

〔報 告〕 計算科学・計算機科学人材育成のためのスーパーコンピュータ無償提供利用報告

東北大学大学院データサイエンスプログラムにおける プログラミングトレーニング演習

瀬川悦生

東北大学大学院情報科学研究科

1. 概要

2017 年 2 月 13 日から 15 日の 3 日間、東北大学情報科学研究科のデータサイエンスプログラムの講義、DSP Training camp I で大量の問題を解きながらプログラミング手法を身につける、または、ブラッシュアップする演習を行いました。

本プログラムに参加した学生は、情報科学研究科所属の学生のみならず、経済学研究科および生命科学研究科と多岐に渡り、様々な能力水準の学生に対して、並行して教育を行う必要がありました。そのため、同じ場所で同一の環境を整え、可能な限り不確定要素を排除するために、サーバーサイエンスセンターによって提供されるスーパーコンピュータ無償提供制度を利用させていただきました。

実際の実習では、端末室をお借りし、その場で受講者 9 名が提供していただいたスーパーコンピュータをはじめとする計算機を駆使することで問題を解決しました。同一の場所・環境で同一の課題に取り組むことで学生間における相乗的な教育効果が得られました。

以下では、その内容について簡単にご報告いたします。

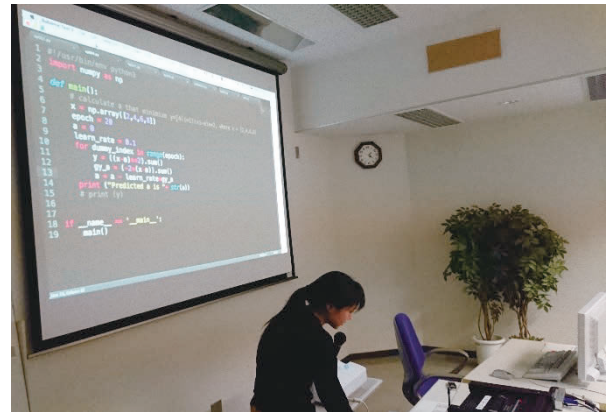
2. 取り組んだ問題

取り組んだ問題は Python プログラムを構築する 50 問です。数値と数値の足し算にはじまり、文字列の出力、繰り返し処理、条件分岐、配列の取り扱い、データ構造の理解、関数の取り扱い、行列の演算、ライブラリの使い方、乱数の取り扱い等の基本的な Python プログラミングの問題を経て、より実践的な文字列処理、簡単な機械学習法の問題や勾配降下法、進化戦略アルゴリズムの実装まで行ないました。

以上の問題はどれも基本的な実装の問題でありましたが、昨今のビッグデータの時代において、今後、大量のデータを取り扱うために必要とされる技術の一部です。ビッグデータに関する計算は、現段階においては、Linux 環境下においてキャラクターベースの操作ができることが前提ですが、この度、サイバーサイエンスセンターのスーパーコンピュータをお借りすることで、特にこの点を意識して技術を習得することができました。

3. 実習の様子および所感

参加した学生のプログラミング能力の水準が多岐に渡るため、実習中には不測の問題が発生することを憂慮していたものの、実際の実習が開始されてからは、プログラミング初心者に対して経験者が教えたり、プログラミング初心者の質問によって経験者に新たな発見があったりと、学生間において教育的な相乗効果が生じ、大きな問題なく実習は終了しました。これは、サイバーサイエンスセンターの端末室をお借りし、同一の場所に一同が会し、同一の計算環境を整えたために生じた良い効果であったように思います。また、利用させていただいた端末室には各デスクに教室前面に設置されているスクリーンと同様の内容を表示させることができるモニターが備え付けられており、これによって、講義部分を円滑に進めることができました。実習最終日に行ったソースコードの解説発表会でも、各々のモニターで詳細にコードを理解することができました。



学生によるコード解説発表会：勾配降下法の説明

4. 最後に

本実習は、サイバーサイエンスセンターのスーパーコンピュータを利用することではじめて実現することができました。また、実習の準備段階においては、Python プログラミング環境の構築をはじめとする技術的なサポートを頂きましたことを感謝申し上げます。