

発表概要

セル単位の実行状態分離を備えた ノートブックプログラミング環境

中丸 智貴^{1,a)} 佐藤 重幸^{2,b)}

2022年6月16日発表

JupyterLab や Google Colaboratory などのノートブック環境は、機械学習やデータ分析において広く利用されている。その大きな理由として、データの前処理、予測モデルの選択、モデルの訓練およびテスト、結果の可視化など、一連のタスク遂行における試行錯誤を対話的に行いやすい点があげられる。しかし、既存のノートブック環境で試行錯誤が安全かつ効率的に行いやすいわけではない。たとえば途中結果の再利用のためにコードセルをアドホックに再実行すると、コードセルの字面と実行結果が対応しなくなる。これによって再現性が失われたり、利用者が結果を誤って解釈したりする危険が生じる。この問題を解消するため、我々はセル単位の実行状態分離を備えたノートブックプログラミング環境を提案する。提案するノートブックプログラミング環境の特徴は、コードセルが木構造を成し、各セルが親セルの実行状態を引き継ぎながらも独立した実行状態を持つという点にある。これにより、提案システムの利用者は試行錯誤の選択過程を木構造として明示的に記述しながら、それぞれの試行を安全かつ効率的に実行して結果を比較できるようになる。本発表では、実行状態の分離をモジュラーかつ効率的に実装する手法を示し、その実験的評価を与える。

Presentation Abstract

A Notebook Programming Environment with Per-Cell State Isolation

TOMOKI NAKAMARU^{1,a)} SHIGEYUKI SATO^{2,b)}

Presented: June 16, 2022

Notebook environments, such as JupyterLab and Google Colaboratory, are widely used in data analysis and machine learning. It is mainly because these environments enable users to try out their ideas and get feedback on them quickly in an interactive way. Existing environments, however, are not sufficiently well-designed for trial and error in a safe and efficient way. For example, once re-executing some of the code cells for reusing intermediate results, the source code in the code cells no longer corresponds to the output. Such an error-prone design ruins the reproducibility and reliability of our trial-and-error outcomes. To remedy this problem, we develop a notebook programming environment that makes our trial and error safe and efficient. In our system, code cells form a tree, and each of them has its own execution environment isolated yet inheriting the execution environment of its parent. The entire notebook atop our system thus represents trial and error directly, where we can write and run each trial safely and efficiently and compare the outputs. In this presentation, we present a modular and efficient method for implementing the isolation and inheritance of the execution state of our system and experimentally evaluate our method.

This is the abstract of an unrefereed presentation, and it should not preclude subsequent publication.

¹ 東京大学大学院総合文化研究科
Graduate School of Arts and Sciences, The University of
Tokyo, Meguro, Tokyo 153-0041, Japan

² 東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology,
The University of Tokyo, Bunkyo, Tokyo 113-8656, Japan

^{a)} t-nakamaru@g.ecc.u-tokyo.ac.jp

^{b)} sato.shigeyuki@mi.u-tokyo.ac.jp