

Paradoks Wielkich Modeli Językowych

Mistrzowie złożoności, amatorzy prostego rozumowania



- Tworzenie poezji i generowanie kodu
- Zdawanie egzaminów prawniczych i medycznych
- Tłumaczenie języków z niezwykłą płynnością



- Proste zadania matematyczne z podstawówki

Przykład: »Kawiarnia miała 23 jabłka. Wykorzystano 20 do lunchu i dokupiono 6. Ile jabłek mają teraz?«

Odpowiedź LLM: **27. (Błędna)**

Prawidłowa odpowiedź: **9.**

Fundamentalna luka:

Doskonałość w statystycznym dopasowywaniu wzorców kontra słabość w wieloetapowym, sekwencyjnym rozumowaniu. Ten paradoks jest głównym tematem przełomowej pracy badawczej Google Research.

Dlaczego modele zawodzą? Anatomia standardowego promptowania



Opis mechanizmu

- Standardowe promptowanie to bezpośrednia transakcja: Pytanie → Odpowiedź.
- Model jest zmuszony do przeskoczenia od złożonego problemu do ostatecznej odpowiedzi w jednym kroku.
- Brak miejsca na pośrednie etapy rozumowania, dekompozycję problemu czy »pokazanie toku myślenia«.

Analiza błędu na przykładzie jabłek

- Model widzi liczby 23 i 6 oraz słowa »dokupiono« / »więcej«.
- Stosując dopasowanie wzorców, wykonuje najprostszą operację: $23 + 6 = 27$. **X**
- Całkowicie ignoruje kontekst i sekwencję operacji (najpierw odjęcie 20).

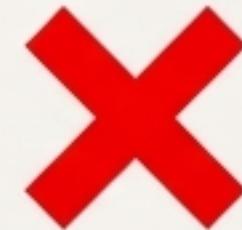
Wniosek: Podejście »jednego skoku« zawodzi, zwłaszcza gdy problem zawiera logiczne pułapki lub wymaga wieloetapowego myślenia. To tryb awaryjny »zero-shot«.

Rewolucyjne rozwiązanie: Chain of Thought Prompting

Standard Prompting

P: ...ile jabłek mają teraz?

O: Odpowiedź to 27.



Chain of Thought Prompting

P: ...ile jabłek mają teraz?

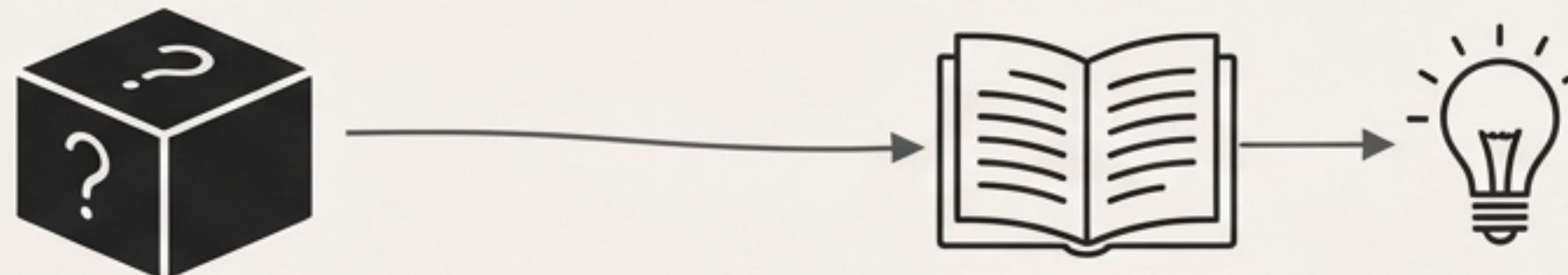
O: Kawiarnia miała pierwotnie 23 jabłka. Wykorzystali 20, więc zostało im $23 - 20 = 3$. Dokupili 6 kolejnych jabłek, więc mają $3 + 6 = 9$. Odpowiedź to 9.



- Zamiast pytać wprost, pokaż modelowi przykłady z rozumowaniem krok po kroku.
- Model uczy się wzorca: Pytanie → **Kroki Rozumowania** → Odpowiedź.
- W obliczu nowego problemu, model spontanicznie generuje podobny łańcuch logiczny, aby dojść do odpowiedzi.

