Institute of SCIENCE TOKYO

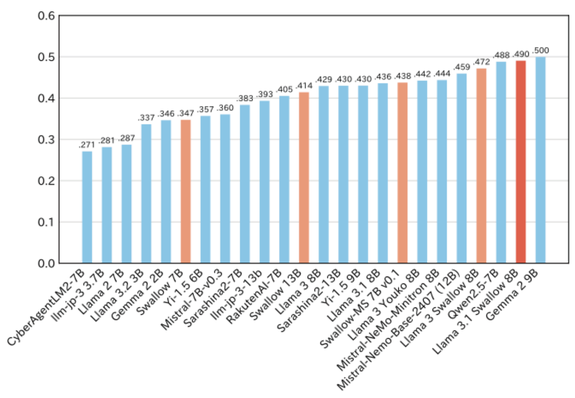
Supercomputing Research Center, IIR

Applications on TSUBAME4.0

Deep Learning and Large Scale CFD

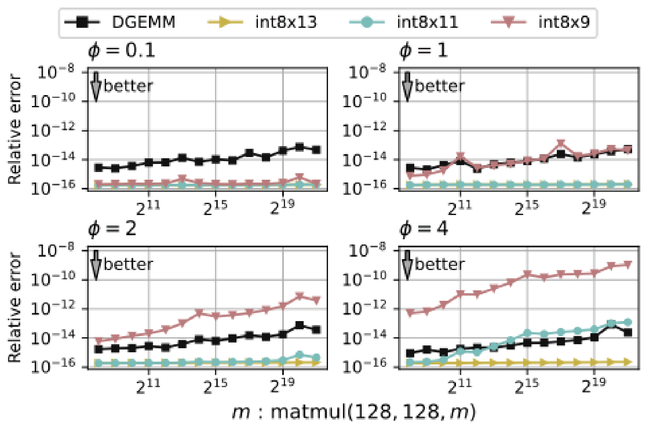
Large Language Models and AI Hardware for Science

Swallow: A Large Language Model Continually Pretrained on Japanese Data



Pretraining of large language models requires enormous computational resources. For enhancing the capability for a specific language, it is much more efficient to start from open models such as Llama, and continually train on the target language. We have been taking this approach since 2023 when we first continually pretrained from Llama2 7B, 13B, and 70B. We have also continually pretrained from Mistral 7B and Mixtral 8x7B. Most recently, we have released a model continually pretrained from Llama3 8B and 70B, and also Llama3.1 8B and 70B. Our models are available on huggingface: https://huggingface.co/tokyotech-llm

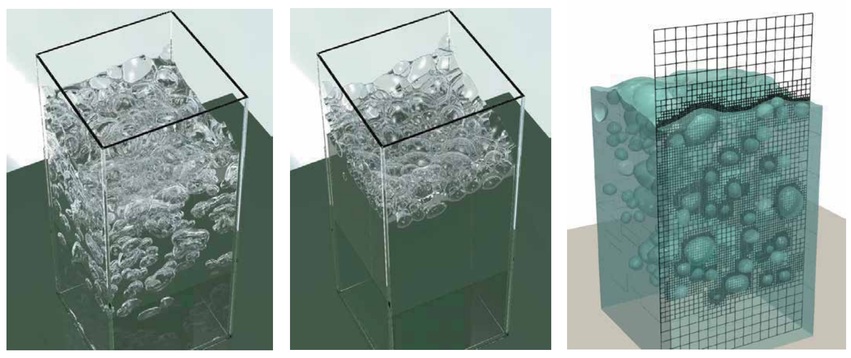
DGEMM emulation on Int8 Tensor Cores



It is of significant interest to find a way to harness the int8 Tensor Cores to improve the performance of HPC applications while maintaining accuracy. We focus on the Ozaki scheme, which computes a high-precision matrix multiplication by using lower-precision computing units, and show the advantages and disadvantages of using IMMU. The experiment using integer Tensor Cores shows that we can compute double-precision matrix multiplication faster than cuBLAS and an existing Ozaki scheme implementation on FP16 Tensor Cores on NVIDIA consumer GPUs. Furthermore, we demonstrate accelerating a quantum circuit simulation by up to 4.33 while maintaining the FP64 accuracy.

