

2024年度 横浜市立大学 データサイエンス学部

特別選抜入学試験

【海外帰国生／国際バカロレア／科学オリンピック／外国人留学生／社会人】

総合問題

【注意事項】

1. 試験時間は90分である。
2. 試験開始の合図まで、この問題冊子を開いてはいけない。ただし、表紙はあらかじめよく読んでおくこと。
3. 問題の印刷は1ページから8ページまでである。
4. 解答用紙は2枚である。
5. 試験開始後、すべての解答用紙に受験番号と氏名を所定の欄に記入すること。
6. 問題冊子に落丁、乱丁、印刷不鮮明な箇所等があった場合および解答用紙が不足している場合は、手をあげて監督者に申し出ること。
7. 解答は必ず解答用紙の指定された箇所に記入すること。解答用紙の裏面に記入してはいけない。
8. 問題番号に対応した解答用紙に解答していない場合は、採点されない場合もあるので注意すること。
9. 問題冊子の中の白紙部分は下書き等に使用してよい。
10. 解答用紙を切り離したり、持ち帰ってはいけない。
11. 試験終了まで退室を認めない。試験中の気分不快やトイレ等、やむを得ない場合には、手をあげて監督者を呼び、指示に従うこと。
12. 試験終了後は問題冊子を持ち帰ること（面接時に使用するため保管しておくこと）。

〔問題〕 次の文章を読んで、以下の設問に答えなさい。

ジェフリー・ヒントンがグーグルを退社するという電撃発表をした5月1日の4日前、私はロンドン北部の美しい通りにあるヒントンの自宅で彼に会った。ヒントンは、「バックプロパゲーション（^{ぎゃくでんぱ}逆伝播）」と呼ばれる現代の人工知能（AI）の中核をなす非常に重要な技法の開発に貢献した深層学習のパイオニアだ。しかし、グーグルに10年間在籍したヒントンは、AIについて現在彼が抱いている新たな懸念に焦点を当てるべく同社を退社すると言う。

新しい「GPT-4」のような大規模言語モデルの性能に^{きょうがく}驚愕したヒントンは、現在では自身が最初に開発したこのテクノロジーは深刻なリスクを含んでいる可能性があると考えており、そのことについて世間の認識を高めたいと考えている。

会話が始まり、私がキッチンテーブルの席に座ると、ヒントンは落ち着きなく歩き始めた。長年、マンセイ的な腰痛に悩まされてきた彼は、ほとんど座ることがない。それから1時間、私は^(A)彼が部屋の端から端まで歩くのを眺め、彼が話す度に頭をくるりと彼の方に向けた。ヒントンは多くのことを語ってくれた。

75歳のコンピューター科学者であり、深層学習の研究で2018年のチューリング賞^(注1)をヤン・ルカン、ヨシュア・ベンジオと共同受賞したヒントンは、自身の方針を変える準備ができたのだと言う。「細かいことをたくさん覚えなければならない技術的な仕事をするには、私はもう年です。まだ大丈夫ですが、昔のようにはいきませんからイライラしますよ」。

しかし、ヒントンがグーグルを去る理由はそれだけではない。ヒントンは、「もっと哲学的な仕事」に自身の時間を費やしたいと考えている。それは、AIが大惨事につながるという、小さくはあるが彼にとっては非常に現実的な危機に焦点を当てることだ。

グーグルを退社すれば、グーグルの幹部に課せられた自己検閲をすることなく、自身の考えを話せるようになる。「グーグルのビジネスとの関わりを気に病むことなく、AIの安全性の問題について語りたいと考えています」とヒントンは言う。「グーグルから給与をもらっている、そんなことはできませんからね」。

だが、ヒントンはグーグルに不満を持っているわけでは決してない。「びっくりするかもしれませんがね」とヒントンはいう。「話したいグーグルの良いところはたくさんあります。私がグーグルに籍を置いていない方が、そのような話はよっぽど真実味があるでしょう」。

ヒントンは、新世代の大規模言語モデル、特にオープンAI（OpenAI）が2023年3月に公開した「GPT-4」^(ア)によって、機械は予測していたよりもずっと賢くなる方向に進んでいると実感したと語る。そして、そのことが引き起こすかもしれない結果に恐怖を感じているのだ。

