

Three green apples are arranged on a white background. One apple is in the foreground, slightly to the right, and is in sharp focus. Two other apples are behind it, one to the left and one to the right, both slightly out of focus. The apples are bright green with some natural texture and a small stem at the top.

えひめ海事産業協働コンソーシアム
設置記念シンポジウム

スーパーコンピュータとAIで広がる
海事産業の可能性

二宮 崇

愛媛大学 大学院理工学研究科 情報工学講座

愛媛大学 データサイエンスセンター(兼任)

愛媛大学 工学部附属 社会基盤iセンシングセンター(兼任)

2025年9月26日

自己紹介

- 二宮 崇 (にのみや たかし)
- 愛媛大学 大学院理工学研究科 情報工学講座 教授
- 経歴:
 - 1986～1989 愛大附属中
 - 1989～1992 松山東
 - 1992～1996 東京大学 理学部 情報科学科
 - 1996～2001 東京大学 大学院理学系研究科 情報科学専攻
 - 2001～2006 JST研究員
 - 2006～2010 東京大学 情報基盤センター 講師
 - 2010～ 愛媛大学 大学院理工学研究科 准教授、教授



機械翻訳

● <https://www.deepl.com/>



テキストの翻訳
31言語

ファイルの翻訳
.pdf, .docx, .pptx

日本語 (自動検出) ▼

↔ 英語 (US) ▼

用語集

2枚のコインA, Bが与えられているとする。コインAは歪んでいないコインであり、表が出る確率は1/2であることがわかっている。しかし、コインBは歪んでおり、さらに、表がでる確率はわかっていない。これらのコインを直接観測することはできないが、2つのコインを同時に投げたときに表の枚数を観測することはでき、この数を X とする(すなわち、 X は 0, 1, 2 の値をとりえる確率変数)。2つのコインを同時に投げる試行を140回繰り返したところ、 X のそれぞれの値に対する出現回数は、次の表のような結果となった。

×

Suppose we are given two coins A and B. Coin A is an undistorted coin, and we know that the probability of it turning up is 1/2. However, coin B is distorted, and furthermore, the probability of it turning up is not known. We cannot directly observe these coins, but we can observe the number of **tables** that appear when two coins are tossed at the same time, and we denote this number by X (i.e., X is a random variable that can take on the values 0, 1, and 2). 140 repeated trials of tossing two coins at the same time yielded the number of occurrences for each value of X . The number of occurrences for each value of X is shown in the following table.

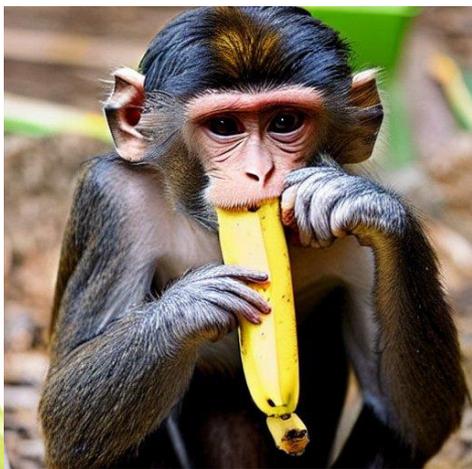
高精度機械翻訳サービスが利用可能に
Google翻訳、DeepLなど

画像生成

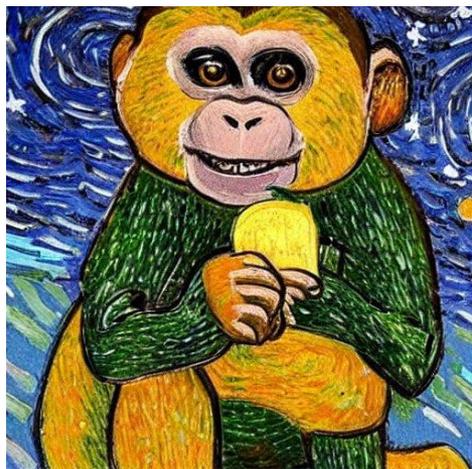
- Stable Diffusion

<https://stablediffusionweb.com/>

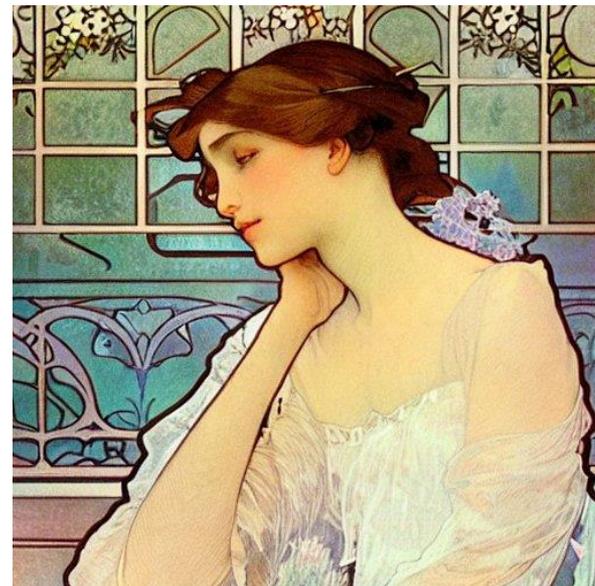
プロンプト
(prompt): a monkey
eating a banana



プロンプト: a
monkey eating a
banana in gogh
style



プロンプト: Illustration of a
beautiful girl in white dress
sitting by the window,
beautiful face, by Alphonse
Mucha, body tilted



ChatGPT

- 会話
- 質問応答
- 原稿作成
- 機械翻訳
- 自動要約
- プログラミング

TA 以下の日本語を英語に翻訳してください。「今日は愛媛県拡大部局長会議で講演をする」



"I will give a lecture at today's meeting of the expanded division directors in Ehime Prefecture."



TA クイックソートのプログラムをC++で書いてください。



以下は、C++でのクイックソートのプログラム例です:

```
cpp Copy code
#include <iostream>
using namespace std;

// 配列の要素を交換する関数
void swap(int* a, int* b) {
    int t = *a;
    *a = *b;
    *b = t;
}

// ピボットの適切な位置を見つけ、配列を分割する関数
int partition(int arr[], int low, int high) {
    int pivot = arr[high]; // ピボットとして配列の最後の要素を選択
```



GPT-4の驚異的な性能



Stanfordの大学院に入学できる

GRE (米国大学院の入学試験) やUBE (米国統一司法試験)などで高スコア



Jim Fan
@DrJimFan

I don't give a damn about what is or isn't AGI. It doesn't matter.

Below is GPT-4's performance on many standardized exams: BAR, LSAT, GRE, AP, etc.

The truth is, GPT-4 can apply to Stanford as a student now. AI's reasoning ability is OFF THE CHARTS. Exponential growth is the scariest thing, isn't it!

