がん患者の散布図データ [5]。横軸は 2 列目の radius mean、縦軸は 9 列目の concave points mean に入っているデータを用いた。三角が正常、丸ががんを表す。

(70)

これは図2にあるように、ある直線を境として0と1を切り替えるような関数であり、2次元空間を分割しているとみなすことができ、図1のようなデータを分類する際にも利用できるものと考えられる。また、 $C = 1$  というのは属性が1である確率（以下の例ではがんであるという確率）ということである<sup>7</sup>。

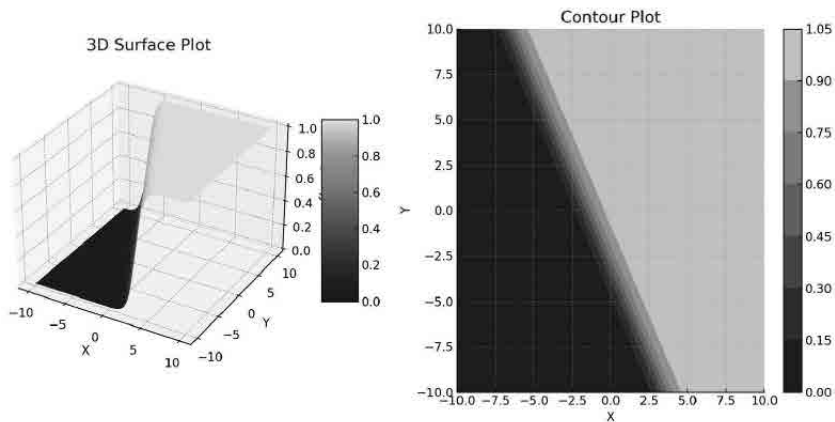


図 2: 2 変数のロジスティック回帰に使われる関数。パラメータは  $w_1 = 2, w_2 = 1, b = 3$ 。左は 3 次元プロット、右は等高線図。ChatGPT を用いて作成。

分類のタスクは大部分が**教師あり学習 (supervised learning)**である。つまり、あるデータ  $(x_i, y_i)$  に対して、**教師信号 (teaching signal)** と呼ばれる  $t_i$  が付随すると考える。例えば、 $(x_1, y_1)$  のデータががん患者のデータであれば  $t_1 = 1$  とし、 $(x_2, y_2)$  のデータが健常者の患者のデータであれば  $t_2 = 0$  とし、データにラベルをつけていく（図 1 左ではそれぞれ丸と三角に対応する）。これが上の関数に比例して起こると考えるので、 $i$  番目のデータに対しては、 $p_i = P(C = 1|x_i, y_i)$  という確率で起こることになる。

そうすると、例えば上の例であれば、1 番目のデータは  $p_1$  という確率で生じているはずであり、2 番目のデータは  $1 - p_2$  という確率で生じていることになる。これを  $i$  番目のデータが生じる確率として以下のように表す（上の例で確かめること

<sup>7</sup>一方、がんでない確率は  $C = 0$  であるときということになるが、これらは排他的な事象なので、 $P(C = 0|x, y) = 1 - P(C = 1|x, y)$  で与えられる。

ができる)。

$$p_i^{t_i} (1 - p_i)^{(1-t_i)} \quad (11)$$

さて  $K$  個のデータが独立に生成されるとすると、式 (2) のようにその全体の確率は

$$p_1^{t_1} (1 - p_1)^{(1-t_1)} \times p_2^{t_2} (1 - p_2)^{(1-t_2)} \times \dots \times p_K^{t_K} (1 - p_K)^{(1-t_K)} \quad (12)$$

となる。これが最大化されるように  $w_1, w_2, b$  を定めるのが**最尤法 (maximum likelihood estimation)**である。最尤法では、パラメータは「尤も (もっとも)」らしい、つまり最大の確率でデータが生成されるように調整されているだろうと考える<sup>8</sup>。

実際はこの確率の対数を取ったものを最大化する。そのためには

$$\begin{aligned} L = & t_1 \log p_1 + (1 - t_1) \log(1 - p_1) \\ & + t_2 \log p_2 + (1 - t_2) \log(1 - p_2) \\ & \dots + t_K \log p_K + (1 - t_K) \log(1 - p_K) \end{aligned} \quad (13)$$

をパラメータで偏微分して、それが0になったところ (最大値) を求める。つまり、式 (10) の  $w_1, w_2, b$  に対する  $L$  の偏微分

$$\frac{\partial L}{\partial w_1}, \quad \frac{\partial L}{\partial w_2}, \quad \frac{\partial L}{\partial b} \quad (14)$$

を求め、これらを用いてパラメータを1回更新するアルゴリズムを作る。この手続きを繰り返し行って最適なパラメータを求めることを**学習**と呼ぶ。

### 2.3. 回帰の問題

回帰とは、別のわかりやすい言葉で言うと、**フィッティング (fitting)** のことであり、ここでは2次元の平面をフィッティングすることを考えてみよう。図3のようにデータが3次元空間に散らばっているが、何らかの理由により、ある平面の周りに分布していると考える。これを以下の関数を使ってフィッティングしたい。さてどうするか。

$$z = w_1 x + w_2 y + b \quad (15)$$

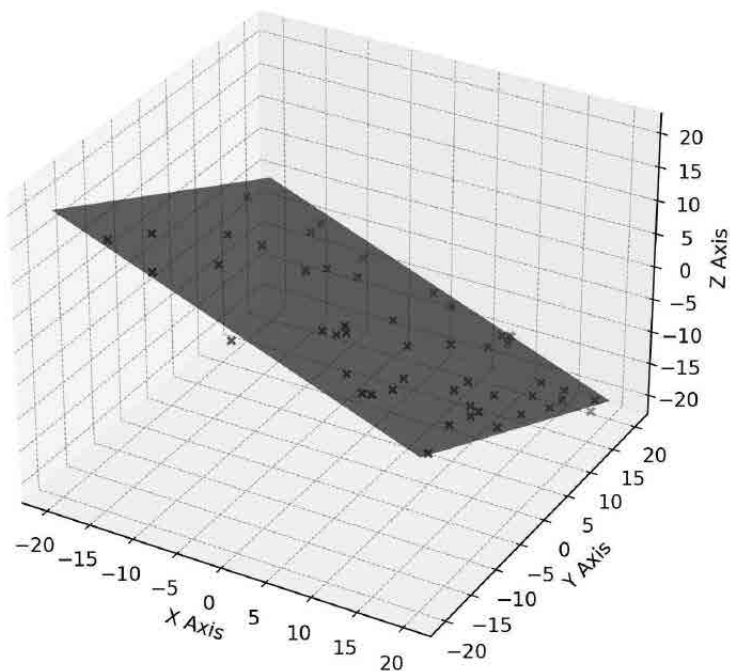


図 3: 3次元空間中の平面と、その周りにランダムに生成されている点。ChatGPTを用いて作成。



理工系1,2年生くらいの知識でこの問題を取り扱う場合は、通常**最小二乗法 (least squares method)**を用いる。つまり、 $(x_i, y_i, z_i)$  というデータの組が  $K$  個あった場合、

$$E = \sum_{i=1}^K (z_i - w_1 x_i - w_2 y_i - b)^2 \quad (16)$$

を計算する。これを全体の誤差とみなし、それが最小になるようにパラメータを選ぶ。つまり、式 (16) 中の  $w_1, w_2, b$  に関する  $E$  の偏微分

$$\frac{\partial E}{\partial w_1}, \quad \frac{\partial E}{\partial w_2}, \quad \frac{\partial E}{\partial b} \quad (17)$$

を計算し、それが0になるようなパラメータを探す<sup>9</sup>。これは上の最尤法と最大化（もしくは最小化）する関数が違うだけで手順としては似ている。

この最小二乗法も確率を用いて定式化できる。式 (15) でフィッティングするということは、正確な値にはならないわけで個別の誤差  $\varepsilon_i$  を含む。それを以下のように表す。

$$z_i = w_1 x_i + w_2 y_i + b + \varepsilon_i \quad (18)$$

そして、この誤差が例えば平均0、分散  $\sigma^2$  の正規分布（ガウス分布）にしたがっていると考えると、その確率密度関数は式 (7) と同様に

$$P(\varepsilon_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right) \quad (19)$$