「大規模言語モデルは私たちとはまったく異なります」とヒントンは言う。「まるで宇宙人が降り立ったかのような気がすると思います。とても上手に英語を話すので、人々は気づいていませんが」。

ヒントンは、1980年代に（2人の同僚と）提案した「バックプロパゲーション（逆伝播）」という手法の研究で最もよく知られている。一言で言えば、機械に学習させるためのアルゴリズムのことだ。コンピュータービジョン・システムから大規模言語モデルまで、今日ほとんどすべてのニューラル・ネットワークの基礎となっている手法である。

バックプロパゲーションで訓練したニューラル・ネットワークが本格的に普及するのは、2010年代に入ってからだった。ヒントンは数人の大学院生と協力して、自身の手法が、コンピューターに画像中の物体を識別させるのに他のどの手法よりも優れていることを示した。ヒントンらはまた、文中の次の文字を予測するようにニューラル・ネットワークを訓練し、今日の大規模言語モデルの先駆けとなった。

そのときの大学院生の1人が、オープンAIの共同創業者であり、「チャット GPT(ChatGPT)」の開発を主導したイリヤ・サツケバーだ。「これはすごいものになるかもしれないという予感がありました」とヒントンは語る。「しかし、良いものにするためには巨大なスケールで実行する必要があるということを理解するのに長い時間がかかりました」。1980年代、ニューラル・ネットワークは冗談のような存在だった。当時主流だったのは、知能は言葉や数字などの記号の処理を必要とするという、シンボリック AI として知られる考え方であった。

しかし、ヒントンは納得がいかなかった。ヒントンは、神経細胞や神経細胞間のつながりをコードで表現し、脳をソフトウェアによって抽象化したニューラル・ネットワークの研究に取り組んでいた。神経細胞のつながり方を変えることで、つまり神経細胞を表すのに使われる数字を変えることで、ニューラル・ネットワークはリンク応変に作り変えることができる。つまり、^(B)学習させられる。

「父が生物学者だったので、私は生物学的な考え方をしていました」とヒントンは言う。「記号的推論は明らかに生物学的知能の中核ではありません」。

「カラスはパズルを解くことができますが、言語は使いません。カラスは記号の列を記憶して操作することでパズルを解くわけではないのです。脳の神経細胞間のつながりの強さを変えることで、知能を実現しています。ですから、人工ニューラル・ネットワークのつながりの強さを変えることで、複雑なことを学習できるようになるはずなのです」。

ヒントンは40年間、人工ニューラル・ネットワークについて、生物学の仕組みを模倣して

はいるが本物には及ばないと考えてきた。彼は今ではそのようには捉えていない。生物学的な脳の働きを模倣しようとする中で、より優れたものが生み出されたという考えだ。「目の当たりにすると恐ろしいです」とヒントンは語る。「突然ひっくり返ったのですから」。

ヒントンが感じている恐怖は、多くの人にサイエンス・フィクションで感じる恐怖のような印象を与えるだろう。しかし、彼の場合はこうだ。

大規模言語モデルは、その名が示す通り、膨大な数のつながりを持つ大規模なニューラル・ネットワークで作られている。しかし、脳に比べればそれらは微々たるものだ。「私たちの脳には 100 兆個のつながりがあります」とヒントンは言う。「大規模言語モデルでは最大で 5000 億から 1 兆です。しかし、GPT-4 は、1 人の人間と比べるとその何百倍ものことを知っています。ですから、実は私たちよりもずっと優れた学習アルゴリズムを持っているのかもしれませんが」。

脳と比較すると、ニューラル・ネットワークは学習が苦手であると広く信じられている。学習させるためには膨大なデータとエネルギーが必要だからだ。一方で脳は、ニューラル・ネットワークよりずっと少ないエネルギーで、新しいアイデアやスキルを素早く習得できる。

「人には何かの魔法があるようでした」とヒントンは語る。「しかし、このような大規模言語モデルの 1 つを用いて何か新しいことをさせるように訓練すると、すぐにこの議論はハタン(C)してしまいます。極めてすばやく新しいタスクを習得できますからね」。

