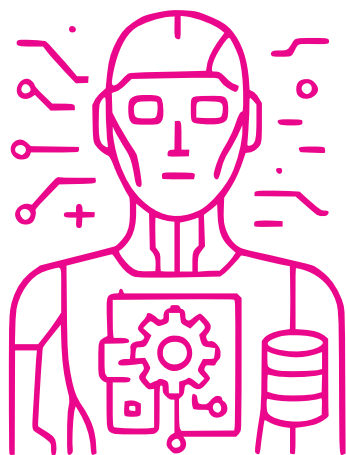


Człowiek o AI (2)

*Instytut Informatyki, Uniwersytet Wrocławski

Autor tego tekstu odnosi się do wywiadu, który można znaleźć pod adresem: www.pap.pl/aktualnosci/prof-sankowski-matematyka-bedzie-pierwsza-dziedzina-gdzie-ai-osiagnie-mozliwosci.
Uwaga: jest pełnym przypadkiem, że akurat w tym samym numerze *Delty* znajduje się inny artykuł – autorstwa właśnie profesora Sankowskiego. Obaj autorzy nie mieli wiedzy na temat tego, kto został poproszony o sąsiedzki tekst dotyczący AI na łamy *Delty*. Oba teksty powstawały więc równolegle, bez wzajemnego wglądu (przyp. red.).



Osiem krótkich uwag z pamiętnika

Jerzy MARCINKOWSKI*

1

Mówią nam różni „prorocy” AI, że matematyka już niedługo podzieli los szachów.

Zdolność do gry w szachy wydawała się kiedyś istotą inteligencji. Nie da się, myśleliśmy, dobrze grać w szachy, nie rozumiejąc pozycji i nie konstruując planów gry. Planów, które dopiero potem wypełniają się mechanicznym liczeniem wariantów.

No i okazało się, jak Czytelnik już wie, że nie ma dziś na świecie nikogo, kto wygrałby w szachy ze swoim telefonem. Chyba, że ma bardzo stary telefon. A telefon nie rozumie pozycji. Tak nam się przynajmniej wydaje, że nie rozumie. On liczy warianty.

Podobnie, mówią nam, będzie z matematyką. Na przykład w tytule wywiadu dla Polskiej Agencji Prasowej z twórcą sławnego ostatnio warszawskiego instytutu IDEAS, profesorem Piotrem Sankowskim, pojawia się stwierdzenie, że „matematyka będzie pierwszą dziedziną (nauki), gdzie AI osiągnie możliwości nadludzkie”.

Może tak będzie, a może nie będzie. Nie wiem. Profesor Sankowski też nie wie. Nikt nie wie.

2

Pisząc w tym tekście „AI”, mam na myśli wielkie modele językowe (LLM) wzbogacone być może (tak jak to się teraz robi) o jakiś mechanizm sprawdzania poprawności. Taki model działa tak, że przeczytał wszystkie dostępne na świecie teksty. I produkuje, gdy go o coś pytamy, nowy tekst, najbardziej podobny, w sensie pewnych skomplikowanych miar podobieństwa, do wszystkiego, co w kontekście naszego zapytania spodziewałby się w tych tekstach zobaczyć. Model nie interesuje się światem i nie wie nic o nim. Ponieważ nie wie nic o świecie, nie zna pojęcia prawdy, bo prawda się definiuje poprzez porównywanie zdań ze światem. A on zna tylko zdania, nie zna świata.

Profesor Sankowski, podobnie zresztą jak wielu innych uczonych, którzy na pewno wiedzą o AI więcej niż ja, uważa, że akurat w badaniach matematycznych ten brak doświadczenia świata nie jest przeszkodą. „Dziedziną badań, na którą sztuczna inteligencja będzie miała największy wpływ” – pisze w cytowanym wywiadzie dla PAP – „jest matematyka. Ta dziedzina istnieje bez konieczności interakcji ze światem fizycznym”.

Czyli że można naczytać się tekstów matematycznych i jeśli się ich przeczyta wystarczająco dużo, to potem można tworzyć nowe teksty matematyczne, zawierające odkrycia matematyczne na poziomie nadludzkim.