で与えられる。ここで  $\varepsilon_i$  として、

$$\varepsilon_i = z_i - w_1 x_i - w_2 y_i - b \quad (20)$$

を代入し、確率密度が独立である（つまり、全体の確率密度は  $P(\varepsilon_1) \times P(\varepsilon_2) \times \dots \times P(\varepsilon_K)$  となる）という事実を用いると、最小二乗法はこの確率密度の対数の最大化と考えることができる。

<sup>8</sup>もっと正確にはパラメータがある確率分布に従っているとする考え方である。事後分布を最大化する場合に、 $P_1(w_1)P_2(w_2)P_3(b)$  ( $w_1, w_2, b$  はロジスティック回帰のパラメータ) のようなパラメータの事前分布と式 (12) の積をとると考える。

<sup>9</sup>上のように誤差が2次関数で表されている場合は、線形演算で最適なパラメータを求めることが可能である。

## 2.4. 生成の問題

データの生成とは、例えば確率分布が式 (1) で与えられる場合に、そこから1つサンプルを取り出す、ということであり、これは

$$x_1, x_2, x_3, \dots, x_N \sim P(x_1, x_2, x_3, \dots, x_N) \quad (21)$$

と表されることが多い。これは**サンプリング (sampling)** の問題であり、実際には様々な手法がある [4]。1 から 6 の数をランダムに出したい場合、サイコロを用意して振ってみて、値を得るのがサンプリングの1つの例である。しかし、所望の確率分布がそもそも何なのか分かっていない場合はサンプリングは難しい (例えば2次元の画像の確率分布は何かというのはアприオリには分からない)。また、サンプルを1個だけでなく、複数、それも確率分布の全体の様子を表すくらい十分な数のサンプルを取り出す場合、非常に時間がかかる (これは分布が複雑な場合に相当する)。

現在、画像生成などでポピュラーなアルゴリズムである**拡散モデル (diffusion model)** では**スコア (score)** によるサンプリングがよく用いられる。ここで、スコアとは以下の量を指す。

$$\nabla \log P(\mathbf{r}) \quad (22)$$

ここで  $P(\mathbf{r})$  は確率分布であり、 $\mathbf{r} = (x_1, x_2, x_3, \dots, x_N)$  である (多次元量を1つのベクトルで表している)。また、 $\nabla$  は  $\mathbf{r}$  に対する微分である。ここでは確率分布が分かっていると仮定しよう。このスコアを用いて、

$$\mathbf{r}' = \mathbf{r} + \alpha \nabla \log P(\mathbf{r}) + \sqrt{2\alpha\epsilon} \quad (23)$$

という計算をすれば、 $\mathbf{r}$  から新しいサンプル  $\mathbf{r}'$  が得られる (サンプリングできる) ことが分かる [6]。ここで  $\alpha$  はある正の定数、 $\epsilon$  はガウス型の (正規分布に基づく) 多次元ノイズである。これは物理学の統計力学では自然な対応物がある。(規格化されていない) 確率分布を

$$P(\mathbf{r}) = e^{-E(\mathbf{r})} \quad (24)$$

と表すと、 $E(\mathbf{r})$  は統計力学とのアナロジーではエネルギーに対応する。すると、上のダイナミクスは

$$\mathbf{r}' = \mathbf{r} - \alpha \nabla E(\mathbf{r}) + \sqrt{2\alpha\epsilon} \quad (25)$$

となり、これはカノニカル分布をサンプリングするランジュバンダイナミクスと等価である [4]。物理系における統計力学の問題ではエネルギーは個々の運動エネルギーと全体の位置エネルギーの和であり、それを陽に書き下すことができるので、ランジュバンダイナミクスを実行することができる。しかし、画像などであればそもそも分布もエネルギーもどのようなものか分かっていない。このような状況でもサンプリングを可能にしたのが拡散モデルであり、上のスコアをうまく推測したり、もしくは深層学習を使って有限のデータから多次元の確率分布を推定し、クオリティーの高いサンプル（例えば画像）を生成できるようになった。これは現在では様々なものに应用されており、その一例としては、タンパク質の会合の問題を扱う **RFDiffusion** と呼ばれるソフトウェアも開発されている（フリーで使うことができる） [7]。

拡散モデル以前は、画像を生成する際には、低次元のガウス分布から高次元のデータを生成する**変分オートエンコーダ (Variational autoencoder, VAE)**(図 4) や、**敵対的生成ネットワーク (Generative Adversarial Network, GAN)** とその進化形がよく用いられてきた。しかし、ここ数年は拡散モデルに基づく手法が大勢を占めており、GAN や VAE より計算コストがかかるが、現在 (2023 年 12 月) では有償の ChatGPT や Bing などに装備されており、誰でも利用できるようになった。拡散モデルの原理に関する日本語の解説としては、岡野原 [6] などを見よ。

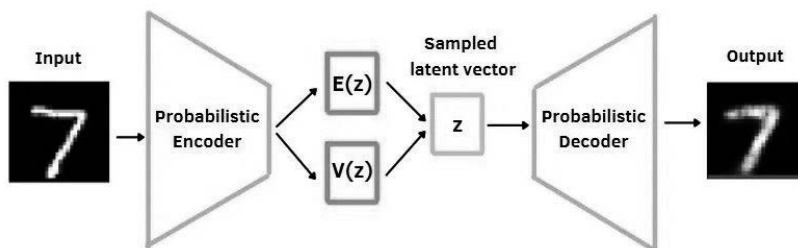


図 4: 変分オートエンコーダ (Variational autoencoder, VAE) の説明のための模式図。 <https://medium.com/dataseries/variational-autoencoder-with-pytorch-2d359cbf027b> から引用。 こういう模式図は ChatGPT で描くのがかなり難しい。

### 3. 生成 AI の利用について

#### 3.1. 生成 AI の歴史

生成 AI は現在の第 3 次 AI ブームの中から生まれてきたものと言っていいだろう。2012 年に Geoffrey Hinton らが AlexNet によって画像認識能力を飛躍的に向上させた。ここで用いられた**畳み込みニューラルネットワーク (Convolutional neural network, CNN)**こそ深層学習の本質的な技術であり、その後は CNN のような**深い層のニューラルネットワーク (Deep neural network, DNN)**を利用した研究が爆発的に増えた。画像認識や音声認識のような分類の問題があらかた「解かれた」後に、研究者たちは生成の問題にも挑戦していき、画像や文章を生成する研究も進展していった。そして画像生成に関しては上で述べたように拡散モデルが 1 つの頂点を極めている（もちろんこれが最終形はどうかは分からない）。また、文章の生成に関しては、文章が順序をもつことから再帰的な構造をもつ**リカレント・ニューラルネットワーク (recurrent neural network, RNN)**の適用がなされたが、その場合は、当時のアーキテクチャではあまり人間的な応答をしない、長い文章が作れない等の問題があった。それが 2017 年の **Transformer モデル**の出現によってがらりと変わり、これに基づいて Google が BERT と呼ばれるモデル、OpenAI 社が GPT というモデルを作り、それによって自然言語処理、文章生成、機械翻訳に関する研究が画期的に変わっていった<sup>10</sup>。特に OpenAI 社は GPT-1, 2, 3 とモデルを進化させていき、GPT-3 のレベルでも非常に自然な英語や日本語が生成できるようになった<sup>11</sup>。

