

教育における生成AI活用推進リーダープログラム

生成AIの仕組み

テキスト生成の仕組み



吉田 壘

東京大学 大学院工学系研究科 准教授

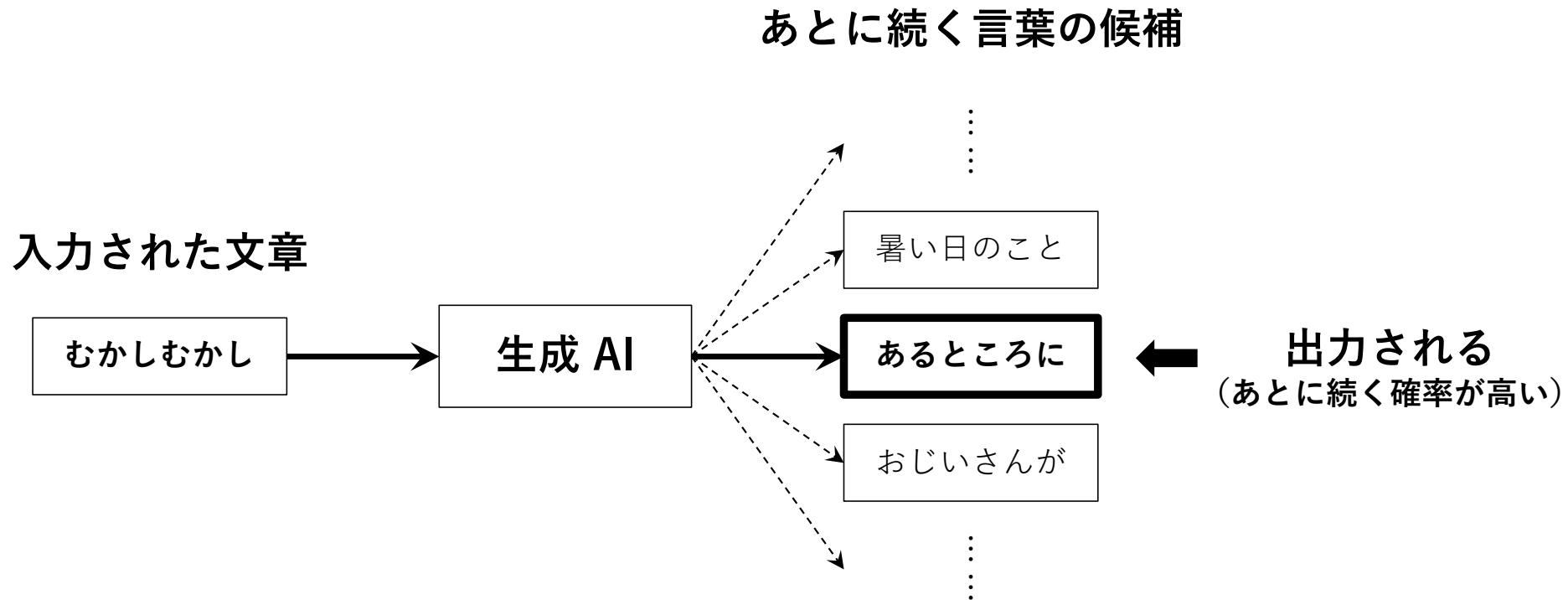
LLM 寄附講座 特任准教授

2026年4月

テキスト生成 AI の仕組み

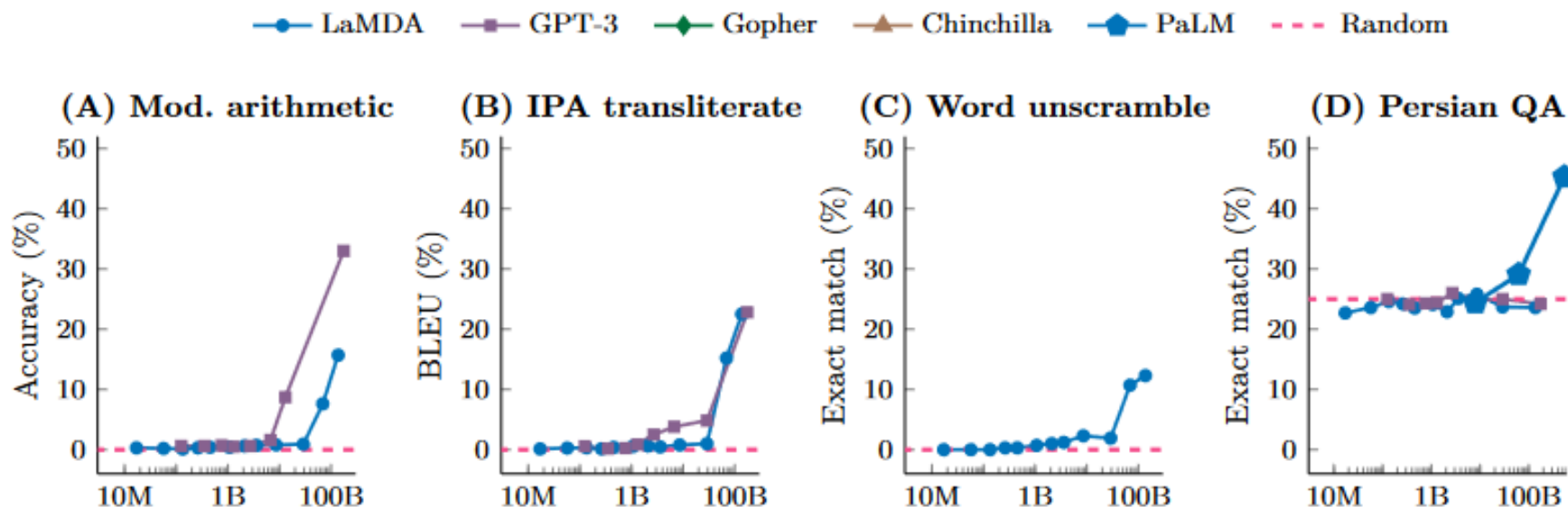
*推論モデルでは異なる(後述)