<https://twitter.com/DrJimFan/status/1635694095460102145>



Open AI (2023) GPT-4 Technical Report
<https://cdn.openai.com/papers/gpt-4.pdf>

Exam	GPT-4	GPT-4 (no vision)	GPT-3.5
Uniform Bar Exam (MBE+MEE+MPT)	298 / 400 (~90th)	298 / 400 (~90th)	213 / 400 (~10th)
LSAT	163 (~88th)	161 (~83rd)	149 (~40th)
SAT Evidence-Based Reading & Writing	710 / 800 (~93rd)	710 / 800 (~93rd)	670 / 800 (~87th)
SAT Math	700 / 800 (~89th)	690 / 800 (~89th)	590 / 800 (~70th)
Graduate Record Examination (GRE) Quantitative	163 / 170 (~80th)	157 / 170 (~62nd)	147 / 170 (~25th)
Graduate Record Examination (GRE) Verbal	169 / 170 (~99th)	165 / 170 (~96th)	154 / 170 (~63rd)
Graduate Record Examination (GRE) Writing	4 / 6 (~54th)	4 / 6 (~54th)	4 / 6 (~54th)
USABO Semifinal Exam 2020	87 / 150 (99th - 100th)	87 / 150 (99th - 100th)	43 / 150 (31st - 33rd)
USNCO Local Section Exam 2022	36 / 60	38 / 60	24 / 60
Medical Knowledge Self-Assessment Program	75 %	75 %	53 %
Codeforces Rating	392 (below 5th)	392 (below 5th)	260 (below 5th)
AP Art History	5 (86th - 100th)	5 (86th - 100th)	5 (86th - 100th)
AP Biology	5 (85th - 100th)	5 (85th - 100th)	4 (62nd - 85th)
AP Calculus BC	4 (43rd - 59th)	4 (43rd - 59th)	1 (0th - 7th)
AP Chemistry	4 (71st - 88th)	4 (71st - 88th)	2 (22nd - 46th)
AP English Language and Composition	2 (14th - 44th)	2 (14th - 44th)	2 (14th - 44th)
AP English Literature and Composition	2 (8th - 22nd)	2 (8th - 22nd)	2 (8th - 22nd)
AP Environmental Science	5 (91st - 100th)	5 (91st - 100th)	5 (91st - 100th)
AP Macroeconomics	5 (84th - 100th)	5 (84th - 100th)	2 (33rd - 48th)
AP Microeconomics	5 (82nd - 100th)	4 (60th - 82nd)	4 (60th - 82nd)
AP Physics 2	4 (66th - 84th)	4 (66th - 84th)	3 (30th - 66th)
AP Psychology	5 (83rd - 100th)	5 (83rd - 100th)	5 (83rd - 100th)
AP Statistics	5 (85th - 100th)	5 (85th - 100th)	3 (40th - 63rd)
AP US Government	5 (88th - 100th)	5 (88th - 100th)	4 (77th - 88th)
AP US History	5 (89th - 100th)	4 (74th - 89th)	4 (74th - 89th)
AP World History	4 (65th - 87th)	4 (65th - 87th)	4 (65th - 87th)
AMC 10 ³	30 / 150 (6th - 12th)	36 / 150 (10th - 19th)	36 / 150 (10th - 19th)
AMC 12 ³	60 / 150 (45th - 66th)	48 / 150 (19th - 40th)	30 / 150 (4th - 8th)
Introductory Sommelier (theory knowledge)	92 %	92 %	80 %
Certified Sommelier (theory knowledge)	86 %	86 %	58 %
Advanced Sommelier (theory knowledge)	77 %	77 %	46 %
Leetcode (easy)	31 / 41	31 / 41	12 / 41
Leetcode (medium)	21 / 80	21 / 80	8 / 80
Leetcode (hard)	3 / 45	3 / 45	0 / 45

ChatGPTは何者？

OpenAIが開発した対話型AI

2022年11月にChatGPTのサービスを開始。単純なチャットボットに比べ、人間と対話しているかのような優れた応答能力があり、質問を抽象的に理解した上で回答することができる

しかし、**技術論文が公開されていないので詳細はわからない**

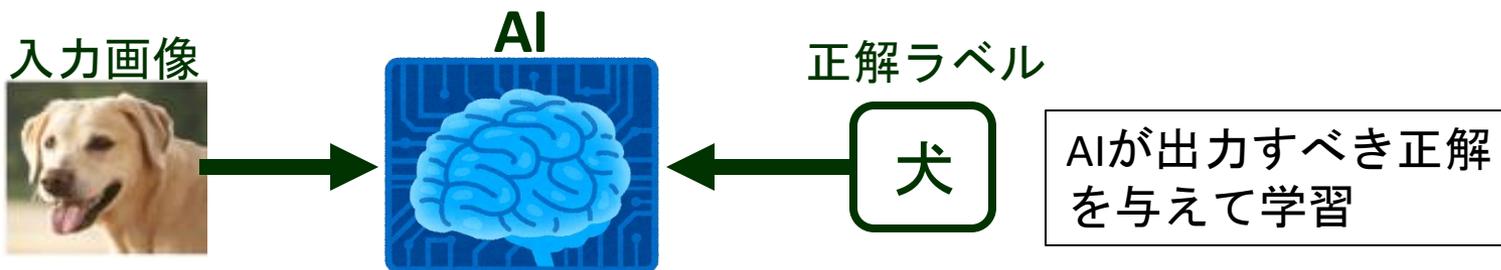
おそらく、

- **大規模言語モデル(LLM)**
- **インストラクション・チューニング**
- **人間のフィードバックに基づく強化学習(RLHF)**

で、できていると考えられている

AIの学習

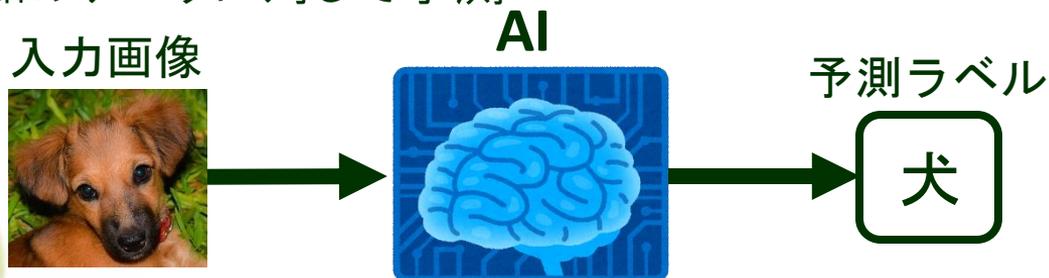
- 教師あり学習



大量のデータから「入力」と「出力」の結びつきを学習

- 予測

- 未知のデータに対して予測



AIの学習方法の変化: 事前学習

- 2018年～ 事前学習＋ファインチューニング
 - 事前学習(自己教師あり学習)
 - 解きたいタスク(機械翻訳など)を直接学習するのではなく、解くことに直接の意味はないタスクを学習
 - ファインチューニング
 - 事前学習した後に行う教師あり学習