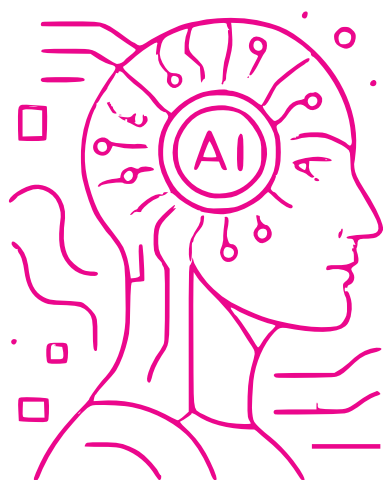
Hipotezy Riemanna jeszcze AI w ten sposób nie udowodniła (ale, bądźmy sprawiedliwi, ludzie też nie). Jednak, jak się ostatnio dowiedzieliśmy, prawie umiałyby wygrać Międzynarodową Olimpiadę Matematyczną (IMO). Rozwiązałyby pięć z sześciu zadań.

Pozwólcie, że się podzielę w związku z tym pewną anegdotą, a następnie pewną myślą. Zapewne nie jestem pierwszy, któremu ta myśl przyszła do głowy.

3

Zacznę od anegdoty. Otóż kiedyś, ponad 40 lat temu, brałem udział w tym IMO-sporcie. Nawet medal zdobyłem, brązowy. Lubiłem rozwiązywać zadania. Mieliśmy w liceum kółko matematyczne. Świętej pamięci profesor Kałuża dawał nam co tydzień listę zadań. I ja potem chodziłem przez tydzień i o nich myślałem. I zwykle gdzieś z połowę zadań z listy udawało mi się przed kolejnym kółkiem (na którym pokazywaliśmy sobie nasze rozumowania) rozwiązać.

A mój kolega, nazwijmy go X, który zdobywał na IMO medale już od drugiej klasy liceum, mówił do mnie z wyższością: „Ja nigdy nie rozwiązałem sam żadnego zadania, ani na kółko, ani na olimpiadzie; ja sobie przypominam



rozwiązania”. Rzeczywiście, w czasie gdy ja chodziłem i myślałem o zadaniach, X studiował sowieckie *Kwanty* (to był taki miesięcznik, w którym były wszystkie możliwe zadania z wcześniejszych olimpiad, na całym świecie) i czytał te zadania, po kolei, z rozwiązaniami. A że pamięć miał dobrą, to znał je potem wszystkie. I na olimpiadzie zawsze się okazywało, że coś się da, z większymi lub mniejszymi zmianami, przypasować. Bardzo to była skuteczna metoda, bo czyta się przecież dużo szybciej niż samemu rozwiązuje. Tylko że ja lubiłem myśleć o tych zadaniach, a czytać nie lubiłem, szczególnie po rosyjsku.

Teraz będzie ta myśl, którą wcześniej zapowiadałem.

Otóż wyobraźmy sobie, że matematycy gawędzący sobie w roku 1935 w kawiarni *Szkocka* we Lwowie dostaliby do rozwiązania zadania z IMO 2025. Nie wiem, czy jeden wieczór by im wystarczył, czy potrzebowaliby powalczyć kilka dni, ale bez wątplenia rozwiązaliby te wszystkie zadania. Nawet gdyby Banach był z jakiegoś powodu akurat nieobecny. Do tego, żeby je rozwiązać, wystarczyłoby im bycie kompetentnym matematykiem z lat trzydziestych XX wieku.

Więc ja bym chciał, żeby twórcy AI uczyli te swoje „potwory”, używając całego tekstu matematycznego, jaki powstał DO ROKU 1935. A potem kazali im rozwiązywać zadania z IMO 2025. To by znaczyło, że w tym korpusie tekstu, na którym maszyna się uczy, nie byłoby tych wszystkich *Kwantów*, tych tysięcy zadań z różnych olimpiad wraz z rozwiązaniami. Bo (mam nadzieję, że nie trzeba tego wyjaśniać) wtedy nie było jeszcze olimpiad matematycznych.