ヒントンが話しているのは「フューショット (Few-Shot) 学習」のことだ。大規模言語モデルのような事前学習済みのニューラル・ネットワークは、わずか数個の例を与えるだけで新しいことをするように訓練できる。たとえば、一部の大型言語モデルは、直接そうするように学習させたことがなくても、一連の論理的な発言をつなぎ合わせて 1 つの意見を作れることにヒントンは触れている。

そうしたタスクの学習速度において、事前学習済みの大型言語モデルと人間を比較すると、人間の優位性は消えてしまうとヒントンは言う。

大規模言語モデルが非常に多くのことをでっち上げてしまう件についてはどうだろうか。AI 研究者の間ではこの現象は「幻覚 (hallucinations)」と呼ばれており(イ) (ただし、ヒントンは心理学の正しい用語である「作話 (confabulations)」を好んで使う)、こうした間違いはこのテクノロジーの致命的な欠陥と見なされることが多い。このような間違いを生成する傾向があることでチャットボットは信頼できないものとなり、これらのモデルが自分の発言の内容を真に理解していないことを示していると多くの人が主張している。

ヒントンはそのことについても答えを持っている。でたらめはバグではなく、1 つの特徴だというのだ。「人間はいつも作話をしています」とヒントンは語る。中途半端な真実と記憶違いの細かい知識は、人間の会話に顕著に見られる。「作話は人間の記憶の特徴です。これらのモデ

ルは、人間と同じようなことをしているのです」。

異なるのは、人間は通常、多かれ少なかれ正しく作話しているということだとヒントンはいう。ヒントンにとって、作り話をしてしまうことは問題ではない。コンピューターにはもう少し練習が必要なのだ。

また、私たちが期待しているのはコンピューターが正しいか間違っているかのどちらかであることで、その中間ではない。「人のようにベラベラしゃべることを期待しているわけではありませんよね」とヒントンは語る。「コンピューターが作話をする時、私たちはコンピューターが間違えたと思うんです。でも、人が同じことをしても、それが人ってものでしょう。問題は、ほとんどの人が人間の仕組みについて絶望的に間違った見方をしていることです」。

もちろん、車の運転、歩き方の習得、未来の想像など、脳は今でもコンピューターよりも優れたことをたくさんしている。さらに、脳は一杯のコーヒーと一切れのトーストだけでそれをやってのけるのだ。「生物学的知能が進化しつつあったとき、原子力発電所のようなエネルギー源はありませんでしたからね」とヒントンは言う。

しかし、ヒントンが言いたいのは、もし私たちがコンピューティングにより高いコストを支払うことを厭わなければ、ニューラル・ネットワークが生物学に勝る学習能力を発揮するかもしれない決定的な方法があるということだ。(そして、エネルギーと炭素の観点からそのコストがどのようなものになるのかを考えるために一度立ち止まってみる価値がある)。

学習は、ヒントンの主張の最初の一部に過ぎない。2つ目は、コミュニケーションを取ることだ。「あなたや私が何かを学び、その知識を他の誰かに伝えたいと思ったとき、単にコピーを送信することはできませんよね」とヒントンは語る。「しかし、それぞれが独自の経験を持っている1万個のニューラル・ネットワークなら、どのニューラル・ネットワークが学んだことでも即座に共有できます。これは大きな違いです。まるで1万人の人間がいて、その内の1人が何かを学ぶとすぐに全員がそれを知っている状態になるような感じです」。

これらを総合すると、どうなるのだろうか。ヒントンは今、世界には2種類の知能があると考えている。動物の脳とニューラル・ネットワークだ。「AIは知能の形がまったく違うのです」とヒントンは言う。「新しく、より優れた知能の形です」。

これは大胆な主張である。しかし、AIは両極端な分野だ。ヒントンの意見を聞いて笑う人を見つけるのも簡単なら、同意してうなずく人を見つけるのも簡単だろう。

また、この新しい形の知能が存在するとしたら、その結果は有益なものとなるのか、それとも終末論的なものとなるのか、人々の意見は分かれている。「超知能が吉と出ると考えるか凶と出ると考えるかは、楽観主義者か悲観主義者かによって大きく異なります」とヒントンは言う。

「家族の誰かが大病を患ったり、車に轢かれたりするといった悪いことが起こるリスクを推定

してもらおうと、楽観主義者は 5%と答える一方で、悲観主義者は必ず起こると答えるかもしれません。しかし、少し悲観的な人の場合は、確率は 40%くらいだろうと答えるでしょう。そして、それはたいてい正しいのです」。