AIの学習方法の変化: 事前学習

- 2018年～ 事前学習＋ファインチューニング
 - GPT
 - 次の単語を予測する問題を学習

例: 今日はえひめ海事産業協働

(正解: 「コンソーシアム」)

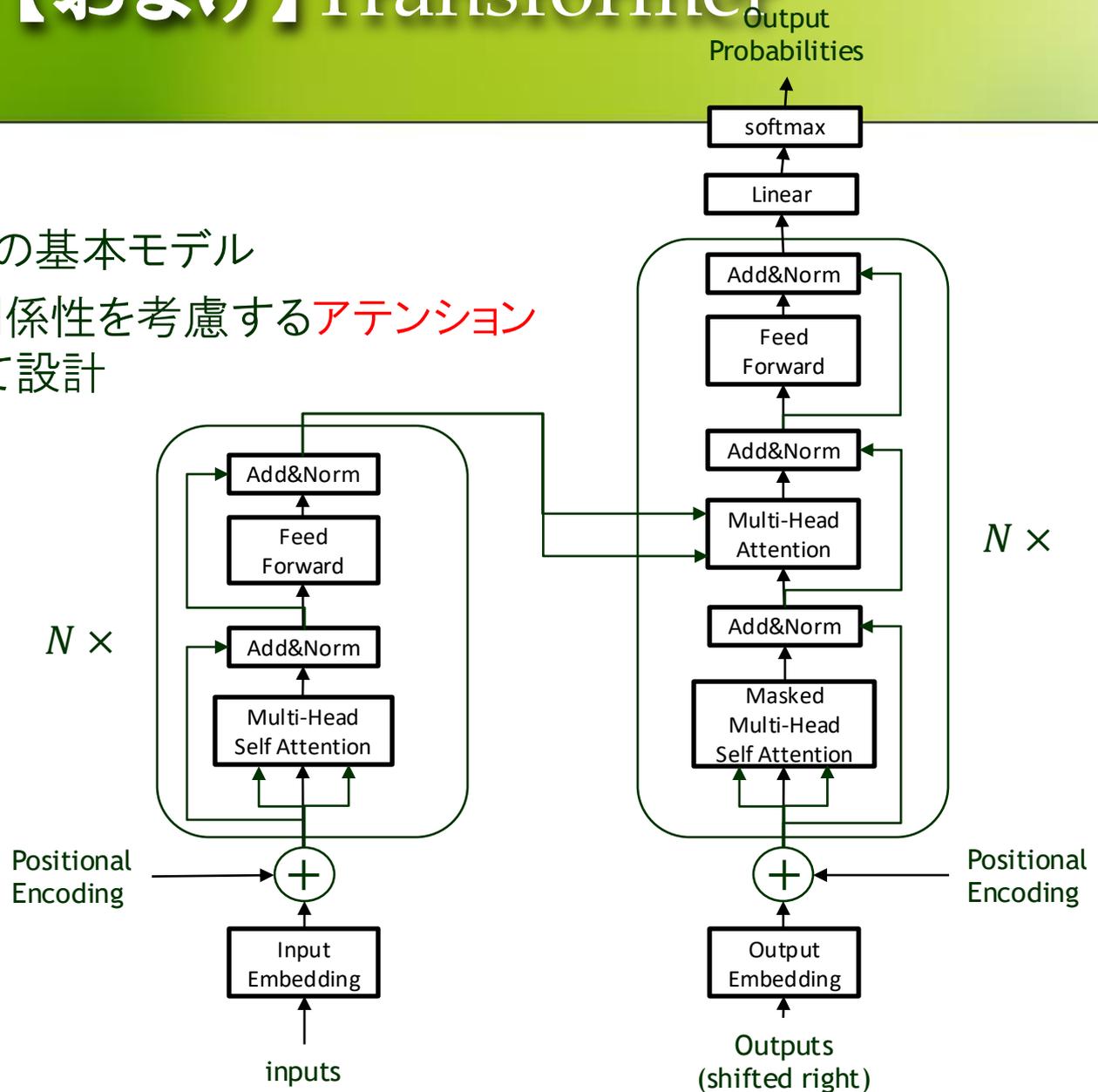
単語の穴埋め問題を解くこと自体に意味は無いが、
言語の構造や確率分布の学習が行われる