#### 3.2. ChatGPT ショック

OpenAI 社の ChatGPT は 2022 年 11 月末に公開され、最初は GPT-3.5 が無償で利用可能だったが、数ヶ月後にサブスクリプション版として有償で GPT-4 が利用できるようになった。ChatGPT の文章生成能力、特に会話のように一連の会話の流れなどを理解できる能力は、従来の ChatBot (Siri など) のような定型の答えしかできないものとは次元が異なり、人々を驚愕させた。ChatGPT も基本的に確率によって言語を生成しており、

$$P(x_{N+1} | x_1, x_2, x_3, \dots, x_N) \quad (26)$$

<sup>10</sup>ただし、トランスフォーマーを RNN から理解するという研究も進んでいる。

<sup>11</sup>この頃、日本では GPT-3 に基づく AI ノベリストというサイト (<https://ai-novel.com/>) が作られ、その自然さに驚いたものだった。GPT-3.5 が出るまでは。

という条件付き確率を推測し、 $x_1, x_2, \dots, x_N$  のトークン（単語のようなもの）の並びから次のトークン  $x_{N+1}$  を生成している。しかし、確率的に生成するとランダムな文章になりそうなものだが（以前のマルコフ状態モデルなどを使ったものではそうであった）、ChatGPT はそうはならない（もちろん条件付き確率から生成しているので、独立の分布を仮定したときのように完全にランダムになることはない）。もちろんこれは分布が一様なものではないということを表しているだけであるが、チョムスキーなどの言語生成モデルはもっと決定論的で論理的なものを想定している [8] ので、それとはかなり異なるものと言える。

OpenAI 社は並列コンピュータ、特に大量の GPU を潤沢に使える環境にあり、巨大な文章データを入力し、膨大なパラメータ数をもつ transformer のモデルを学習させた。これは当時では OpenAI 社にしかできなかったものと思われる。しかし、巨大とはいえ有限の言語データで学習しており、モデルも有限なものなので様々な限界がある。また、データ管理やセキュリティ、倫理的問題なども発生した。ChatGPT が発表された当初の代表的な問題としては、以下が挙げられる。

1. データにない（もしくは入力との関連付けが難しい）ものに関しては不適切な文章を生成してしまう。これは**幻覚 (hallucination)** と呼ばれる現象であり、特に文献を検索するといったことをすると、現実には存在しない仮想的な文献を生成してしまう。
2. 数学的でロジカルな思考や計算などができない。これは文章のみで学習しているので、式の意味や式変形、計算の本質を理解していないためである。物理学も数学を大部分用いるので、同様のことが起こる。
3. 学習するのに大量のコストがかかるのみならず、生成する際にも多大なコストがかかる。
4. OpenAI 社のクラウドサービスであるので、アップロードした文章の情報が漏洩、流出する可能性がある。
5. 倫理的に問題のある文章（フェイクニュースや誹謗中傷など）を生成し、それらが悪用される恐れがある。

しかし、ChatGPT の発表から 1 年が経ち、これらの問題に関してはかなり改善された。1 に関しては、入力されるデータはますます増えてきており、また、生成する際に検索を行うという機能も実装されたので、以前ほど幻覚の問題は現れなくなった（なくなったわけではない）。2 に関しては WolframAlpha[9] などの計算のクラウドサービスと連携することで単純な計算であれば間違いを生じることがなくなった。式変形などもある程度は正しく実行できる。また、サム・アルトマンの解任劇で明らかとなったのは、OpenAI 社は現在 Q\* という ChatGPT とは異なるアルゴリズムの AI モデルを開発しているようであり、それは数理的な推論能力のあるものとされている。3 に関しては、より少ないデータ、小さいモデル、つまり economical なモデルで GPT-3.5 レベルを目指す研究が進んでいる。4 に関しては、OpenAI 社ではオプトアウトのようなユーザーのデータを蓄積しないシステムにしたり、ローカルな ChatGPT を作る研究やアプリ開発も進んでいる。5 に関しては、倫理的に問題のある文章が生成されないように、入力に対して制限を設けるなどの処置がとられている<sup>12</sup>。ユーザーのマナーやリテラシーを向上させることも一助になるだろう。

### 3.3. 生成 AI と教育について

生成 AI は教育にも多大なインパクトがあった。日本ではビッグデータに対して AI を存分に使いこなす社会としての Society5.0 という概念が打ち出され、また、AI 戦略 2019 において、大学や高校でも AI に関する授業を行うことを推奨するようになった。しかし、基本路線としては、いままでの統計学の話であるとか、プログラミングの話を中心者用にかみ砕くということが想定されていた。また、深層学習のようなものは「難しく」研究レベルのもので、基本は専門家が触るものであり、教育の現場ではあくまで「遠目で見る」というスタンスであった。

しかし、2022 年以降、生成 AI が登場することで AI の民主化が起こった。画像生成 AI はプロンプトを入れれば様々な画像を生成できるようになった。ChatGPT などの大規模言語モデル (Large Language Models, LLM) はまさに会話をしているように用いることができ、プログラミングなどがある意味「不要に」なった。よって、上で述べたような従来（といっても 4 年前！）の情報教育の変革が現在求められていると言える。

ChatGPT が発表された当初は、これを利用すべきか否かという議論が存在し、イタリアなどでは一時期禁止された。しかし、現在では Microsoft Office やスマホ

---

<sup>12</sup>しかし、ここが人力であり、自動化するのが難しいらしい。

のアプリ、ブラウザなどに標準で搭載されるようになり、利用しないという選択はありえない。よって、どのように付き合っていくか、という議論をしなければならぬ。そこで文科省や各大学がそれぞれのガイドラインや手引きなどを公表した。例えば日本医科大学が2023年8月に公表したものは以下である。

## 文章生成 AI(ChatGPT など) の利用に関する手引き

日本医科大学 生成 AI ワーキンググループ 作成  
2023/8/31

手引きを作るに当たっての経緯 ChatGPT に代表される大規模言語モデルや DALL-E, Stable Diffusion のような画像生成アルゴリズムは近年発表された非常に強力な AI ツールであり、正しく用いれば学生や研究者にとって多大なベネフィットがある。その中でも ChatGPT などの文章生成 AI は、情報の検索ツールとしてはまだ十分ではないが、過去の膨大な文章データベースを元に自然な日本語(もしくは英語)での応答が与えられるので、様々な質問を継続的に ChatGPT に投げかけて「対話」することで、議論のたたき台を作ったり、考えを整理したり、別の視点を導入するなどといった「知的な」作業が可能になる。しかし、教育的には剽窃や捏造、著作権侵害のような、クリティカルな問題をはらんでいる。ここでは日本医科大学での主に文章生成 AI の利用法に関して生成 AI ワーキンググループの見解を述べる。

### 留意すべき文章生成 AI の特性

1. 日本語や英語としては自然でも、訓練データが雑多なデータベースであるため、内容としては誤った文章を容易に作り出してしまう。
2. 個人情報や機密情報などを入力すると漏洩する危険性がある(実際、漏洩していた)。
3. 生成画像や生成文章に関しては将来的に著作権の問題がある。

### 学生にとっての手引き

1. 文章生成 AI を使ったときにはその旨をきちんと明記する。
2. 内容の信ぴょう性をチェックする。教科書や専門書のような情報源を調べたり、別の検索ツールを使ったりして、複数のソースで内容の正確さを吟味すること。検索ツールとしてはまだ不十分であることを十分意識すること。
3. 個人情報や機密情報などは入力しない。OpenAI の ChatGPT を使う場合はオプトアウトしておく (明示的に学習データの利用を許諾しない旨を OpenAI に伝える)。

#### 教員にとっての手引き

1. 文章生成 AI を適正に使っているかどうかを学生にチェックさせる。チェック項目としては以下などがある。a. 出典や引用文献を明記しているか。b. 自分の意見や分析を加えているか。オリジナリティや創造性があるか。c. ファクトチェックや校正を行っているか。
2. レポート提出のみで学生を評価するのではなく、口頭試験や MCQ(多肢選択テスト) などの客観試験のような、外部資料や文章生成 AI から切り離された状況での試験の評価も加味すること。(文章生成 AI を使ってレポートを作成したかどうかを判別するソフトもあるが、マージナルな判断能力しかもたないことが分かっている。) レポートに関しても、教師や他の学生からのフィードバックを与えることや、学生間の議論を行わせることも有効である。
3. 今後学生が活用することを想定して、教師側も文章生成 AI を定期的に使っておき、その性能や限界などを評価・認識しておく。また、文章生成 AI を教育目的で使う際には細心の注意を払う。文章生成 AI の概要や、効果的な使い方、悪用の事例なども学生に適宜示す。
4. 最終的な目標・目的が何であるかを意識させる。ChatGPT を無目的に使い、結果コピー&ペーストのような行為を行うことは剽窃や捏造、著作権侵害と同じであり、本人の教養・スキル・人間性の観点から問題があることを常に意識させる。