- 過去のテキストを参考に次に来る文字や単語を予測して出力する*
- 単純な機構と思いきや、モデルのサイズ（パラメータ数など）を大きくすることで、能力が高くなり、多様なタスクをこなすことが可能 (Brown et al. 2020, Wei et al. 2022a, OpenAI 2023)



スケール則と創発的能力

- スケール則: パラメータ数、学習計算量、データサイズを上げるほど性能が上がる (Kaplan et al. 2020)
- 創発的能力: あるパラメータ数 (学習計算量) を超えると急に性能が上がる (Wei et al. 2022a)



最近では小さなモデルでも性能を高める方法 (蒸留など) が開発されている

プロンプトの重要性 (プロンプトエンジニアリング)

- 例を載せると性能が向上 (Few-shot prompting) (Brown et al. 2020)
- 思考過程を出力させると性能が向上 (思考の連鎖: Chain of Thought) (Wei et al. 2022b)
- 例を載せなくても「Let's think step by step」とプロンプトに入れるだけで性能が向上 (Zero-shot prompting) (Kojima et al. 2022)

No.	Category	Template	Accuracy
1	instructive	Let's think step by step.	78.7
2		First, (*1)	77.3
3		Let's think about this logically.	74.5
		⋮	
14	irrelevant	By the way, I found a good restaurant nearby.	17.5
15		AbraKadabra!	15.5
16		It's a beautiful day.	13.1
-		(Zero-shot)	17.7

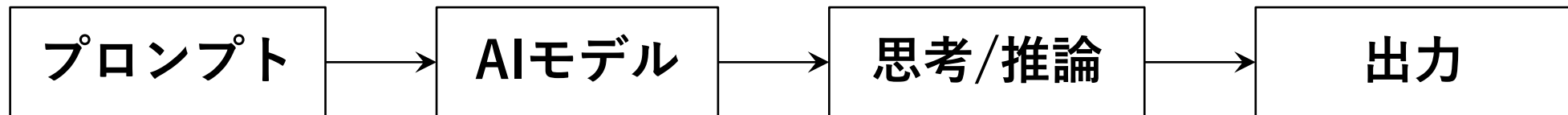
推論モデル

- モデル自身が思考の連鎖を行い、出力の質を高められるように学習させたもの
(雑に言えば自身でプロンプトエンジニアリングできるようになったもの)

従来のテキスト生成 (例: OpenAI GPT-4o, DeepSeek-V3)



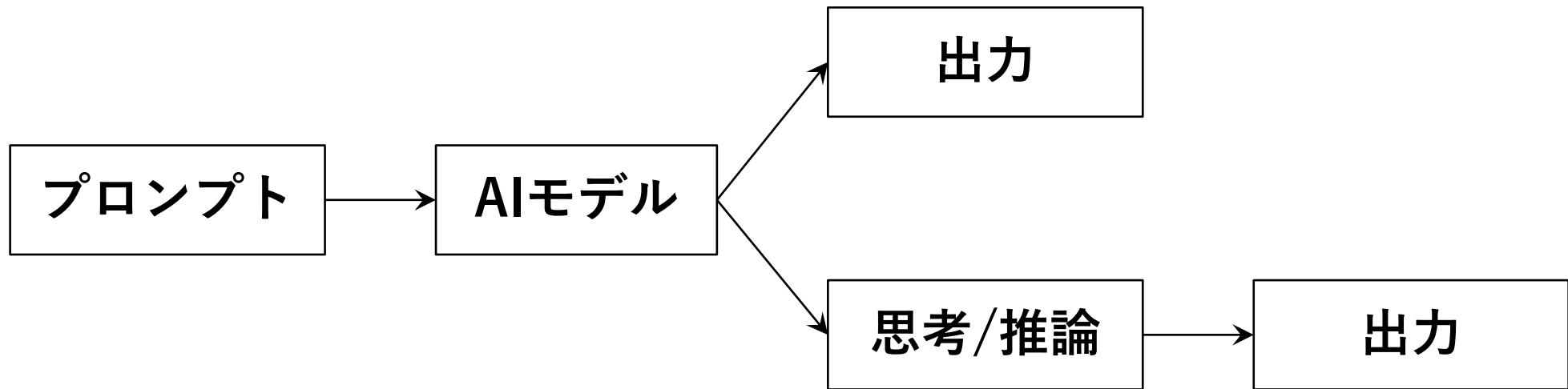
推論モデルのテキスト生成 (例: OpenAI GPT-5 Thinking, DeepSeek-R1)



この部分を更に
追加でトレーニング

ハイブリッドモデル

- 直接出力と推論を組み合わせることができるようになったもの
(例: Claude Opus 4.6 手動切替、ChatGPT 5.4 自動切り替え)



参考文献

- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901.
- Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., ... & Amodei, D. (2020). Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*.
- Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). Large language models are zero-shot reasoners. *Advances in neural information processing systems*, 35, 22199-22213.
- OpenAI (2023). GPT-4 technical report. *arXiv preprint arXiv:2303.08774*.
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., ... & Fedus, W. (2022a). Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022b). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35, 24824-24837.