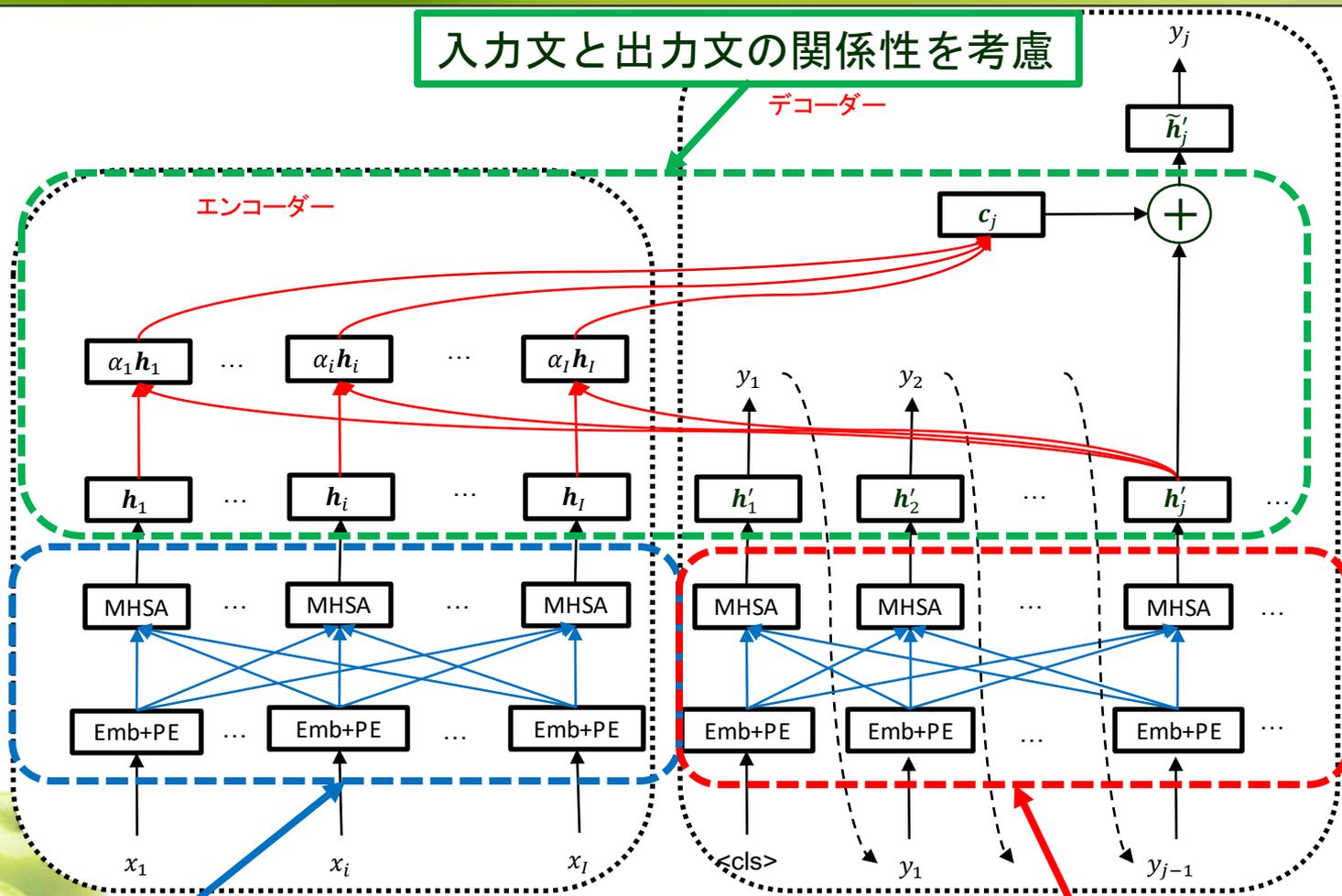
【おまけ】Transformer

● Transformer

- 現在のLLMの基本モデル
- 単語間の関係性を考慮するアテンションを基本として設計



【おまけ】Transformer

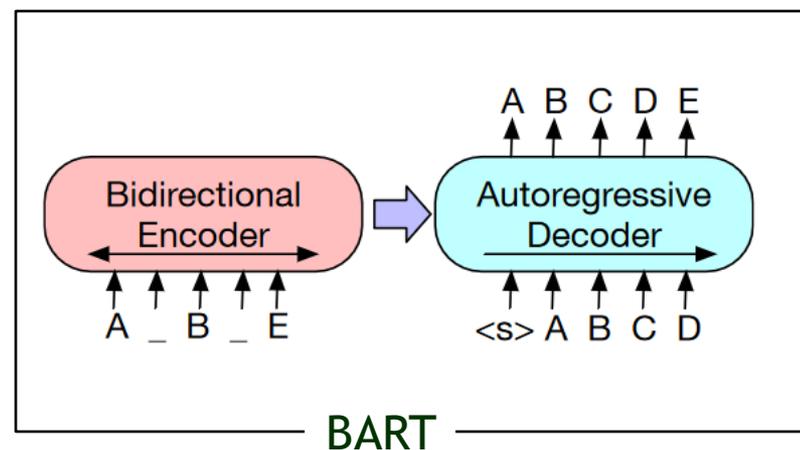
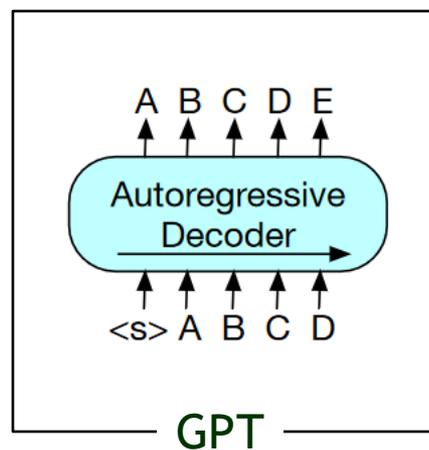
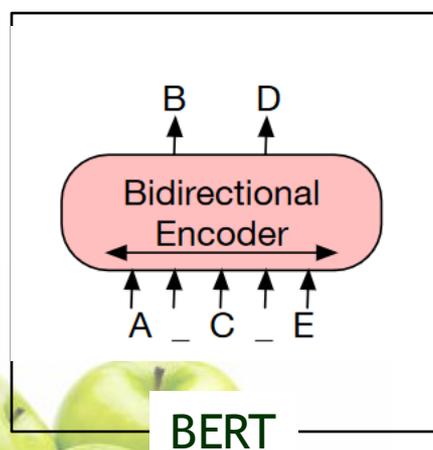


入力文の単語間関係性を考慮

出力文の単語間関係性を考慮

【おまけ】3種類の事前学習モデル

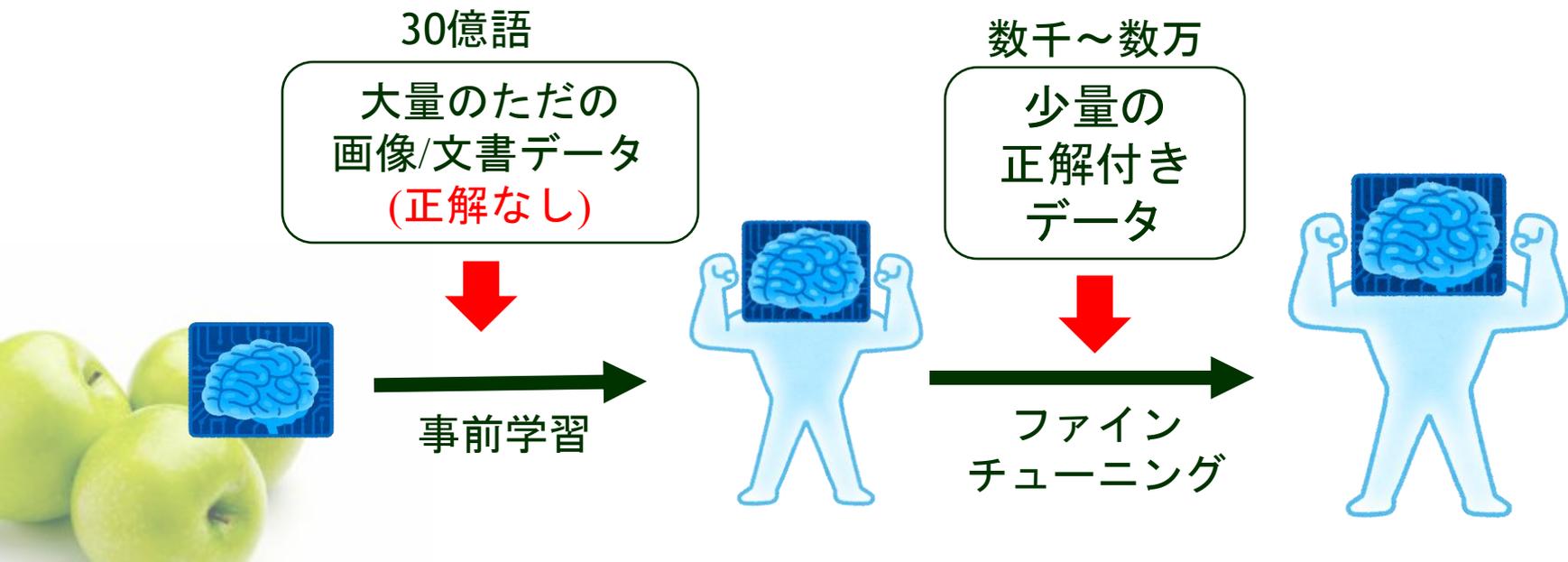
- **BERT**
 - Transformer(エンコーダー)による事前学習モデル
- **GPT**
 - Transformer(デコーダー)による事前学習モデル
- **BART**
 - Transformer(エンコーダー+デコーダー)による事前学習モデル