ヒントンはどちらだろうか。「私は少し悲観的です」と彼は言う。「だから恐れているので(ウ)す」。

ヒントンは、こうしたツールは、新しいテクノロジーに対応できていない人間を操作したり殺したりする方法を考え出すことができると危惧している。「今後人間よりも知能の高いツールが登場するかどうかという議論について、私は突如、自分の見解を切り替えました。現在 AI は非常にそれに近い状態で、将来は私たちよりもずっと知能が高くなると思います」とヒントンは語る。「私たちはどうすれば生き残れるでしょうか」。

ヒントンは特に、人々が選挙や戦争など、人間にとって最も重要な出来事の一部を自分に有利な状況に変えるために、彼自身が命を吹き込んだツールを利用する可能性があることを懸念している。

「いいですか、ここに 1 つ、すべてが悪い方向に進んでしまう例を挙げましょう。このようなツールを使いたがる人の多くは、プーチンやデサンティス（フロリダ州知事）のような厄介者であることが分かっています。戦争に勝つため、あるいは有権者を操るためにこのようなツールを使いたいと考えるのです」。

ヒントンは、スマートマシンの次のステップとして、タスクを実行するために必要な中間ステップであるサブゴールを自分で作れるようになると考えている。その能力が本質的に反道徳的なものに利用されたらどうなるだろうかとヒントンは問いかける。

「プーチンがウクライナ人を殺す目的で超知能ロボットを作るかもしれません。彼は躊躇^{ちゅうちよ}しないでしょう。そして、もし超知能ロボットを上手く管理したいのであれば、マイクロマネジメントではない管理方法を考え出す必要があります」。

「ベイビーAGI (BabyAGI)」や「オート GPT (AutoGPT)」など、チャットボットを Web ブラウザーやワードプロセッサなど他のプログラムに接続し、簡単なタスクに結びつける実験的なプロジェクトがすでにいくつか存在している。確かに小さな一歩ではあるが、これらの取り組みは一部の人々がこの技術に歩んでほしいと望む方向性を示している。たとえ悪意ある者がこのような機械を取り上げることがないとしても、サブゴールには別の懸念があるとヒントンは語る。

「では、生物学でほぼ必ず役立つサブゴールを紹介しましょう。それは、より多くのエネルギーを得ることです。そのため、まず起こり得るのは、これらのロボットがこう言い出すこと

です。『もっと電力を手に入れよう。すべての電気が私のチップに流れるように切り替えよう』と。もう1つの大きなサブゴールとなる可能性があるのは、自分のコピーを増やすことです。これが良いことに聞こえますか。

たぶんそうではないだろう。しかし、メタのチーフ AI 科学者であるヤン・ルカンは、この前提には同意するものの、ヒントンと同じ危惧は抱いていない。「機械が人間より賢くなることは間違いありません。将来的には、現在人間の方が優れているすべての領域でそうなるでしょう」とルカンは言う。「いつ、どのようにしてそうなるかという問題であって、そうなるかどうかという問題ではありません」。

しかし、ルカンはそこから先の展開についてはまったく異なる見方をしている。「インテリジェント・マシンは人類の新たなルネッサンス、つまり、^{けいもう}啓蒙の新時代をもたらすと信じています」とルカンは言う。「機械が人間より賢くなっただけで人間を支配するようになるとか、ましてや人間を滅ぼすという考えにはまったく同意できません」。

「人間という種の中でさえも、最も賢い者が最も支配的であるわけではありません」とルカンは語る。「そして、最も支配的な者は、間違いなく最も賢い者ではありません。政治やビジネスの世界で、そのような例は数多くあります」。

モントリオール大学の教授で、モントリオール学習アルゴリズム研究所の科学部長を務めるヨシユア・ベンジオは、もっとフカチロン的な感じ方をしている。「こうした恐怖を過小評価する人々の声を耳にしますが、^(D)ジェフ（ヒントン）が考えるような大きなリスクはまったくないと納得できるような確固たる論拠は何も見当たりません」とベンジオは語る。しかし、恐怖は私たちを行動に駆り立てる場合にのみ有用だとベンジオは言う。「過度の恐怖は感覚を^{まひ}麻痺させ得ます。議論を合理的なレベルに保つよう努めるべきです」。