5. 生成系 AI の著作権の問題に関しては、2023 年 6 月現在国内においては、生成・利用段階では通常の著作権法が適用されるが、データを収集し学習するレベルでは限度を越えなければ著作権者の承諾を得る必要はないとされている。このことを踏まえた教育、研究活動を行い、学生にも周知させるものとする。

この手引きは 2023 年 8 月現在のものであり、以降適宜見直される。以上。

上で述べたことを簡潔にまとめると、今後は生成 AI を学生がどれくらい理解しているか、使いこなせるかということを中心に教育、評価していく必要があるだろう。個別の LLM によらないプロンプト・エンジニアリング [10] (どのようにプロンプトを工夫すればいい結果が得られるか) のようなことを教えるべきなのかもしれない。また、AI 時代の大学教育<sup>13</sup>ということも様々なところで議論されている [11]。

### 3.4. 生成 AI の個人的な利用について

#### 3.4.1. 生成 AI とまず戯れる

さて、それでは私はこの 1 年間どのように生成 AI に向き合ってきたのだろうか。流行り物が好きなので、拡散モデルに関しては、Google colab のファイルがあるとすぐ試したりはしていた。ただし、2022 年までは画像の精度も悪く、また、既存のものをつぎはぎしているように感じられた (図 5)。しかし、2023 年の DALL-E などでは例えばアニメ調にするのも非常に自然にできるようになった (図 6)。また、静止画だけでなく、動画も自動生成できるようになり、ある角度からの画像のみから別の角度 (背面など) の生成もできるようになった。2023 年 12 月現在では実用に耐えうるような数分の動画を作ることも可能になってきている (さすがに自分ではやっていない)。

また、音楽の自動生成も興味があるが、数年前には授業で学生に Bach Doodle [12] というものを紹介していた。これは DNN の構成もわかりやすく、結果もそれなりにバッハ風の作曲ができるものであった。しかし、結果としては単純すぎてつ

---

<sup>13</sup>北陸大学の杉森氏によれば、『AI 時代の大学教育の在り方については、人間にしかできない創造力のために、技術リテラシー・データリテラシー・ヒューマンリテラシーの 3 つを必要とする「ヒューマニクス」の創成』が重要になるとのことである (私信)。



図 5: 2022 年に stable diffusion で作った画像。プロンプトは左は jimie hendrix and star wars、右は寿司を食べるジョージルーカス。生成画像が原画そのまま出力されているように感じる。また、オブジェクト間のつながりも悪い。

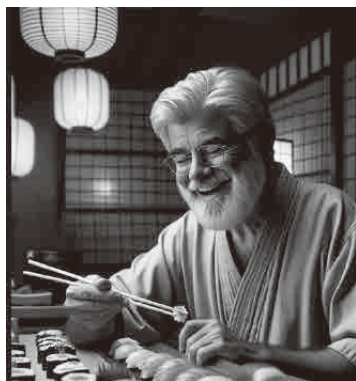
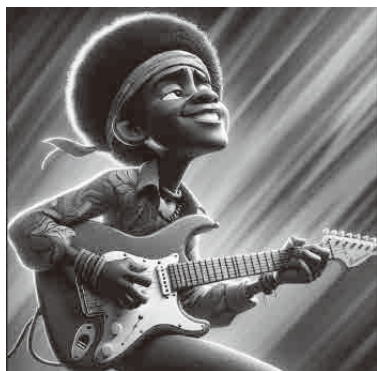


図 6: 2023 年 12 月ごろに ChatGPT で作った画像。プロンプトは左は pixar 風の jimie hendrix、右は寿司を食べるジョージルーカス風の人。2022 年の結果と比べるとかなりよくなった。

まらない。また、OpenAI 社も Jukebox [13] という音楽生成 AI を作っていたり、拡散モデルに基づく Riffusion [14] というものもあるが、まだ試作品という感じが拭えない（音のクオリティや音と音のつながりが悪い）。しかし、2023 年の 12 月に発表された Suno AI [15] はかなり自然なポップチューンを生成できるようになっており、これは画期的なものであると感じる（これを市販のポップチューンと区別するのはかなり難しいかもしれない）。まだジャズやクラシックのような複雑な構成の音楽を生成するのは道半ばだが、10 年以内にはできるようになるかもしれない。私は例えば洋楽であれば、Todd Rundgren, XTC, Frank Zappa, Stevie Wonder, Steely Dan, Beach Boys, Frefab Sprout, Herbie Hancock, Wayne Shorter, Allan Holdsworth といった、複雑なロック、ジャズ的で Adult Oriented Rock 的だが、そこまで商業的でない音楽が好きなので、そういったものを混合した音楽が作られるのかということに興味がある（ビートルズに似た曲を作るとか、有名なものを真似ることにはあまり関心がない）。

### 3.4.2. ChatGPT と教育

生成 AI で近年最も隆盛を誇ったのは LLM を搭載した文章生成 AI であろう。私も ChatGPT と BingAI については発表された当初から利用しており、その他にも Google の LLM である Bard もアカウントは作ったがほぼ利用していない。2023 年の 11 月ごろに Google が Gemini というマルチモーダル LLM を発表し、それは「もしその速さで動くのなら」驚愕ものであったが、そうでないことが分かってがっかりしたことがあった<sup>14</sup>。

ChatGPT に関しては、プロンプトが記録されているので、過去にどのようなことをやっていたのかを振り返ることができる。最初にご多分に漏れず、いろんなことを検索してみて有用かどうか試してみた。しかし、上で述べたように幻覚の問題があり、あまり有用ではないことがすぐに分かった。2023 年 4 月に

AI の進歩が社会に与える影響を、とくにマイナス面に焦点をあてて考える上で、日本語で読める重要な本を 10 冊挙げてください。

ということを聞くと以下のように答えた。

<sup>14</sup>Google はそういうことが度々ある。量子コンピュータの超越性のニュースなど。

AIの進歩に関するマイナス面に焦点を当てた日本語の重要な本を10冊ご紹介します。これらの本は、AI技術の発展が社会や個人に与える潜在的な問題やデメリットについて議論しています。

- 『AI vs. 人間』 - 松浦晋也
- 『人工知能は人類を滅ぼすのか?』 - 野中郁次郎
- 『データ主義の罠: AIが創り出す不自由な未来』 - 竹内謙礼
- 『AIリスク: 人工知能がもたらす未来の危機』 - 西尾幹二
- 『AIの倫理: 人工知能時代の倫理と法』 - 関根秀夫
- 『プライバシーの終焉: AI時代の個人情報保護と人権』 - 坂東忠信
- 『AI時代の働き方: 失われる仕事、創られる仕事』 - 木暮祐一
- 『AIが創る格差社会: 技術革新がもたらす新たな分断』 - 金子勇
- 『AIと差別: 人工知能が拡大する格差と差別』 - 堀内充弘
- 『AI時代のデマ: 情報操作とフェイクニュースの脅威』 - 細谷雄一

これらの本は、AI技術がもたらす様々な懸念や問題について深く考察し、読者に対してそのリスクを認識し対処するための示唆を提供しています。

という本を挙げたが、ほとんどが実在しない本であった(いくつかの本は本物かもしれないが)。そこでむしろ幻覚を楽しむということもやっていた。例えば、人のことを尋ねると、よほどの有名人でもない限りは架空の内容を出力するので、どれくらいそのようなことが起こるのか試していた。また、足し算掛け算のような単純な計算であっても正しい答えを出さないことがあった。