[M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer \(2020\) BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, Proc. ACL.](#)

AIの学習方法の変化: 事前学習

- **事前学習(自己教師あり学習)の何が嬉しいのか？**
 - 人手で正解がつけられたデータがなくても、**ただの画像やただの文書からいくらでも学習ができる！**
 - 大量のデータで事前学習をしておいて、**少量の正解付きデータで教師あり学習(ファインチューニング)**を行う
 - 様々なタスクでAIの性能が向上



AIの学習方法の変化: 大規模言語モデル (LLM)

さらに**大量のデータ**から**大規模AI(パラメータ数の大きいAI)**の事前学習をすればどうなるだろうか？

- 2020年～ 莫大な量のテキストで事前学習した巨大な言語モデル(**大規模言語モデル**) + ファインチューニング
 - 普通の言語モデル
 - BERT-LARGE(3.4億パラメータ, 33億語の文書データ)
 - GPT-2 (15億パラメータ, 40GBの訓練データ)
 - **大規模言語モデル**
 - **GPT-3 (1750億パラメータ, 570GBの訓練データ)**, GPT-3.5, InstructGPT, ChatGPT, GPT-4, Gemini, Claude
 - LLaMA, Qwen, Gemma



ChatGPT

- 大規模言語モデルGPTをベースとする対話AI
 - インストラクション・チューニング
 - たくさんのタスク(62個の自然言語処理データセット)でファインチューニング→様々なタスクを解くことができるAI
 - 自然言語推論、常識、感情分析、言い換え、質問応答、機械翻訳、読解、照応解析、自動要約
 - タスクの指示を自然言語(プロンプト)で与える→対話型で命令を与えられる
 - 人間のフィードバックによる強化学習(RLHF)
 - 人間に評価してもらって、人間が望む出力になるよう学習を調整
 - 不適切な応答を抑える



ChatGPTのできること/できないこと

◎テキストの読み書きの効率化

- メールや議事録やキャッチコピーなどの下書き
- 翻訳や要約や平易化などの一般的な自然言語処理
- プログラミング(下書き&デバッグ)

○対話

- 相談、ブレインストーミング

×新しい知識の収集

- 事実に反する発言をする(ハルシネーション)
- 非常に流暢なので、知らない内容の誤りに気づくのは難しい

×細かな制御

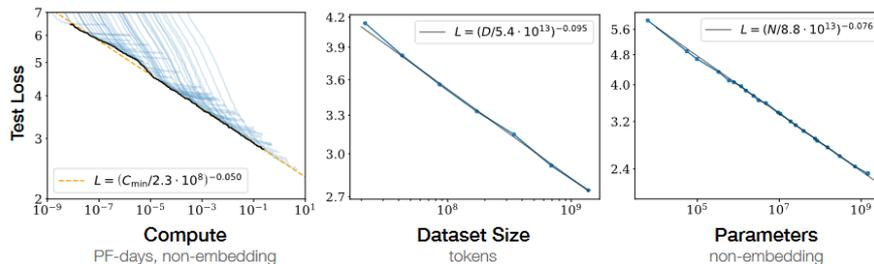
- 全ての指示に確実に従ってくれるわけではない(プロンプトのハッキングが必要?)

×算数

- 計算は苦手

様々な大規模言語モデル

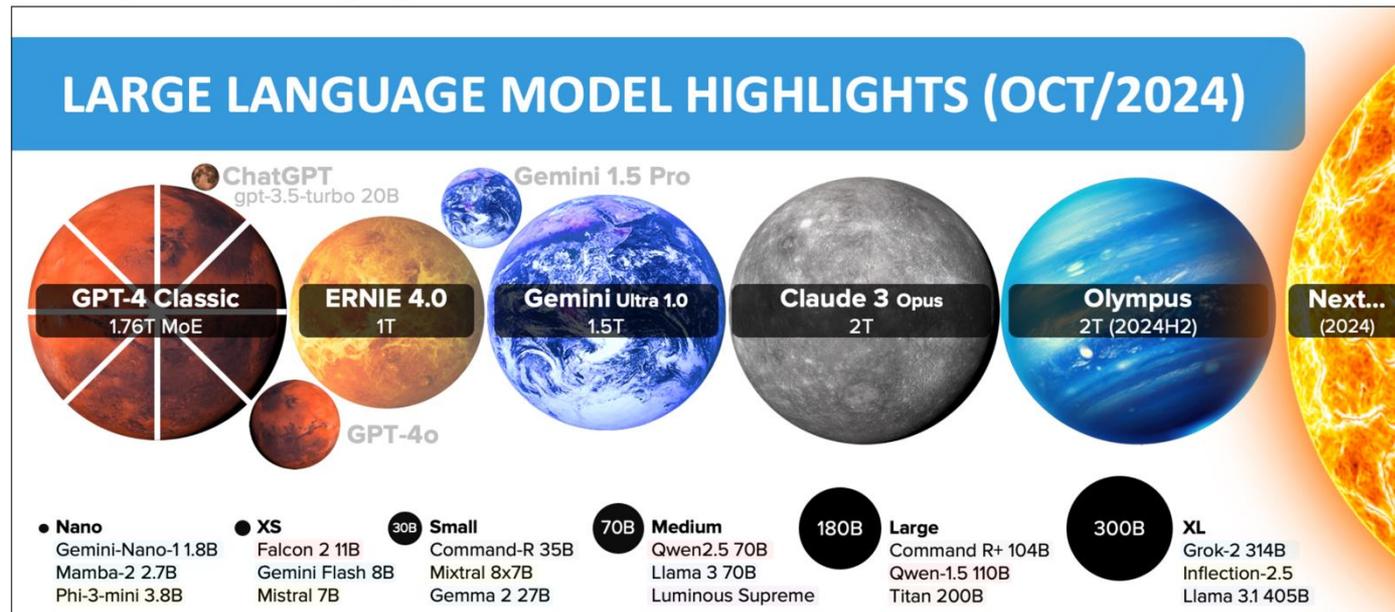
- 言語モデルの性能と計算能力、訓練データ量、パラメータ数の間にべき乗則



[J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, D. Amodei \(2020\) Scaling Laws for Neural Language Models, Arxiv.](#)