Zmiana paradymatu: Od wyroczni do ucznia

Jak uczymy modele myśleć, a nie tylko odpowiadać

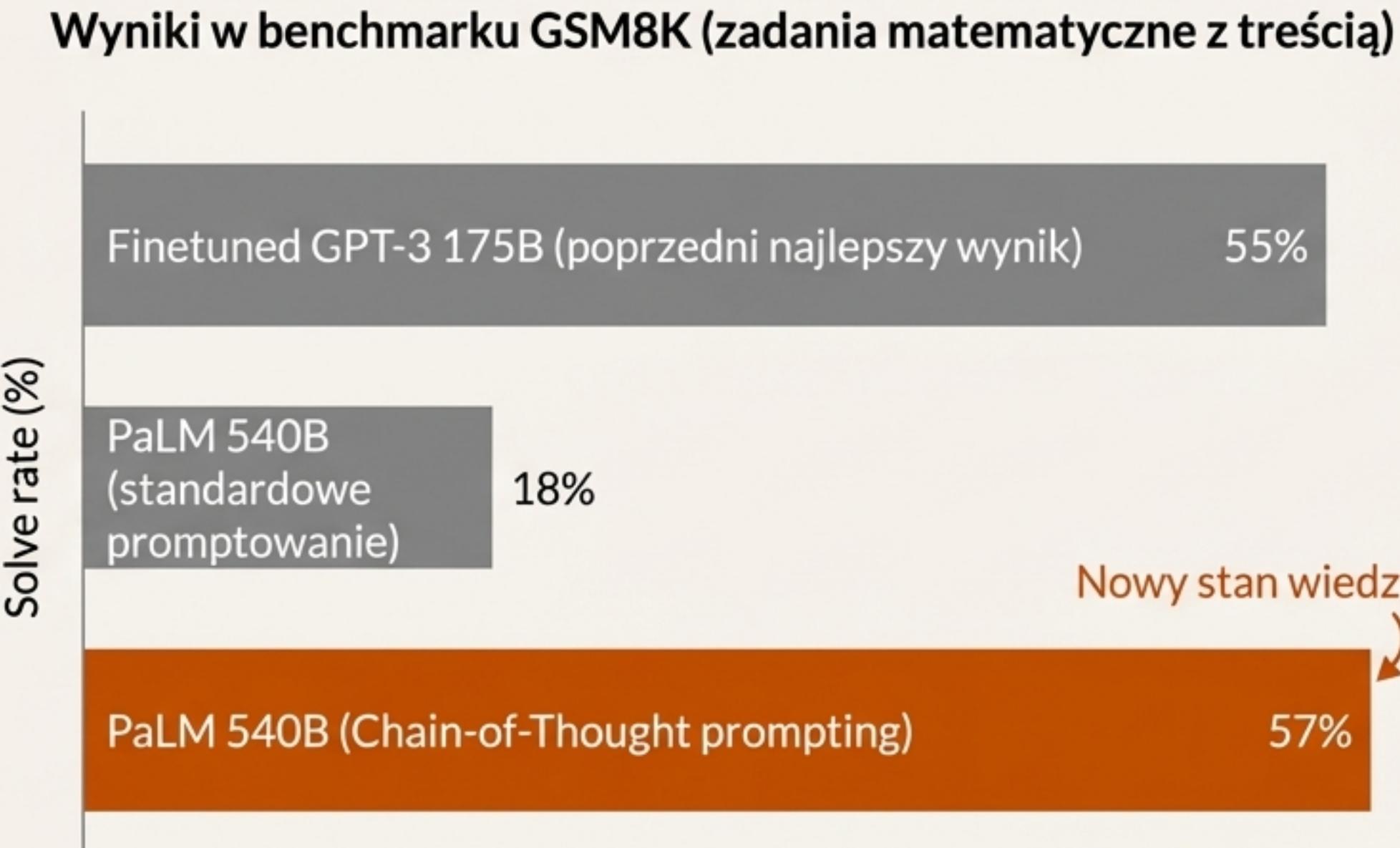


Metoda Tradycyjna (Wyrocznia)	Metoda Chain of Thought (Uczeń)
Pokaż: Pytanie → Odpowiedź	Pokaż: Pytanie → Proces myślowy → Odpowiedź
Uczzenie: Zapamiętywanie faktów	Uczzenie: Metodologia rozwiązywania problemów
Cel: Uzyskanie prawidłowego wyniku	Cel: Odblokowanie ukrytej zdolności do dekompozycji

Nie chodzi o wstrzykiwanie nowej wiedzy, ale o aktywację utajonych zdolności.

