TSUBAME 共同利用　令和6年度　産業利用　成果報告書

利用課題名　LLM エージェントの研究開発
英文：Development of LLM agents

利用課題責任者
黒木　颯（So Kuroki）

所属
Sakana AI 株式会社

邦文抄録（300 字程度）

大規模言語モデル（LLM）に多様なエージェントスキルを獲得させるためには，複数タスクの学習データ比率や目的関数設定が難しいという課題がある．本研究では，品質多様性（QD）の一種である MAP-Elites に着目し，タスク性能を交互に最適化する CycleQD の提案に向けた予備実験を実施した．具体的な一連のプロセスとしては，各タスクに特化した専門モデルを初期に用意し，モデルパラメータを進化的にマージすることで，コード生成や OS 操作，DB 操作などで高い性能を獲得しつつ，一般言語能力の維持に成功した．さらに画像セグメンテーションモデルへの適用も示し，多様な領域への拡張可能性を明らかにした．

英文抄録（100 words 程度）

We propose CycleQD, a novel method for enabling large language models (LLMs) to acquire diverse agent skills. By treating each task's performance as a quality metric in a rotating fashion, our approach eliminates the need for intricate data mixing ratios. We employ model merging as a crossover step and introduce an SVD-based mutation to escape the convex region formed by the initial expert models. In empirical evaluations of coding, operating systems, and database tasks, CycleQD surpasses traditional fine-tuning baselines and reaches performance comparable to GPT-3.5-TURBO. Moreover, it retains strong language capabilities and extends easily to vision tasks.

Keywords: 5つ程度
Evolutionary Algorithm, Model Merging, Quality Diversity, LLM-based Agent, Skill Acquisition

背景と目的

大規模言語モデル（LLM）は，対話や推論において高い性能を示す一方，複数の専門的タスクを同時に学習させる際には，タスク間のデータ比率調整や目的関数の設計といった難題が生じることが知られている．さらに，単一タスク特化の複数モデルを単純にマージする場合には，過学習や性能の部分的劣化が見られることがある．本プロジェクトでは，品質多様性（Quality Diversity: QD）を用いた CycleQD という手法を提案し，これらの課題を解決するとともに，多様なタスクへの適用可能性を示した．

概要

提案手法である CycleQD は，MAP-Elites を基盤とした進化的アルゴリズムを用いてモデルマージをする手法である．まず各タスクに対して専門特化したエキスパートモデルを初期個体群として準備する．そのうえで，タスク性能である質（Quality）と行動特性（Behavior Characteristics: BC）を切り替えながら世代交代を行う．具体的には，(1) エリートサンプリングにより優秀なモデル個体を 2 つ取得し，(2) モデルマージにより 2 つのモデルを重み空間で線形結合し，(3) SVD を用いた突然変異を加えることで線形結合の範囲外へ探索を広げる．この過程で各タスクにおける最適化と多様なモデルの保持が同時に実現され，最終的に多数のタスクを高水準でこなせる統合モデルを得ることができる．本研究では，コード生成，OS 操作，DB 操作などのコンピュータサイエンス系タスクに加え，一般言語性能や画像セグメンテーションタスクでも本手法の有効性を示した．図 1 は手法の概要図である．

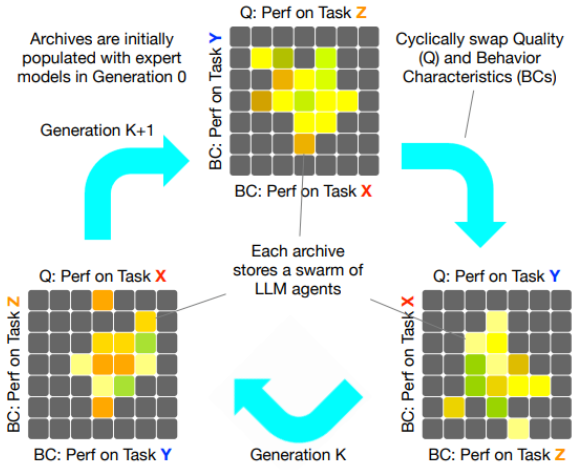

図 1. CycleQD では Quality と BC を周期的に切り替えながら世代交代を実施する.

結果および考察

実験では、LLAMA3-8B-Instruct をベースモデルとし、まずコード生成・OS 操作・DB 操作の三つのタスクに対して専門モデルを作成した。その後 CycleQD を適用することで、従来の単純なファインチューニングや従来型のモデルマージ手法を上回る性能を示す LLM モデルの獲得した. 特にコード生成と OS 操作タスクでは専門モデルよりも高いスコアを達成した。また、GPT-3.5-TURBO と比較しても同等程度の性能に迫る結果が得られ、言語能力に関しては大きな劣化が見られないことが確認できた。さらに画像セグメンテーションモデル（SAM）にも適用し、タスク間の類似性が高いほどモデルマージが容易であることなど、手法の有効性と限界を分析した. 表 1 に CycleQD と比較手法との比較を示した.

表 1. CycleQD は比較手法より高い性能を示した.

#	Methods	MBPP	DB	OS	Avg
0	GPT-4	83.6	36.5	63.7	61.3
1	GPT-3.5-TURBO	82.0	41.6	38.5	53.7
2	Llama3-8B-Instruct (base model)	67.3	5.3	25.2	32.6
Fine-tuning Based Methods					
3	Fine-tuning (Coding expert)	70.4	21.2	20.7	37.4
4	Fine-tuning (DB expert)	65.8	**42.4**	28.5	45.6
5	Fine-tuning (OS expert)	66.3	0.0	30.4	32.2
6	Fine-tuning (All)	67.3	37.1	36.7	47.0
Merging Based Methods					
7	Merging (w/o learning)	72.9	24.7	**42.6**	46.7
8	Merging (learning w/ GD)	69.3	41.2	29.6	46.7
9	Merging (learning w/ CMA-ES)	69.3	41.2	30.2	46.9
10	Merging (learning w/ NSGA-II)	75.9	**42.4**	36.4	51.6
11	CycleQD (Ours)	**76.4**	38.2	**42.6**	**52.4**

まとめ、今後の課題

本研究では、QD アルゴリズムを活用し、周期的に最適化対象タスクを切り替える CycleQD を提案するための予備実験を実施し、最終的に複数タスクを同時に高い水準で実行可能な LLM を構築できることを示した。従来の手法では困難だったデータ比率調整や目的関数設計を簡略化しつつ、多様で高性能なモデル群を得られる点が本手法の大きな利点である。一方で、学習に用いる専門モデル間の類似度が低い場合には性能向上が制限される可能性も示唆された。今後は専門モデルの事前学習工程における類似度制御や、CMA-ME など高度な進化的手法との統合による効率向上、さらにマルチエージェント環境への応用などが課題として挙げられる.