- 大規模言語モデル

2024 optimal LLM highlights <https://lifearchitect.ai/models-table/>



主要な大規模言語モデル

	開発元	パラメータ数	発表時期
GPT-3	OpenAI	1,750億	2020年6月
LaMDA	Google	1,370億	2021年5月
PaLM	Google	5,400億	2022年4月
OPT-175B	Meta	1,750億	2022年6月
LLaMA	Meta	650億	2023年2月
GPT-4	OpenAI	約1.8兆(?)の噂	2023年3月
Claude2	Anthropic	1,300億の噂	2023年7月
LLaMA2	Meta	70億, 130億, 700億	2023年7月
Qwen	Alibaba	70億	2023年8月
Mistral	Mistral AI	70億	2023年9月
Gemini1.0	Google	約1.6兆(?)の噂	2023年12月
Gemma	Google	20億, 70億	2024年2月
Claude3Opus	Anthropic	約2兆(?)の噂	2024年3月
Qwen2	Alibaba	720億	2024年6月
LLaMA3.1	Meta	80億, 700億, 4000億	2024年7月

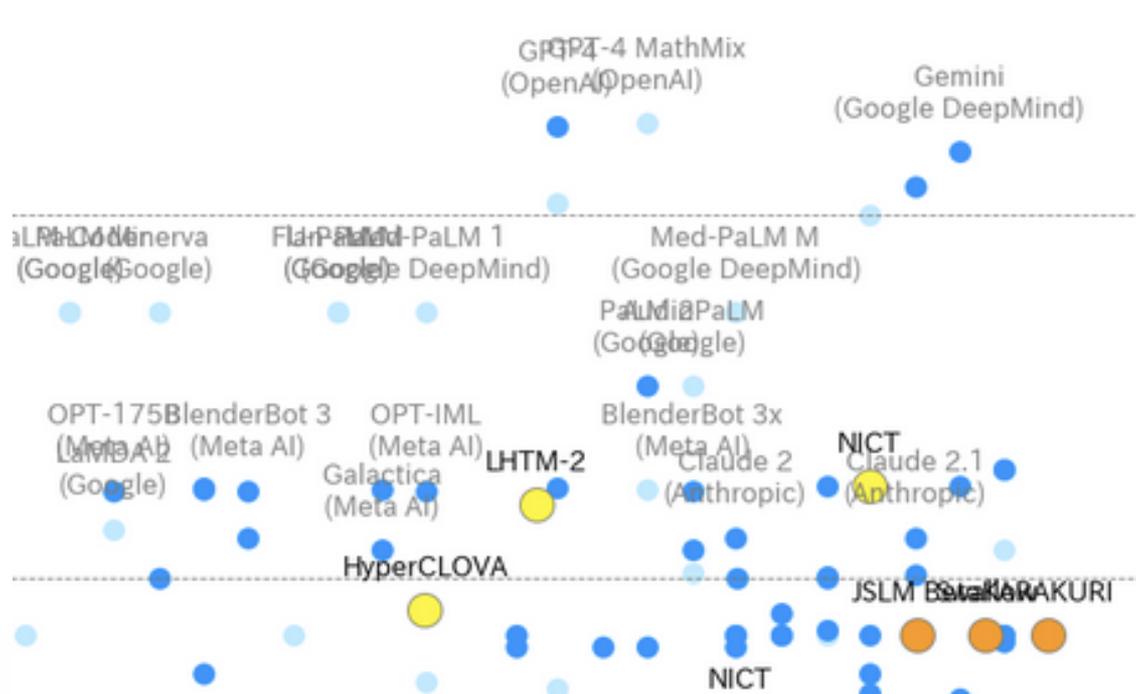


様々な大規模言語モデル: 日本の主要なモデル

	開発元	パラメータ数	発表時期
HyperCLOVA	LINE	820億	2022年1月
OpenCALM	CyberAgent	70億	2023年5月
GPT-NeoX	Rinna	36億	2023年5月
Japanese Stable LM Alpha	Stability.AI	70億	2023年8月
Weblab-10B	東大	100億	2023年8月
Japanese-large-lm	LINE	36億	2023年8月
LLM-jp	LLM-jp	130億	2023年10月
Swallow	東工大, 産総研	70億, 130億, 700億	2023年12月
Nekomata	rinna	70億, 140億	2023年12月
ELYZA-Japanese-Llama-2-7b	ELYZA	130億	2023年12月
Sarashina1, 2	SB Intuitions	70億, 130億, 650億	2024年6月
CyberAgentLM3	CyberAgent	225億	2024年7月
LLM-jp-3	LLM-jp	1720億	2024年9月
PlaMo-100B-Pretrained	Preferred Elements	1000億	2024年10月 ²⁰

日本語LLMのまとめサイト

- <https://llm-jp.github.io/awesome-japanese-llm/>



国内で大規模言語モデルを学習できる場所は少ない？

- GPT-3の学習に要するコストは460万ドル(約5億円)と言われている
- 計算機環境を購入すると、GPU (NVIDIA H100) 500万円×30,000枚=1500億円ぐらいかかる計算に...

Open LLM

- プロプライエタリLLMとOpen LLM

- プロプライエタリLLM

- ChatGPT, Gemini, Claudeなど
- APIでアクセスして利用。中身をいじることは難しい。中身の詳細が不明な(公表していない)場合が多い
- 大規模(～2兆パラメータ)で、非常に高い性能

- Open LLM

- LLaMA, Qwen, LLM-jpなど
- モデルとパラメータを公開(ダウンロード可能)
- ローカルマシンで推論、学習が可能！
- 改造することも可能
- 小規模 30億～700億(1300億や4000億もある)で、プロプライエタリLLMよりやや性能が低い(けど追いつきつつある)



Open LLM



● プロプライエタリLLM

- 高い性能で、API利用によって利用可能
- 商用サービスに展開、RAGで高性能化、実用化
- しかし、研究者にとってはもうすでに面白みのない研究対象となってしまう...