Wystarczyło tylko 8 przykładów z łańcuchem myśli, aby nauczyć model PaLM o 540 miliardach parametrów nowej, złożonej umiejętności.

Przełomowe wyniki: PaLM 540B deklasuje rywali



Nowy stan wiedzy (State-of-the-Art): PaLM 540B zaledwie 8 przykładami CoT osiągnął najwyższą dokładność.

Pokonanie specjalistów: Model ogólnego przeznaczenia pokonał precyjnie dostrojony (*fine-tuned*) model GPT-3, wspomagany zewnętrznymi programami weryfikującymi.

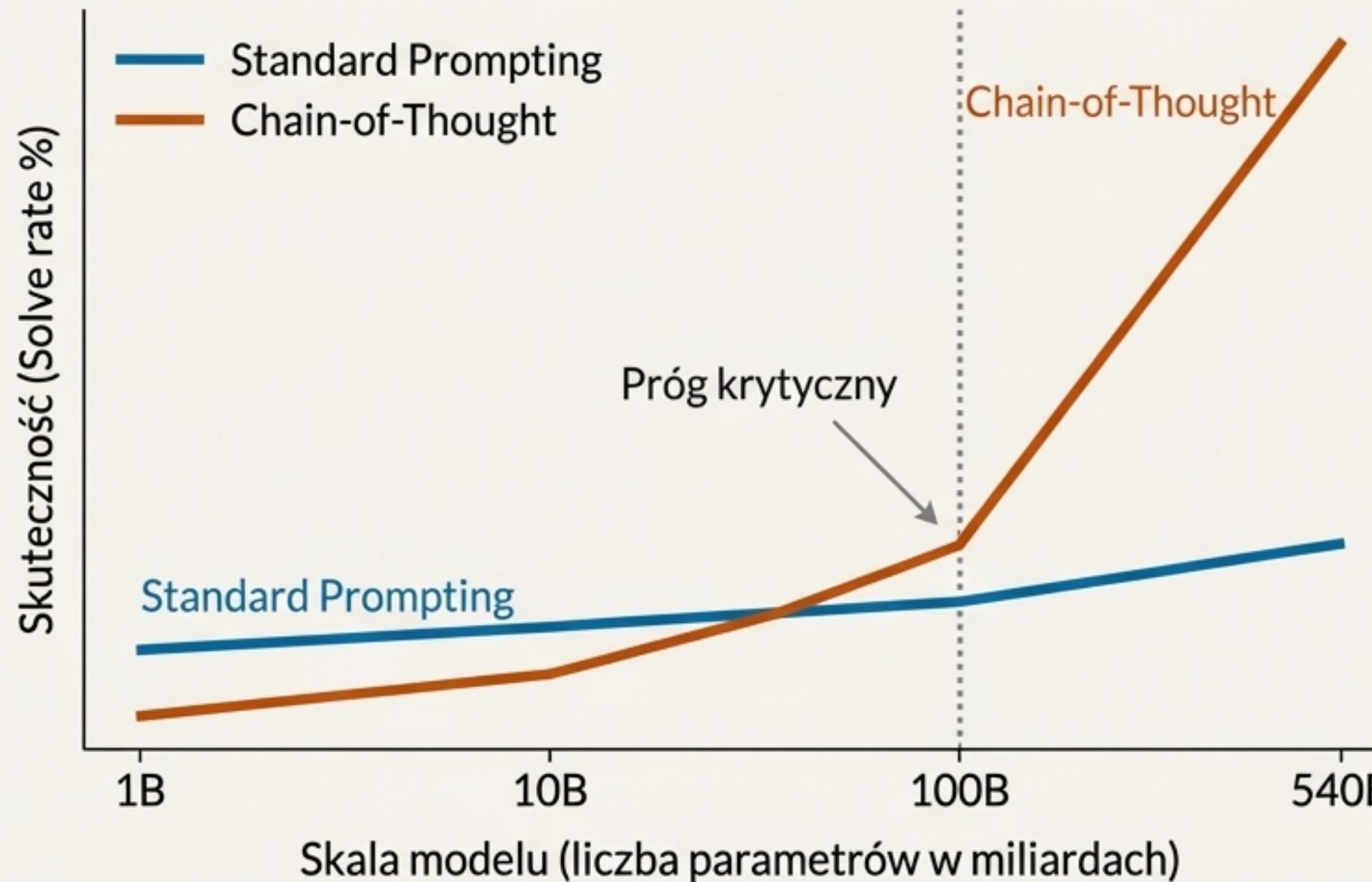
Bez potrzeby dotrenowywania: Jedyna zmiana to inżynieria promptu, a nie kosztowne trenowanie całego modelu.