Myślę (mogę się oczywiście mylić, okropne by było, gdybym się mylił, ale wątpię żebym się mylił), że sztuczna inteligencja nauczona na takim ograniczonym korpusie tekstu rozwiązałaby najwyżej jedno zadanie z IMO 2025, to z geometrii. Bo geometria jest z nami od bardzo dawna.

4

Ciągnąc dalej tę myśl: jeśli wytworzony dotąd, przez ludzi, korpus tekstu matematycznego jest (jak wydają się twierdzić prorocy AI) wystarczający do wyhodowania modeli, które będą już w stanie zostawić nas w tyle i iść z matematyką dalej, z nadludzką mocą, to kiedy był pierwszy moment, gdy ten korpus tekstu zrobił się dla tego celu wystarczający?

Czy był on już wystarczający w 1935 roku? Jeśli tak, to – powtarzam – dlaczego nie nauczycie modeli na tym korpusie i nie pokażecie mi, że one tak nauczzone rzeczywiście umieją rozwiązywać zadania z IMO?

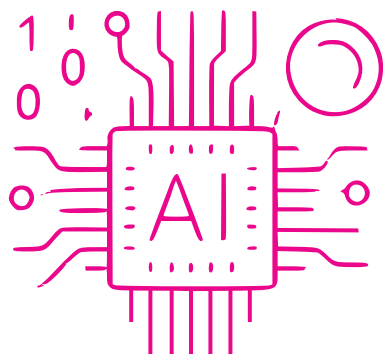
A jeśli stał się wystarczający dopiero niedawno, to zdumiewa mnie ten zbieg okoliczności, że nie tylko żyjemy w czasach, kiedy człowiek zaprojektował nadludzko zdolną sztuczną inteligencję, ale że w dodatku powstała ona akurat wtedy, gdy ludzie wyprodukowali korpus tekstu matematycznego dostatecznie duży, aby ta sztuczna inteligencja mogła się na nim nauczyć swoich nadludzkich mocy.

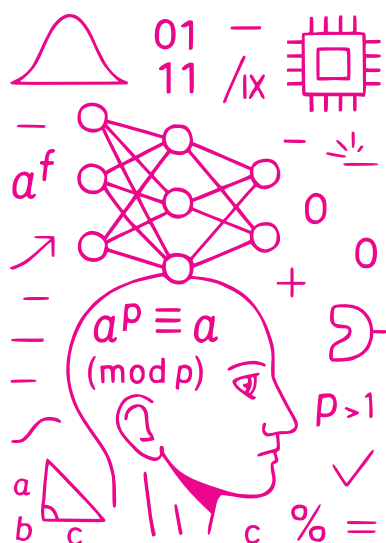
5

Równolegle z tym tekstem piszę artykuł zawierający dowód pewnego twierdzenia, niezbyt trudnego. Dowód opiera się na czterech łatwych lematach i kilkunastu obserwacjach. Artykuł zawiera dowody lematów. Te dowody zaś odwołują się do obserwacji i są krótkie, bo – jak wspomniałem – to są łatwe lematy. Obserwacje nie mają dowodów, bo wydają mi się na tyle proste, że Czytelnikowi będzie łatwiej wyobrazić sobie odpowiedni dowód, niż go czytać.

Przy jednej z tych obserwacji zawahałem się. Być może – pomyślałem – powinna ona być lematem. Z dowodem na kilka linijek.

I kiedy chciałem napisać ten dowód, uświadomiłem sobie, że choć widzę go doskonale, to język, w którym sformułowane jest całe moje rozumowanie, cały formalizm, który wprowadziłem, żeby ponazywać różne obiekty, że to jest wszystko niewystarczające, aby ten (zupełnie trywialny) dowód napisać. I żeby go zatem napisać, nie tylko musiałbym się cofnąć o kilka stron i przerobić od tego miejsca cały tekst (co byłoby do zniesienia, jeśli chcecie napisać dobry artykuł, to musicie się pogodzić z tym, że będziecie się cofać i przerabiać), ale





w dodatku musiałbym skazać Czytelnika na czytanie kilku stron cięższego, mniej czytelnego tekstu, tylko po to, żeby raz użyć tego cięższego formalizmu dla zapisania dowodu lematu, który to dowód można sobie łatwo wyobrazić bez nazywania występujących w nim obiektów.