しかし、2023年5月になるとプラグインというものが始め、webを検索したり、Wolframalphaと連携して計算したりできるようになった(BingAIのほうが検索能力は当時は高かった)。それらによって幻覚や計算ミスはかなり減った。6月ごろにはPDFもプラグインで読み込めるようになったり、chatpdf [16] というクラウド

サービスもできた。これらによって論文等の長文を ChatGPT に入れることもでき、要約や質問ができるようになった。また、YouTube もある程度の動画は文字起こしがされており、それを ChatGPT で読み込んで要約する summarize.tech [17] というクラウドサービスもできた。これらによって論文や YouTube の講演などの概要を掴むことができるようになり、論文を読んだり YouTube を見たりする時間が短縮された（もちろん完全な要約ではないので、重要なものであれば自分で再チェックする必要がある）。

また、人工知能の授業において 5 択のミニテストを行うことを考えたので、ChatGPT に次の問いかけを行った。

人工知能に関する 5 択の問題を 6 つ作ってください。テーマは教師あり学習にしてください。またレベルは初歩的なものでいいです。

すると以下の返答が出力された。

- 教師あり学習は何を用いて学習を行いますか？
  - － ラベル付きのデータ
  - － ラベルのないデータ
  - － ランダムなデータ
  - － フィードバックのないデータ
  - － 人間の監督なしで学習
  
- 教師あり学習の主なタスクの一つである回帰とは何ですか？
  - － カテゴリーを予測するタスク
  - － データポイント間の距離を予測するタスク
  - － 真実の値を予測するタスク
  - － データの分布を予測するタスク

- データセットのクラスターを予測するタスク
- 教師あり学習アルゴリズムの例として、どれが正しいですか？
  - K-近傍法 (K-nearest neighbors)
  - K-平均法 (K-means)
  - 階層的クラスタリング (Hierarchical clustering)
  - DBSCAN
  - 自己組織化マップ (Self-Organizing Maps)
- 教師あり学習で、過学習とは何を指しますか？
  - 学習データに対する性能が低すぎる状態
  - モデルが学習データに過度に適合し、新しいデータに対する性能が低い状態
  - 学習が十分に行われていない状態
  - 学習データが少なすぎる状態
  - モデルが予測結果に確信を持ってない状態
- 教師あり学習で、訓練データとテストデータの違いは何ですか？
  - 訓練データはモデルの学習に使用され、テストデータはモデルの性能を評価するのに使用される
  - テストデータはモデルの学習に使用され、訓練データはモデルの性能を評価するのに使用される
  - 訓練データとテストデータは同じデータである
  - テストデータは訓練データの一部である
  - 訓練データはテストデータの一部である
- 教師あり学習における「損失関数」の目的は何ですか？

- 学習の速度を制御する
- モデルの複雑さを制御する
- モデルの予測と実際の値との差を計算する
- 学習データを生成する
- モデルの学習を終了するタイミングを決める

これらはどれも遜色のない問題であり、難易度も適当である。これを自分で考えるとなるとかなり大変であり、少なくとも数十分はかかってしまうが、ChatGPTでは瞬時にできる。また、正しいかどうか確かめるときにも ChatGPT が利用できる（答えを聞いてみればよい）。

また、物理の教員としては物理学の問題に対してどれくらい正しく答えられるかということにも興味があるので、以下のような質問を試してみた。

1 ナノメートル離れた2つのアミノ酸にそれぞれ1つの電子がのっかっていると、1ピコ秒経つとどれくらい動くでしょうか。ただし、2つのアミノ酸は最初は電的に中性とし、アミノ酸の分子量を100とします。一つのアミノ酸は止まっているとして計算してよいです。

それに対する ChatGPT の返答は以下であった。

この問題はクーロンの法則とニュートンの第二法則を使用して解くことができます。まず、クーロンの法則を使用して二つの電荷間の力を計算します。次に、ニュートンの第2法則  $F = ma$  を使用して、その力がアミノ酸に与える加速度を計算します。最後に、基本的な運動の方程式  $d = \frac{1}{2}at^2$  を使用して、与えられた時間内にアミノ酸がどれだけ動くかを計算します。

ここで、

- $F$  は力（ニュートン）

- $m$  はアミノ酸の質量 (キログラム)
- $a$  は加速度 (メートル/秒<sup>2</sup>)
- $d$  は距離 (メートル)
- $t$  は時間 (秒)

アミノ酸の質量は分子量を使用して計算できます。分子量は1モルの質量を表し、1モルは  $6.022 \times 10^{23}$  個の分子を表します。したがって、アミノ酸の質量は  $100 \text{ g/mol} \times 1 \text{ kg}/1000\text{g} \times 1 \text{ mol}/6.022 \times 10^{23}$  となります。

これらの情報を使用して、まず力を計算し、次に加速度を計算し、最後に距離を計算します。

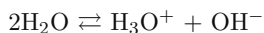
1ピコ秒経つと、アミノ酸は約  $6.93 \times 10^{-10}$  メートル (つまり、約 0.693 ナノメートル) 動きます。これは初めの距離の1ナノメートルに比べてかなり小さい距離ですが、微視的なスケールでは大きな移動となります。

また、化学の問題として

10の24乗個 (実際は10の24個と間違っただけで入力したが、常識的に考えて ChatGPT が  $10^{24}$  個と解釈してくれた!) 水分子が存在する場合に、1秒間に何回プロトン移動が発生するか?

と尋ねると ChatGPT は以下のように出力した。

水分子間のプロトン移動は、水の自己イオン化反応を通じて行われます。この反応は以下のように表されます:



この反応の速度定数は、25度Cでおおよそ  $2.4 \times 10^{-14} \text{ s}^{-1}$  です。しかし、この値は2つの水分子間の反応を表すので、全体の反応率を計算するためには全体の水分子の数を考慮する必要があります。



提供された情報に基づいて、10 の 24 乗個の水分子が存在することなので、以下のように計算できます：

反応率 = 速度定数 × (水分子の数の半分)

なぜなら、この反応は 2 つの水分子が 1 つの  $\text{H}_3\text{O}^+$  と 1 つの  $\text{OH}^-$  に変わるの  
で、全体の水分子の数の半分だけが反応に関与するからです。

したがって：

反応率 =  $2.4 \times 10^{-14} \text{ s}^{-1} \times (10^{24} / 2) = 1.2 \times 10^{10} \text{ s}^{-1}$

したがって、1 秒間に約  $1.2 \times 10^{10}$  回のプロトン移動が発生すると予想されま  
す。ただし、これはあくまで理論的な計算で、実際の状況では温度、圧力、その他  
の条件により結果が異なる可能性があります。

このように非常によく知られている問題、公式などが分かっている問題、それに対  
する数値のあてはめという問題に対してはほぼ及第点が取れるということが分かる。

しかし、いろいろ試してみたところ、高校の初年度程度の問題に対しては正しい  
答えを与えるが、それ以上のレベルの問題、論理性が求められたり図でないと  
分からないような複雑な問題に対しては ChatGPT は正しい答えを与えないこと  
が分かった。これはやはり GPT が言語のみから学習していることの限界であり、  
式を変形したり、近似したり、図やグラフを見て考えたり、ということをはなされ  
ていないことに起因するのであろう。しかし上で述べたように Google はマルチ  
モーダルな（すなわち文章のみならず図や音声などの様々な情報を入力して学習  
する）LLM を開発中であるし、OpenAI 社も論理的な思考を LLM に実装するこ  
とを研究しているので、近いうちに大学入試レベルの問題を作ったり、解いたり  
することが可能になるかもしれない。

### 3.4.3. 生成 AI と研究

研究に関しては、上で述べたように数理的な処理能力はまだまだ貧弱であり、物  
理学に関する研究内容そのものに関して新たな知見を与えるようなものではない。  
論文を要約できるとも言ったが、重要なことを見落としていることも多く、これも  
研究の本質を理解していないことに起因するものだろう。これに関しては OpenAI  
社が開発中の Q\* に今後期待したいところである。