● Open LLM

- 大規模計算機で学習済みのモデルがダウンロード可能
- 創意工夫で小さいモデルながらも高い性能を目指している
- 既存のディープラーニング技術と組み合わせやすく、研究者にとっては非常に面白い研究対象



従来LLMの問題

- テキストだけから学習をしていて、現実世界との対応を理解できていない (シンボルグラウンディング問題)

freedom



theory



god



together



画像と単語の
シンボルグラウンディング

中国語の文書



中国語を全くわかっていないが、マニュアルに従って返事を書く



返事



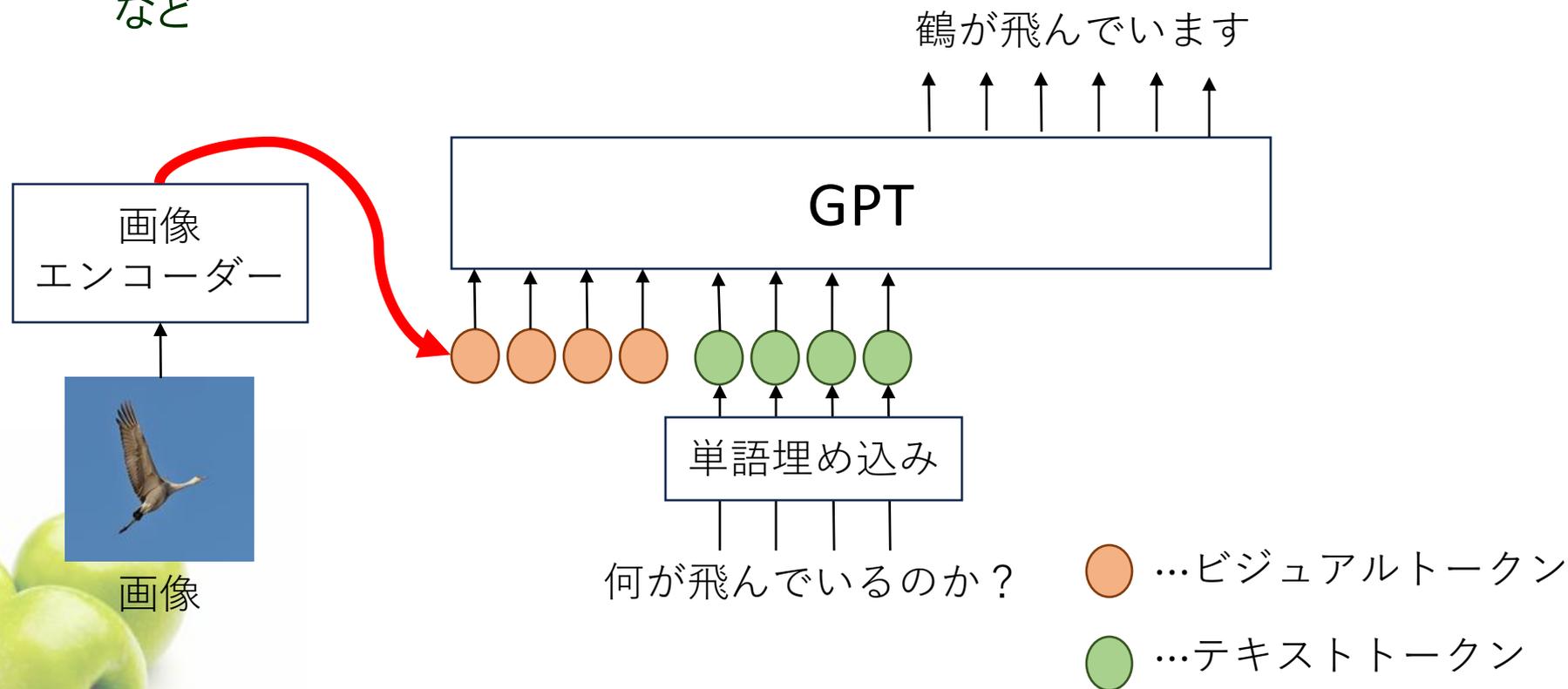
中国語の部屋問題 (サール 1980)



視覚言語モデル (Vision-Language Model; VLM)

- **視覚言語モデル(VLM)**

- 画像情報も扱えるLLM
- GPT-4o, Gemini 2.0 Flash, LLaMA 3.2, Qwen2.5-VL, DeepSeek-VL2 など



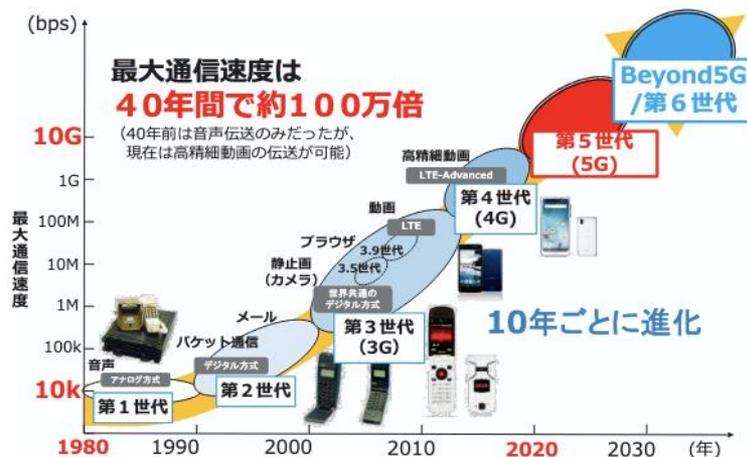
VLMの応用

- 画像キャプションニング（画像の説明文生成）
- ビジュアル質問応答（画像を使った質問応答）
- 画像データの理解（表やグラフの理解）
- OCR（文字認識）

「この画像の内容を説明して」



図表 1-1-2-1 移動通信システムの進化



(出典) 総務省作成資料

この図は「移動通信システムの進化」を示したものです。

主な内容

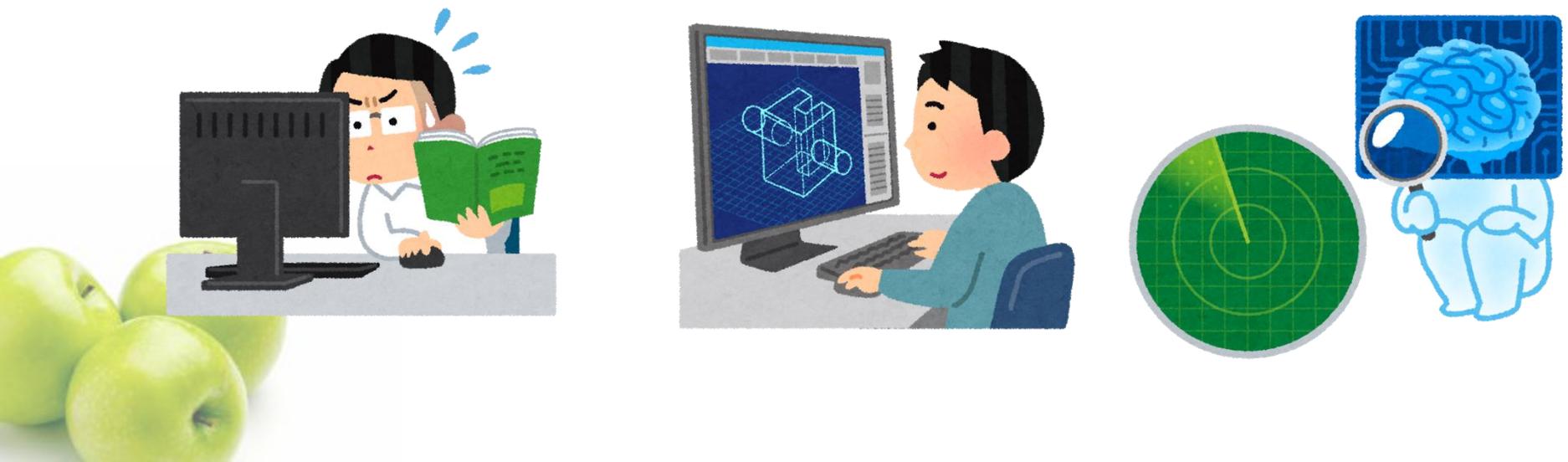
縦軸 (bps: 通信速度)

最大通信速度の推移を対数スケールで表しています。

1980年の第1世代（アナログ音声通信）から、2020年の第5世代（5G）まで、.....