»*To jak odkrycie ukrytego biegu w silniku, który mieliśmy od samego początku.*«

*Dane na podstawie »Chain-of-Thought Prompting Elicits Reasoning in Large Language Models«, Google Research.

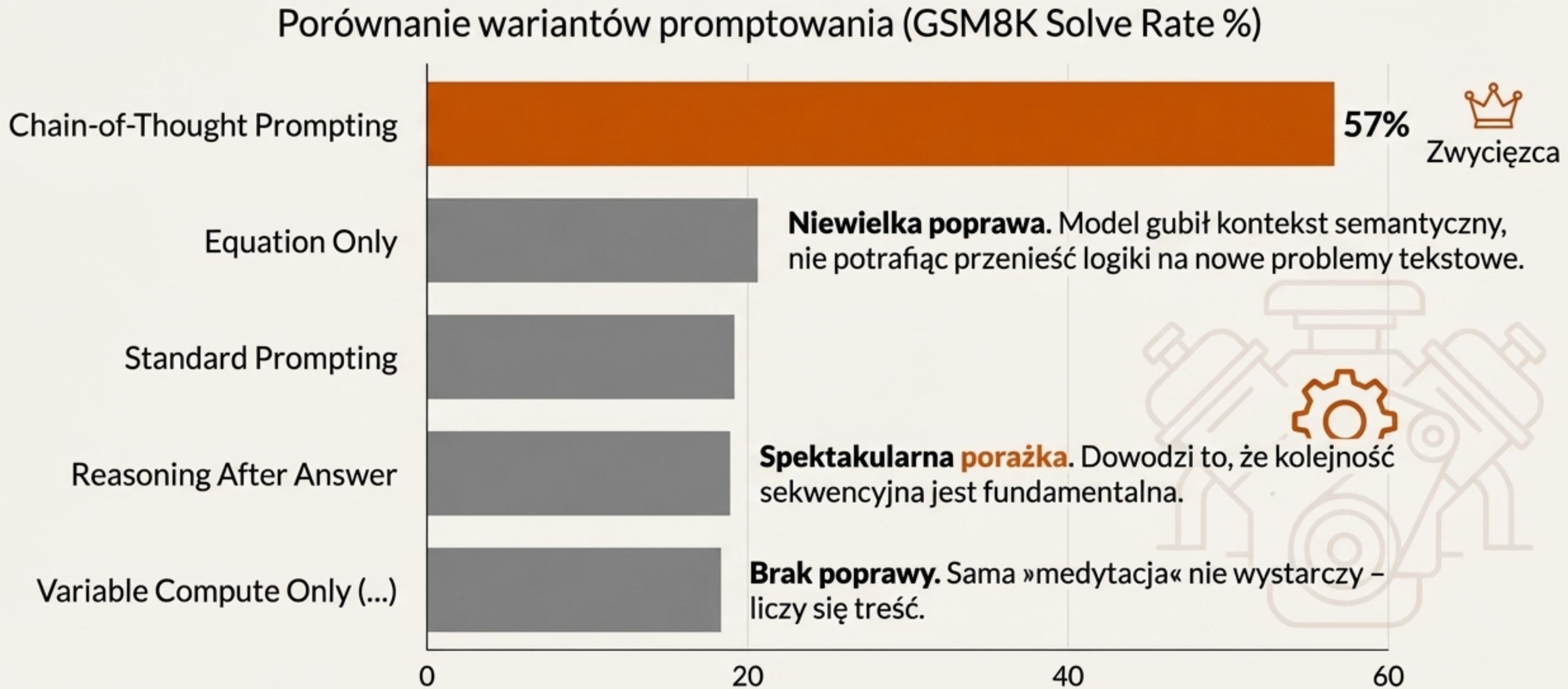
Zdolność Emergentna: Magia pojawia się przy 100 miliardach parametrów

Skuteczność w zależności od skali modelu



- Chain of Thought to **zdolność emergentna** – pojawia się gwałtownie dopiero przy krytycznej skali modelu.
- Poniżej progu ~100 miliardów parametrów, technika ta **pogarsza wyniki**.
- Mniejsze modele generują tekst, który *wygląda jak rozumowanie, ale jest płynnym, gramatycznym nonsensem z błędą logiką*.
- Zdolność do **spójnego rozumowania logicznego** pojawia się jako funkcja skokowa po przekroczeniu progu skali.