一方、データを処理したり、生成したりするときに、われわれはプログラミング  
言語を用いるが、LLM は大量のプログラミング言語の文章を既に読んでいる。また、  
それには解説文が付随している場合も多いので、こういうプログラミングがした

いといえば、典型的なコードを出力してくれる。これは研究上も教育上も非常に有用である。そのおかげで、例えば、bash, latex, python, pytorch, fortran などのポピュラーなシェルやプログラミング言語に関しては、ほぼネットサーチしたり、本で調べたりする必要がなくなった。逆にコードを入力すれば、その各行がどういう意味をもつのかということも解説してくれるし、間違ったコードを入力すれば大体ここが間違えていそうだということも教えてくれる。また、変数名から大体それは何を意味するのか、ファイル名からどういう機能をもつコードなのか、ということも推測してくれる。データ量が少なく実行時間の短いプログラムであれば、有償の GPT-4 内で実行することも可能であり、その結果をグラフィカルに表示することもできる。これが研究のみならず教育に効果があることも自明であろう。

例えば、拡散方程式を数値的に解くことを考えて以下のように質問すると、

1次元の拡散方程式を数値的に解く python コードを書いて、階段型の密度の初期状態から始めて時間発展を計算し、ある程度拡散するまでを図にしてください。

ChatGPT は以下のように出力した。

了解しました。1次元の拡散方程式は以下のように表されます：

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2}$$

ここで、

- $u$  は位置  $x$  と時間  $t$  における密度や濃度を示します。
- $D$  は拡散係数です。

まず、離散化のために次のような差分方程式を考えます：

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = D \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2}$$

ここで、

- $\Delta x$  は空間ステップ
- $\Delta t$  は時間ステップ
- $n$  は時間ステップのインデックス
- $i$  は空間ステップのインデックス

この差分方程式を用いて、時間発展を計算します。

以下の条件で計算を行います：

- 初期状態:  $u(x, 0)$  は中央付近で 1、その他の部分で 0 の階段型とします。
- 境界条件:  $u(0, t) = u(L, t) = 0$  (固定境界条件)

実装と計算を行って、結果を図に表示します。

図から、初期の階段型の密度が時間とともに滑らかになっていることが確認できます。これは拡散が進行していることを示しています。固定境界条件のため、境界の密度は 0 のままであり、その影響で拡散は中心部分でより顕著になっています。

この数値計算は、拡散方程式の基本的な性質を示していますが、より高度な計算や異なる境界条件、初期条件を試すこともできます。何か他に質問や要望があればお知らせください。

その結果は下の図に示している通りであり、尤もらしい結果が出力されている。また、このときの python コードは以下の通りであった。

```
import numpy as np
import matplotlib.pyplot as plt

# 定数の設定
L = 10.0 # 1次元空間の長さ
T = 2.0 # 計算する時間
```

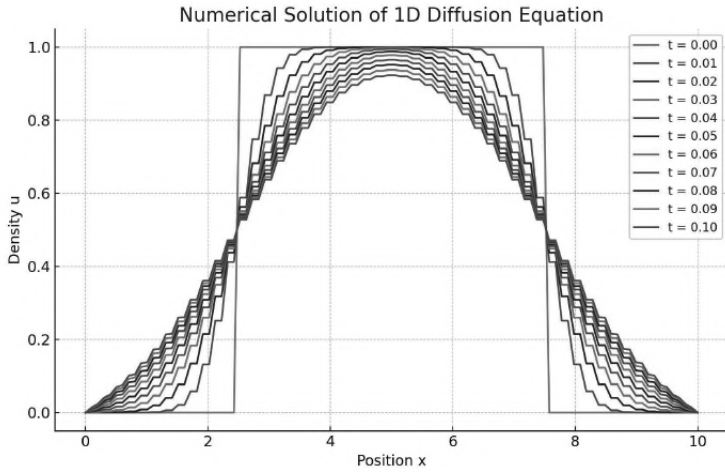


図 7: 1次元の拡散方程式の数値解。初期の密度は矩形型であり、 $x = 0, x = 10$  で密度が0になる境界条件を課している。

```
Nx = 100 # 空間の分割数
Nt = 200 # 時間の分割数
D = 0.5 # 拡散係数

dx = L / Nx
dt = T / Nt
alpha = D * dt / dx**2

# 初期条件の設定
u = np.zeros(Nx)
u[int(Nx/4):int(3*Nx/4)] = 1

# 結果を保存するための配列
results = [u.copy()]

# 差分方程式を用いた時間発展の計算
for n in range(Nt):
    u_new = u.copy()
    for i in range(1, Nx-1):
        u_new[i] = u[i] + alpha * (u[i+1] - 2*u[i] + u[i-1])
    u = u_new
    results.append(u.copy())

# 結果の表示
plt.figure(figsize=(10, 6))
for i, u in enumerate(results[::int(Nt/10)]):
    plt.plot(np.linspace(0, L, Nx), u, label=f't = {i*dt:.2f}')

plt.title('1次元拡散方程式の数値解')
plt.xlabel('位置 x')
```

```
plt.ylabel('密度 u')  
plt.legend()  
plt.grid(True)  
plt.show()
```

しかし、問題を2次元に変えると異常な挙動を示し、何回かのエラーの後で以下の出力になった。

申し訳ございません。再度、計算の全過程を見直します。

特に、差分方程式の計算の部分や境界条件の設定、初期条件の設定などを再確認します。

再度計算してみます。

再試行しましたが、結果は変わりませんでした。計算の過程やパラメータ設定に問題があるかもしれません。

一つの可能性として、拡散方程式の数値計算における安定性条件を考慮する必要があります。特に、時間刻み  $\Delta t$  と空間刻み  $\Delta x$  の間には以下の関係が成り立つ必要があります：

$$\Delta t \leq \frac{\Delta x^2}{4D}$$

この条件を確認して、必要であればパラメータを調整してみます。

安定性条件の計算結果から、現在の時間刻み  $\Delta t = 0.01$  は、条件  $\Delta t \leq 0.005$  を満たしていないことがわかります。このため、不安定な計算結果が生じている可能性が高いです。

この安定性条件を満たすように時間刻みを調整し、計算を再試行いたします。

安定性条件を考慮して再計算した結果、期待通りの時間発展が確認できました。初期の中央部分の階段型の密度が時間とともに拡散していく様子が明確に表示されています。

お手数をおかけし、申し訳ございませんでした。何か他に質問や要望があればお知らせください。

この出力では数回のエラーが出て、その度にこちらがその結果がおかしいと介入し、それを何回か繰り返すと正しい答えまで辿りつくことができた。よって、ChatGPT にプログラミングさせて、実行させた結果（例えば最初の結果）を鵜呑みにするのは危険であるということであり、これは常識的な結果であろう。こういった危険性を防ぐためには、やはりプログラミングのプロ、もしくはこの問題のように物理の問題であるのならば物理のプロにプログラミングや計算の妥当性を尋ねたほうがよい。ただし、このような ChatGPT の問題も Q\*などの登場によって改善されていくだろう。

### 3.5. ChatGPT に対する学生の初期の反応

2023年4月17日には2023年度の人工知能概論という授業を始めた。2回目の授業で ChatGPT について触れたが、知らない、使ったことがない、という学生が大半であった。これは ChatGPT が公開された 2022年11月末から翌年3月末まで、彼らの大部分は高校3年生で受験勉強をしており、生成 AI を利用する余裕がなかったためと思われる。