Wspominam o tym, bo wydaje mi się bardzo dobrą ilustracją tezy, że matematyka nie jest tożsama z tekstem matematycznym i że tekst matematyczny jedynie (lepiej lub gorzej) oprowadza nas po pewnym świecie, który to świat możemy czasem zobaczyć lub wyobrazić sobie, bez pośrednictwa języka, a więc (jak mogłoby się wydawać) w sposób niedostępny dla modeli językowych.

6

Wierzę, tak jak wierzył Platon 2400 lat temu, a potem wielu współczesnych sławnych matematyków, między innymi Kurt Gödel, Paul Erdős i Roger Penrose, że świat matematyczny istnieje niezależnie od tekstu. Według tego poglądu, matematyka istnieje obiektywnie i odwiecznie, niezależnie od naszej obecności w świecie i od naszej zdolności do wyrażania myśli w języku. Jeśli gdzieś daleko w kosmosie istnieje inteligentne życie, to – jeśli myślimy jak Platon – spodziewalibyśmy się, że tamtejsi matematycy odkrywają matematykę bardzo podobną do naszej.

Ale są też tacy matematycy, którzy uważają, że matematyka jest jedynie precyzyjnym tekstem. Tak myślał, na początku XX wieku, wielki David Hilbert i tak, zdaje się, myśli profesor Sankowski. Ja mam z tym poglądem problem, bo po pierwsze on się kłóci z moją głęboką intuicją, a po drugie on nie wyjaśnia, czemu matematyka tak dobrze tłumaczy fizyczną rzeczywistość. Skoro matematyka nie ma nic wspólnego z rzeczywistym światem, to po co uczą się jej inżynierowie? Czy nie jest jednak tak, że te mosty wiszą dzięki matematyce?

7

Platon mógł się oczywiście mylić, a ja mogę się mylić wraz z nim. Ale jeśli się jednak nie mylimy, to wątpliwe jest, aby sztuczna inteligencja, oparta na pomysle generowania najbardziej prawdopodobnie wyglądającego tekstu, mogła rzeczywiście dokonywać wielkich odkryć matematycznych.

Ale to nie znaczy, że nie zmieni ona sposobu, w jaki rozwija się matematykę.

Bo większość z tych tysięcy artykułów pisanych i publikowanych przez matematyków nie zawiera wielkich odkryć, ani w ogóle nie zawiera istotnie nowych idei.

Czasem są to wariacje na temat odkryć, których ci sami ludzie dokonali wcześniej. Mogą oni na przykład stosować wcześniej odkrytą przez siebie technikę do rozwiązania jakiegoś nowego problemu. Taki charakter ma artykuł, o którym wspominałem parę akapitów temu.

Czasem zaś, myślę, że bardzo często matematycy wyważają drzwi, które są już bardziej lub mniej otwarte. Potrzebna im technika została już odkryta i użyta w jakimś odrobinę innym kontekście. Człowiek nie jest w stanie tego dostrzec, bo nie może śledzić wszystkiego, co kiedykolwiek opublikowano. AI natomiast, która zna cały Internet, będzie w stanie (w jakiś magiczny sposób, którego nikt dobrze nie rozumie) zobaczyć analogię i albo podpowiedzieć nam pomysł, albo samodzielnie stworzyć dowód.

8

Matematyka zostanie zatem niejako „domknięta” na operację automatycznej analogii. Można myśleć, że to dobrze. Że stanie się to, co od wieków dzieje się w różnych innych dziedzinach życia, że maszyny zastępują nas w pracy nudnej i rutynowej, pozostawiając nam to, co twórcze i ciekawe.

Ale to nie jest takie proste.

Bo pisząc te mało twórcze prace i ponownie odkrywając znane już w innym kontekście techniki, ludzie nabywają umiejętności potrzebnych do robienia rzeczy trudniejszych. I ja nie wiem, czym ten proces uczenia się będzie można zastąpić.

Zauważmy, na zakończenie, że powyższe odnosi się analogicznie do AI zastępującej ludzi w programowaniu.