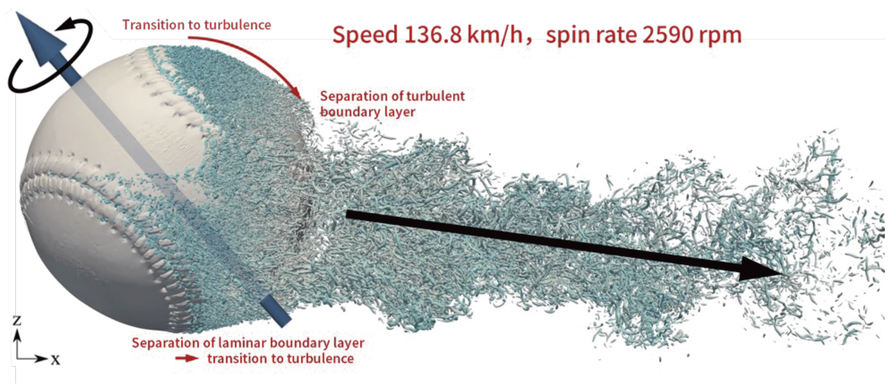
Large-scale Mesh-based and Particle-based Simulations

A Large-scale Foam Simulation using a Multi-phase Field LBM and AMR



Foam formation with a stable thin liquid film is very difficult to simulate using conventional methods due to the limitation of mesh resolution that can be used. We addressed this issue by using a Multi-phase Field Lattice Boltzmann Method and Adaptive Mesh Refinement. Using the Multi-Phase Field LBM, we can prevent a “numerical coalesce” phenomenon that leads to bubble break-up. Adaptive Mesh Refinement has been introduced so that the thin liquid film can be simulated efficiently. Herein, we demonstrate a simulation of foam formed from 200 air bubbles using our proposed method.

Unraveling Mystery of Sweeper



It is well known that MLB Los Angeles Angels Shohei Ohtani shows incredible performances for both batter and pitcher. He throws a kind of slider “Sweeper” (breaking ball) and its ball trajectory has more than 40cm horizontal movement but small movement in vertical direction. We study the aerodynamics of the sweeper by using a simulation based on the Lattice Boltzmann Method with a cumulant collision term suitable for a large-eddy simulation model. It is found that a lift force appears when the spin axis inclines to batter direction with 50-60 degree.

TSUBAME Grand Challenge

History of TSUBAME Grand Challenge

The TSUBAME Grand Challenge solicits proposals for grand challenge problems that can utilize all nodes of TSUBAME4.0 and has two categories.

Category A: Exclusive use of all nodes for 24 hours

Category B: Exclusive use of 1/3 of the nodes for up to 1 week.

Number of Accepted Proposals

	2024		2023		2022		2021		2020		2019		2018		2017		2016		2015		2014		2013		2012		2011		Total
	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	F	S	
Category A	0	1	0	0	0	0	0	0	0	0	0	0	0	1	2	0	1	1	1	2	1	2	0	1	2	2	3	4	24
Category B	0	2	0	1	0	1	0	2	0	1	1	2	0	2	0	1	0	1	1	3	2	2	1	1	0	0	2	0	26
Total	0	3	0	1	0	1	0	2	0	1	1	2	0	3	2	1	1	2	2	5	3	4	1	2	2	2	5	4	50

We started this program in 2011, and have continued to perform the Grand Challenge runs every year. Under this program, we have accepted a total of 50 grand challenge proposals, some of which were awarded the Gordon Bell prize as shown below.

Grand Challenge Projects for 2024

For 2024 we had 3 grand challenge projects as shown below:

Exploring Effective Continual Pre-training Methods for MoE Models
Takuya Akiba, Kazuki Fujii, Taishi Nakamura, Noriyuki Kojima, Jungo Kasai, Hiroto Kurita, Keisuke Sakaguchi, Naoki Okazaki, Masanari Ohii, Sakae Mizuki, Takumi Okamoto

Development of a detailed analysis method for the membrane permeation process of cyclic peptides based on multi-dimensional reaction coordinates
Yutaka Akiyama, Keisuke Yanagisawa, Masatake Sugita, Takuya Fujie, Kei Terakura

Large-Scale Vision-Language Models for Understanding and Generating Diagrams
Keisuke Sakaguchi, Itsumi Saito, Kazuki Fujii, Taishi Nakamura, Takumi Okamoto, Shigeki Ishida, Haruto Yoshida

https://www.isct.ac.jp