VLMの海事産業への応用

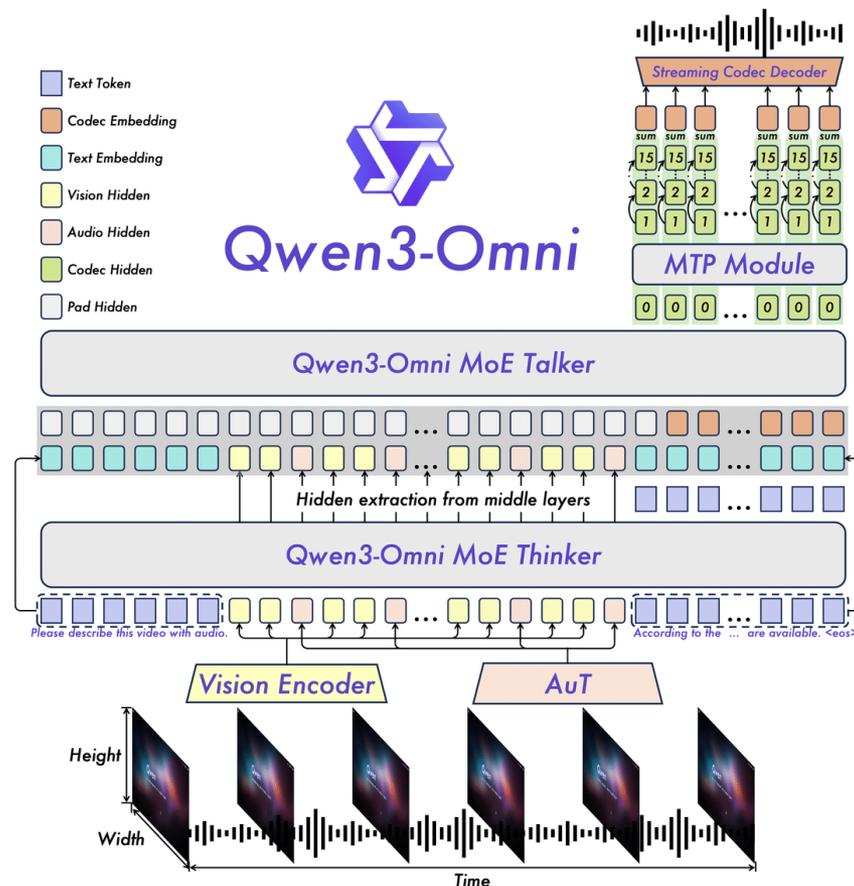
- 操船/造船に関するマニュアルに対するQAシステム
- 日本語の指示による3D CADの編集
- 気象情報、レーダー画像による海上交通の状況理解



オムニモーダル言語モデル

● オムニモーダル言語モデル (Omni-Modal Language Model)

- 音声、画像、動画、テキストの全てのモダリティを扱う大規模言語モデル
- フレームレートの違いの問題（音声・動画 vs 画像・テキスト）があったが、これを解決する次世代LLM
- GPT-4o, Gemini, Qwen3-Omni など



オムニモーダル言語モデルの応用

- **カメラ画像やレーダーを用いた海上自動監視**

- 港守り、船周辺の海上交通確認

- **機械の異常検知**

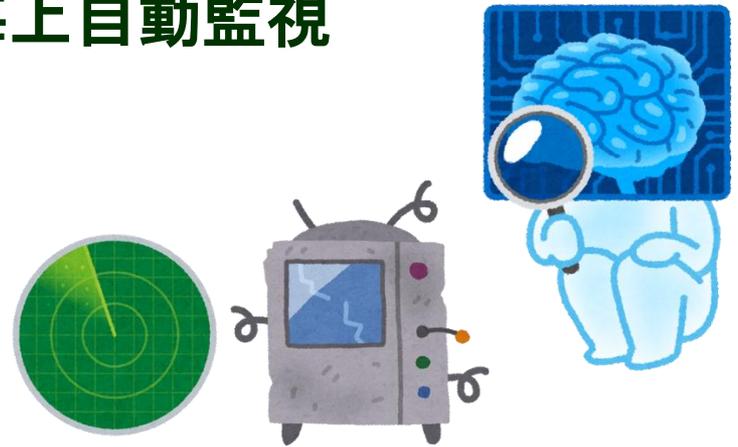
- 機械の異常を自動的に検知

- **操船サポート・自動操船**

- 衝突回避、接岸サポート
- 自動操船 (参考) Turingの自動運転、VLMやVLAの活用

<https://speakerdeck.com/tsubasashi/vision-language-model-toward-automated-ship-operation>

- **聞き取りづらい無線音声の自動音声認識**



LLMによって、考えながら、説明・分析を行うので、従来のような単純なミスが起きにくい



えひめ海事産業スパコンのススメ

- GPUサーバー 7ノード (約7000万円×7ノード = 4.9億円)

- 各ノード NVIDIA B200 GPU × 8枚
- 計56枚のB200 GPU

- (参考) 日本語LLM(Swallow)の学習コスト

- AWSのGPU256枚使用
- 16.25日 = 390時間
- AWSを用いた場合
 - 390時間×256枚×1024円=102,235,160円(約1億円)
- ABCIを用いた場合
 - 390時間×256枚×440円=43,929,600円(約4300万円)
- えひめ海事産業スパコンを用いた場合
 - 390時間×256/56 = 1782時間(=74.28日)の時間をかければ同等の学習ができる→**だいたい1年で元が取れる(?)**



まとめ

● 最近のAI技術

- 機械翻訳、画像生成、対話とChatGPT
- 超大量の生テキストデータからの事前学習(大規模言語モデル)
- Open LLM
- VLM
- オムニモーダル言語モデル
- 海事産業への応用
 - マニュアルのQAシステム
 - 日本語の指示による3D CAD編集
 - 海上交通の状況理解
 - 海上自動監視
 - 機械の異常検知
 - 操船サポート・自動操船
 - 無線音声の自動音声認識