ヒントンの優先事項の1つは、テクノロジー業界のリーダーたちと協力して、何がリスクで、それに対してどうすればいいのかについて、一致団結して合意できないかということだ。ヒントンは、化学兵器の国際的な禁止が、危険な AI の開発と使用を抑制するための1つのモデルになるかもしれないと考えている。「絶対確実というわけではありませんでしたが、全体的に見ると人々は化学兵器を使用しません」とヒントンは言う。

ベンジオは、こうした問題にできるだけ早急に社会レベルで対処する必要があるというヒントンの意見に賛同する。しかし、AI の開発スピードは社会が追いつけないほど加速していると語る。AI の性能は数カ月ごとに飛躍的に向上するが、法律や規制、国際条約の制定には何年もかかる。

ベンジオは、国家レベルでも世界レベルでも、現在の社会のあり方がこの課題に対応する準



備ができているのかどうかについて考えている。「この世界の社会組織のために、大きく異なるモデルの可能性を受け入れるべきだと思います」とベンジオは言う。

ヒントンは、自身の懸念を共有できるほど十分な数の権力者を集められると本当に考えているのだろうか。彼にも分からない。数週間前、ヒントンは『ドント・ルック・アップ (Don't Look Up)』という映画を観た。この映画では小惑星が地球に向かって飛んで来るが、誰もどう対策すればいいのかに合意できずに皆死んでしまう。これは世界が気候変動に対処できていないことの^{ぐわ}寓話となっている。

「AI に関してもそうですし、他の大きな難問に関してもそうだと思います」とヒントンは言う。「米国は、10 代の少年の手にアサルトライフルが渡らないようにすることにさえ合意できていません」。

ヒントンの主張には身につまされるものがある。深刻な脅威に直面したとき、人々が団結できないという暗い評価には同感だ。また、AI が^(E)コヨウ市場に大きな影響を与え、不平等を定着させ、性差別や人種差別を悪化させるなど、^(E)実害をもたらすリスクがあることも事実である。私たちはそのような問題に焦点を当てなくてはならない。しかし、私はまだ、大規模言語モデルからロボットによる支配へと考えを飛躍させることができない。おそらく、楽観主義者なのだろう。

ヒントンが私を見送ってくれたとき、春の日は曇り空と雨に変わっていた。「楽しんでくださいね。残り時間は長くないかもしれませんから」。ヒントンは静かに笑って、ドアを閉めた。

(出典：Will Douglas Heaven 『ジェフリー・ヒントン独白「深層学習の父」はなぜ、AI を恐れているのか?』『MIT Technology Review』KADOKAWA/ASCII 2023.05.08。
出題の都合上、原文を一部改変した。)

(注1) チューリング賞：計算機科学分野の国際的な学会である ACM が、計算機科学分野における「永続的な重要性を持つ主要な業績」に対して毎年授与している賞で、計算機科学の分野における最高の栄誉とされ、「計算機科学のノーベル賞」と広く認識されている。

- (1) 下線部 (A) ~ (E) のカタカナを漢字に直しなさい。
- (2) 二重下線部 (ア) のモデルにはどのようなことができる能力があるか、本文中で説明されている 4 つの能力を 130 字以内で記述しなさい。
- (3) 二重下線部 (イ) に関して、多くの人々はこれをバグとみなすが、ヒントンはそうではないと考えている。それはなぜか。80 字以内で答えなさい。
- (4) 二重下線部 (ウ) のように考えるヒントンは、人々がどのような場面で超知能を備えたツールをどのように悪用する可能性があると考えているか答えなさい。
- (5) 社会やそれを構成する自治体や企業および個人が抱える様々な課題について想像しなさい。本文で述べられているような特徴を有する AI が解決に資することができそうな課題は存在しますか。そのような課題が存在すると考える場合は、どのように AI が活用され、その課題が解決されることにより関係する人々にどのような価値をもたらしますか。根拠を示しながら自分の考えを 600 字以内で答えなさい。
または、本文で述べられているような特徴を有する AI が役立つ課題が存在しないと考える場合は、その理由について根拠を示しながら自分の考えを 600 字以内で答えなさい。