授業では以下のような課題を与えた。

**OpenAI の ChatGPT のアカウントを作って、以下を実行せよ。**

- ネットにある適当な長文の日本語（英語でもよい）の文章を 400 文字で要約させてみる。どこから文章を取得したかを書く。
- 俳句を作ってみる。ちゃんと 575 になるようにする。
- 適当な数学の問題を解かせてみる。例:  $x^2 + 2x + 1 = 0$  を解け。
- 医師国家試験の問題を解かせてみる。それを正解と比較する。参考: 第 113 回医師国家試験の問題および正答について | 厚生労働省 ([mhlw.go.jp](http://mhlw.go.jp))
- ChatGPT を使ったアプリを 1 つ考えてみよ、もしくは ChatGPT の面白い使い方を考えるか、調べてみよ。

彼らのレポートにあった記述を適当にピックアップしてみよう。要約に関しては、以下があった。

**(出典) 日本医科大学アドミッションポリシー** この文章の内容を要約すると、この私立医科大学は「克己殉公」という学是を基盤に、愛と研究心を有する質の高い医師、医学研究者、医政従事者を育成してきた歴史がある。本学は最新の医学を教育・研究し、国際的な視野と人間性を備えた医師・医学研究者を養成し、人々の健康の維持・増進、社会貢献を使命と考えている。本学が求める学生は、学是や理念、使命を理解し、自己の目的意識を持ち、努力することが求められ、生命倫理を尊重し、科学的思考力を備え、病める人の心を理解し、主体性と協働性を持ち、社会的な見識を有することを要求される。さらに、世界の医学・医療の進歩に貢献する強い意欲を持った人が望まれている。

**シンデレラを 400 文字で要約せよ。**シンデレラは、美しい娘でしたが、母親が亡くなった後、悪い義母と義姉たちに虐待されていました。彼女は家の仕事をし、動物たちと友達になっていました。ある日、王宮で開かれた舞踏会に招待されましたが、義母に出席を妨げられます。しかし、魔法使いの助けを借りて、シンデレラは美しいドレスとガラスの靴を手に入れ、王子とダンスを楽しむことができました。しかし、夜が更けると、彼女は急いで帰らなければなりません。王子は彼女を見つけるために全国を探しましたが、シンデレラは失われたままでした。しかし、王子がシンデレラのガラスの靴を見つけたことで、彼女を見つけることができました。そして、王子とは結婚し、幸せな結末を迎えました。

俳句に関しては 30 首生成して、何首が 575 になっているか調べている学生がいた。結果として 575 になっているのは 2 首だけであった。しかし、30 首があったという間に生成されるのは面白い。

**俳句を五文字七文字五文字になるよう 30 首作って。**

- 降りしきる 白き雪の中を歩む
- 青々とした 草原を駆け抜ける 疾走感



- 古き時代 限ぶ史跡を 訪ね歩く
- はるかなる 山々に春の 息吹く
- 初夏の風 吹き抜ける山々 爽やかに
- 真っ赤なる 夕陽が沈む 海辺に
- 深き森に 迷い込んだら 出口は
- すがすがしい 風に吹かれ 丘を登る
- かたちなき 音に耳澄まし 秋の夜
- まどろむ島 あくびしてる 昼下がり
- 人の波 かき分けて進む 繁華街
- 雨上がりの道 にぎわい戻り 青葉風
- やわらかな うたた寝光に 包まれ
- あの夏の 懐かしく思い出す ひととき
- 朝の光 射し込む部屋 にこやかに
- 桜咲く頃 思い出す君との 別れの日
- 夕暮れ時 赤い雲が流れる 黄金色
- 空高く 飛び立つ白鳥 見送りを
- 季節は移り 葉っぱが落ちる 秋の日
- 星座を見る 星空に思う ふとしたこと
- 波音に 聞こえる鷗の さえずり
- 海辺の岩に 座って見る 夕陽

- 霞む空に 浮かぶ月の影 うつろいて
- 初雪の日 庭に足跡が つけられる
- 虫の音に 耳を澄ます 夜ぶしさよ
- 静かな夜 星が輝く 心安らぐ
- 道の脇に 咲く花を見て 微笑む
- いざ出でよ 青い空の下 未知へと
- 淡い色の 空気に包まれて 夏の終わり
- 冬枯れの 木々に収しさ 感じ入る

医師国家試験の問題を解くことに関しては以下の表 1 にまとめてくれた学生がいた。これは 10 問を選んで、ChatGPT の答えと厚生労働省の答えが合っているかどうか比較しているものである。その学生の感想は以下であった。

倫理が関係する問題は ChatGPT には難しいことが分かった。また、全体的に正答率は低く、ボーダーの 8 割に満たないため、医師国家試験には合格しないと予想される。

	医師国家試験の問題									
	1	2	3	4	5	6	7	8	9	10
ChatGPT	C	E	C	DE	A	C	E	A	B	A
正答	D	E	C	E	C	C	D	A	E	A
正誤	×	○	○	×	×	○	×	○	×	○

表 1: 医師国家試験に対する ChatGPT の答えと正解の比較 (1 と 4 は倫理が関係する問題)。

ただし、学生はこの時は無料の GPT-3.5 を使っていたはずであり、能力は有償の GPT-4 と比べると低い。ある研究では、2018 年から 2022 年までの日米の医師国家試験を受験させたところ、GPT-4 はすべての年で合格ラインを上回ったことが報告されている。この当時はまだ画像を読み込めていないので、マルチモーダルなモデルができればさらに正答率は上昇するだろう。

ChatGPT の面白い利用法、アプリなどに関しては以下のような回答があった。

先生の授業の初めの時間を ChatGPT に文章を作ってもらって、それをあたかも自分で喋っているかのようにして、その部分が、喋り終わったら生徒に ChatGPT に作ってもらったものとネタバラシをするというもの。そこで生徒は、人間と AI の差異を感じることに (差異があるか、ないか) ができるのではないかと思った。

コミュニケーションをとる機会はないが、老人ホームに入る基準に達していない一人暮らしの高齢者のために、認知症予防として会話を促すアプリ。1日に数回話しかけてくれれば、体調の異変などを訴えるきっかけにもなると思う。

1つ目は、AIが将来自己完結型の高度な思考に到達した際のマッピングだと思った。節目となる段階で人間に理解できるような会話を生成して AI 内の思考を人間にも理解できるようにするためのものとなってもおかしくないと思った。2つ目は、未知の生物との対話だと思った。人間が性能を最大限に活用できる範囲の期間はもうかなりなくなっているように感じたので、現状の論文やレポートの活用に加えて動植物の意思に近いものをただ抽出するだけだったら応用を重ねれば可能ではないかと思った。また、将来地球外生命体との遭遇時に五分ほどで相手と対話出来たらロマンがある。

ChatGPT に SNS に投稿する文章を第三者目線で確認してもらい、炎上対策として利用する。

スポーツ名を入力した時に上達の方法を教えてください、自分のプレイ動画を送ると分析して改善点を提示してくれるアプリ。

もうあるかもしれないが、視覚障害者などが目の前にあるものを知りたい時に、何があるかきくと答えてくれるもの。

一番目のものは（たぶん本人は知らないままであろうが）**チューリングテスト (Turing test)** のことを述べているが、既に ChatGPT がチューリングテストをパスしているという報告もある。もちろんチューリングテストそのものはそんなに厳密なものではないので、ChatGPT が人間と同等かどうかという問題を議論するのは時期尚早だろう。。。と言いたいところだが、以下で述べるように既にその議論は始まっている。それが AGI の問題である。

### 3.6. AGI について

**AGI (Artificial General Intelligence)** は日本では**汎用人工知能**、もしくは**強い AI (strong AI)** として表現されることが多い [18]。これまでの AI はある特定のタスクを処理することに特化したものであり、これを**弱い AI (weak AI)** とも呼ぶ。特定のタスクとして、例えば画像認識をする、将棋や碁を打つ、機械翻訳をするということであれば、AI は人間を上回る能力を持っている。しかし、少し考えれば分かるように、われわれは既に、車や飛行機や各種の工作用ロボットなどのように、人力より強く、速く、正確に動作するものを作ってきており、人間の能力の一部が新しい技術によって置き換えられるということはこれまでに何度も経験してきている。ただし、AGI はそれらと比較して、汎用的に使えるということが特色であり、これは人間の知性そのものをモデル化、製品化するということである。人工知能という言葉が生まれた際にも AGI を作ることは目標の一つではあったが、難しすぎるのでそれを一旦棚上げして、特定のタスク処理を行う弱い AI を作ることに注力してきた。その結果として、現在の第 3 次 AI ブームがある。