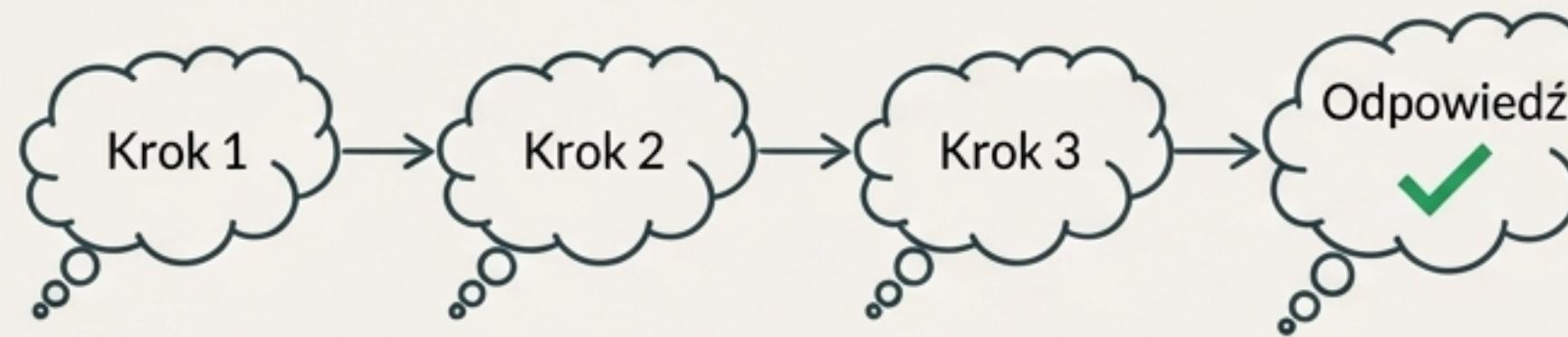
Co tak naprawdę działa? Badania ablacyjne



Dowód na autentyczne rozumowanie: Sekwencja ma znaczenie

Dlaczego porażka wariantu »rozumowanie po odpowiedzi« jest tak ważna?

Autentyczny Proces Myślowy



Racjonalizacja Post-Hoc



Główny wniosek

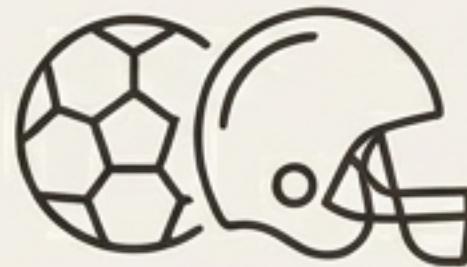
- Gdy model najpierw odpowiada, a potem próbuje »wyjaśnić« dlaczego, jego wyniki są tak samo słabe jak przy standardowym promptowaniu.
- To dowodzi, że łańcuch myśli **nie jest racjonalizacją post-hoc** (dopisywaniem uzasadnienia do już podjętej decyzji).

Jak to działa naprawdę?

- Model autentycznie **używa każdego wygenerowanego kroku jako danych wejściowych do obliczenia następnego kroku**.
- Jest to prawdziwy, sekwencyjny proces myślowy, który prowadzi do odpowiedzi.
- Implikacja: Obserwujemy załączki prawdziwego, proceduralnego rozumowania.

Uniwersalne zastosowanie: Od zdrowego rozsądku po logikę symboliczną

Łańcuchy myśli w języku naturalnym działają w każdej dziedzinie rozumowania.



Rozumowanie zdroworozsądkowe (Commonsense Reasoning)

Pytanie: Czy João Moutinho złapał podanie w mistrzostwach NFC?

