



## Desenvolvimento de uma biblioteca para Python utilizando multiprocessamento em CPU e GPU

Pedro Rodrigues Lima da Silva, João Luiz de Almeida Filho, Jorge Hernandez Fernandez  
LQFPP, CBB, UENF

A Bioinformática é um campo da ciência que emprega técnicas e ferramentas computacionais para o estudo de problemas e questões biológicas. Estas aplicações normalmente possuem um alto custo computacional e requerem tecnologia de ponta. Atualmente, a computação de alto desempenho utiliza-se de estratégias como a computação distribuída e a tecnologia GPGPU (*General-purpose computing on graphics processing units*). No entanto, as APIs para GPGPU foram desenvolvidas para linguagens de baixo nível, inibindo a utilização de este poder computacional em linguagens de alto nível, como por exemplo, o Python. Neste contexto, estamos desenvolvendo uma biblioteca Python que forneça o escalonamento do multiprocessamento utilizando CPUs e GPUs, permitindo o fácil desenvolvimento de softwares de bioinformática que necessitam de tecnologia de alto desempenho. Para avaliar esta abordagem estamos implementando os algoritmos clássicos de alinhamento de sequências Smith-Waterman e Needleman-Wunsch utilizando CPU e GPU, através as bibliotecas de *Multiprocessing*, *Biopy* e *Numpy* do Python para o processamento em CPU e PyCUDA em GPU. Após 6 meses de projeto, desenvolvemos um ambiente testes que consiste avaliar o tempo necessário para realizar o alinhamento de um *pool* de sequências de proteínas utilizando um ou mais núcleos de processamento em CPU. Para isto, utilizamos um arquivo fasta contendo 1372 sequências de proteínas, um computador Intel Xeon E5405 2Ghz com 8 núcleos de processamento e Linux. Nossos resultados demonstram que a utilização da biblioteca *multiprocessing* torna o alinhamento até 7 vezes mais rápido. Atualmente, estamos desenvolvendo os testes utilizando o PyCUDA, Biopy e Numpy para examinar o ganho de desempenho utilizando GPU. Utilizaremos a mesma máquina, porém, com o acréscimo de uma placa de vídeo NVIDIA GTX560Ti.

Palavras-chave: Bioinformática, Python, GPU, Paralelismo.

Instituição de fomento: FAPERJ, UENF.