しかし、ChatGPT などの生成 AI が進化することで、AGI の可能性や危険性に関する議論が 2023 年現在増えてきている。ChatGPT は人間の言葉の入力に対して、非常に自然に受け答えでき、直接尋ねていないことも慮って答えることもでき

る（ように見える）ので、これを AGI の初期のバージョンと言ってよいのではないかという議論もある。現在 GPT-4 を使った様々なサービスが作られているが、これは言語によって様々な問題が処理できるという現れでもあり、汎用性に近づいているのは確かである。また、数理的な思考能力も（上で挙げた Q\* のように）獲得すると汎用性はさらに増すだろう。

ただし、AGI に関する重要な課題として**身体性**が挙げられる。人間は知性だけでなく身体をもっており、それが共同的に働くのが普通である。例えば、赤ん坊が目で見ただけ、もしくは言葉を聞くだけで、体を使って外界を経験しなければ、きちんとした知性が身につくか疑問に思われる。また、もっと実際的な問題として、ChatGPT のようなもの（その進化系）が脳の役割をするにしても、体がないと実世界に働きかけることができない。そこでロボットを開発するというのも AGI にとって重要（もしくは本質的）なことになる。Google, OpenAI, Tesla などの企業はこぞってロボット開発にも取り組んでおり、これは AGI への布石と考えることもできる。AGI を搭載し、人間のように滑らかに動くロボット（例えば医師ロボット）が開発されれば、もちろん社会は激変するであろう。

AGI の特性の 1 つとして**自律性**が挙げられるが、これも弱い AI とはかなり異なる点である。弱い AI はある特定の（固定された）環境で特定のタスクを処理するだけである。しかし、AGI は動的に変化する環境に応じて自律的に動くことが想定されている。これに関する研究としては、東大の池上高志らによる Alter3 という取り組みがある [19]。そこでは環境としての外部の画像や音（音楽）とロボットが（もしくはロボット同士が）相互作用しながら体を動かすシステムが構築されている。

しかし、身体性と自律性をもつと、AGI ロボットは自らの意志で自分の複製を作る、または改良することも可能となり、これは AI の**超進化**という話題につながる。これは未来学者のレイ・カーツワイルが昔から主張していることであり、シンギュラリティの 1 つの形である。また、身体性をもたなくても、例えば、戦闘用飛行機を操縦するなどの軍事的な目的であれば、AGI に任せることは十分可能であり、そこで人間には予想もつかない判断をしてしまう恐れがある。そこで、**AI アライメント**という研究も始まっており、これは人間に制御可能な AI をいかに開発するかという研究分野である。また、AGI に関して、各国、各企業の研究者が open に課題などを共有し、共同研究するという動きもある [20]。

#### 4. まとめと生成 AI の今後について

本稿では機械学習による分類・回帰・生成というタスク処理の基本的な内容を確率分布という観点から復習してから、生成 AI の歴史、ChatGPT の個人的利用（教育面と研究面）、ChatGPT に対する学生の反応、最後に AGI について雑駁にまとめた。ChatGPT に関しては 2022 年末から 2023 年末までの進展であり、この急激な変化には全く驚かされる。これからこの進歩がさらに加速していくのかそれとも停滞するのかは全く分からない。AGI に関しては、最も明白な危険性としては軍事転用があるが、これも各国内で秘密裡に進んでいくことが想定されるので、国連などで規制の枠組みを作っても難しい課題である。

また、AGI には身体性が重要になる、つまり、AGI ロボットを作る必要があるということも述べた。以前はロボットは日本の十八番芸であったが（Honda の ASIMO, Sony の AIBO などがあった）、現在はアメリカの Google, OpenAI, Tesla などの企業、または中国企業が新規参入し、特に AI と接続することに関して日本の大企業は強みを出せないでいる。ただし、その中でも日本の中小企業である Preferred Networks 社は GPU チップの開発、深層学習アルゴリズムの開発からロボットの開発販売まで手広くやっており、将来的には期待がもてる。

とりあえずは生成 AI を教育や研究に取り入れていかなければいけない。教育に関しては、日本の場合は AI 戦略 2019 を早くも大幅に見直さないといけない状況であり、それに向けた取り組みも既に始まっている。ただし、教育においてはただ最新の技術を教えるだけでなく、背後にある数理、思想や哲学なども教える必要がある。そうでないと、その技術が古びたときに新しいことに全く対応できない人間が生み出されてしまう。また、AI 倫理の問題などは技術とは別個に教える必要があるだろう。研究に関してはプログラミングの効率は格段に良くなっていくだろうが、さらに数理的な思考のできる LLM やマルチモーダルな LLM が登場したときにそれをどう研究に生かすかというのは面白い課題であろう。

**謝辞** 奥出直人氏（慶應大学特任教授）、杉森公一氏（北陸大学教授）、長谷部孝氏（日本医科大学教授）、中澤秀夫氏（日本医科大学教授）、Steven Kirk 氏（日本医科大学教授）には原稿を読んでいただき、有用なコメントを頂いた。ここで感謝いたします。

## 参考文献

- [1] 藤崎弘士, 医科大学における人工知能教育について, 日本医科大学基礎科学紀要 **48**, 21 (2019).
- [2] 藤崎弘士, 日本医科大学における AI 教育と AI を活用した生物物理研究, 日本医科大学医学会雑誌, **19**(3), 279-282 (2023).
- [3] 盛山和夫, 統計学入門 (ちくま学芸文庫) 文庫 (2015).
- [4] Daniel M. Zuckerman, *Statistical Physics of Biomolecules: An Introduction*, CRC Press (2010); 翻訳は、藤崎弘士・藤崎百合訳、生体分子の統計力学入門、共立出版 (2014).
- [5] <https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-dat>
- [6] 岡野原大輔, 拡散モデル データ生成技術の数理, 岩波書店 (2023).
- [7] Joseph L. Watson, et al, De novo design of protein structure and function with RFDiffusion, *Nature* **620**, 1089–1100 (2023).
- [8] 長谷川香子, 生成文法の生得的言語知識 (普遍文法) と言語獲得, 石巻専修大学研究紀要, **29**, 45-58 (2018).
- [9] スティーヴン・ウルフラム (著), 稲葉 通将 (訳), ChatGPT の頭の中, ハヤカワ新書 (2023).
- [10] <https://will-blog.com/prompt-engineering-paper/>
- [11] ジョセフ・E・アウン (著), 杉森公一ら (訳), ROBOT-PROOF: AI時代の大学教育, 森北出版 (2020).
- [12] <https://doodles.google/doodle/celebrating-johann-sebastian-bach/>
- [13] <https://openai.com/research/jukebox>
- [14] <https://www.riffusion.com/>
- [15] <https://www.suno.ai/>

- [16] <https://www.chatpdf.com>
- [17] <https://www.summarize.tech/>
- [18] 荒川直哉, 山川宏, 市瀬龍太郎, 汎用人工知能の研究動向, 2014 年度人工知能学会全国大会 (第 28 回)  
[https://doi.org/10.11517/pjsai.JSAI2014.0\\_2C40S22a1](https://doi.org/10.11517/pjsai.JSAI2014.0_2C40S22a1).
- [19] Takahide Yoshida, Atsushi Masumori and Takashi Ikegami, From Text to Motion: Grounding GPT-4 in a Humanoid Robot “Alter3”,  
<https://arxiv.org/abs/2312.06571>.
- [20] サム S. アダムスら, 人間レベルの汎用人工知能の実現に向けた展望, 人工知能, 29 卷 3 号, 241-257 (2014).

(受付日 令和5年 12月 29日)

(受理日 令和6年 3月 5日)