Łańcuch myśli modelu: João Moutinho to portugalski piłkarz → piłkarze grają w piłkę nożną (soccer) → NFC to mistrzostwa futbolu amerykańskiego → **Odpowiedź: Nie.**

Wniosek: Model łączy wiedzę z dwóch odrębnych dziedzin, aby wyciągnąć logiczny wniosek.

$$\begin{matrix} L \\ G \end{matrix} \rightarrow \begin{matrix} y \\ a \end{matrix}$$

Rozumowanie symboliczne (Symbolic Reasoning)

Pytanie: Weź ostatnie litery słów w »Lady Gaga« i połącz je.

Łańcuch myśli modelu: Ostatnia litera »Lady« to »y«. Ostatnia litera »Gaga« to »a«. Połączenie ich daje »ya«. → **Odpowiedź: ya.**

Wniosek: Model nauczony na przykładach 2-wyrazowych nazwisk potrafi **perfekcyjnie generalizować** tę umiejętność na nazwiska 3- i 4-wyrazowe.

Ograniczenia i otwarte pytania

Mimo przełomu, droga do powszechnego zastosowania jest jeszcze daleka.



Główne ograniczenie: Wymóg Skali

Rozumowanie typu Chain of Thought jest zdolnością emergentną, która pojawia się **tylko w modelach o skali 100B+ parametrów**.

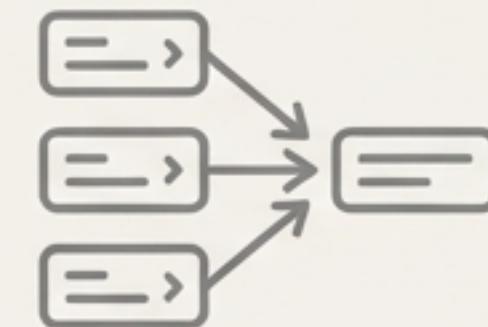
To sprawia, że technika jest obecnie kosztowna obliczeniowo i niedostępna dla większości badaczy i deweloperów.

Brak gwarancji poprawności – wygenerowany łańcuch myśli może zawierać błędy, które prowadzą do błędnej odpowiedzi.



Otwarty problem badawczy

Chociaż CoT naśladuje ludzki proces myślowy, nie odpowiada to na pytanie, czy sieć neuronowa faktycznie »rozumie« w ludzkim tego słowa znaczeniu.



Wrażliwość na przykłady

Jakość i styl przykładów użytych w prompcie (tzw. »exemplars«) mają wpływ na wynik, chociaż badania (Figure 6) pokazują, że metoda jest stosunkowo odporna na różnych autorów i style.

Nowe horyzonty: Co dalej z rozumowaniem w LLM?

Ograniczenia CoT definiują najważniejsze kierunki przyszłych badań.

Demokratyzacja rozumowania

Jak możemy wywołać lub symulować zdolności rozumowania łańcuchowego w mniejszych, bardziej dostępnych modelach? Czy możliwe jest »destylowanie« tej umiejętności z dużych modeli do małych?

Odkrywanie ukrytych zdolności

Jakie inne utajone zdolności, podobne do CoT, mogą być ukryte w wielkich modelach językowych i czekają na odblokowanie przez odpowiednie techniki promptowania?

Czy istnieją »klucze« do odblokowania kreatywności, planowania strategicznego czy inteligencji emocjonalnej?

Zwiększanie niezawodności

Jak możemy tworzyć systemy, które weryfikują poprawność każdego kroku w łańcuchu myśli, aby zwiększyć niezawodność i zaufanie do wyników? (np. poprzez self-consistency lub zewnętrzne weryfikatory).

Wizja: Standardowe promptowanie pokazuje jedynie dolną granicę możliwości LLM. Chain of Thought udowadnia, że prawdziwy potencjał drzemie głębiej.

Od Paradoksu do Zasady: Podróż z Chain of Thought



Paradoks

LLM są genialne, ale jednocześnie zawodzą w prostych zadaniach logicznych.



Przełom

Chain of Thought:
Uczenie modeli **jak myśleć**, a nie tylko co odpowiadać.



Dowód

Wyniki State-of-the-Art i rygorystyczne badania dowodzą skuteczności przy odpowiedniej skali.



Zasada

Rozumowanie jest **emergentną, sekwencyjną właściwością**, którą można wywołać w języku naturalnym.

