

Área: ENGENHARIAS E CIENCIA DA COMPUTACAO

Projeto: EVOLUÇÃO GRAMATICAL MASSIVAMENTE PARALELA

Autores: IGOR LUCAS DE SOUZA RUSSO (XXII PIBIC/XXVI BIC/UFJF); HELIO JOSÉ CORRÊA BARBOSA; HEDER SOARES BERNARDINO (ORIENTADOR);

Resumo:

Regressão simbólica é um problema da área de identificação de sistemas que consiste em determinar um modelo simbólico a partir de um conjunto de dados, podendo auxiliar especialistas na descoberta de conhecimento. Neste domínio, técnicas de Programação Genética (PG) têm sido utilizadas com sucesso, alcançando resultados expressivos documentados em diversos artigos na literatura. A PG é uma metaheurística capaz de gerar programas automaticamente, por meio da evolução de uma população de indivíduos. Neste trabalho foi adotada a Evolução Gramatical, um tipo de PG que, por meio de uma gramática formal, evolui modelos em linguagem arbitrária.

Apesar de sua robustez, as técnicas de PG possuem alto custo computacional, sobretudo na etapa de avaliação dos modelos candidatos. Apesar disso, estas técnicas são consideradas naturalmente paralelas, pois há vários componentes que podem ser computados de forma independente, possibilitando a utilização de computação paralela.

O foco da pesquisa foi a utilização de GPUs para aceleração da técnica utilizando OpenCL (*Open Computing Language*), um *framework* para construção de sistemas paralelos em ambientes heterogêneos, que podem ser formados por CPUs (Unidades centrais de processamento) e GPUs, por exemplo. O modelo adotado aqui consiste em utilizar duas abordagens em conjunto: alguns indivíduos são avaliados em paralelo, sendo as amostras de treinamento computadas concomitantemente.

A avaliação dos modelos é realizada comparando-se a saída gerada pela execução de cada programa e aquela presente no conjunto de treinamento. Inicialmente, a avaliação foi realizada utilizando um interpretador e experimentos mostraram que o desempenho computacional desta metodologia em uma GPU foi até 66,3 vezes superior à execução sequencial do algoritmo em CPU.

Na Evolução Gramatical as soluções candidatas são programas gerados através de uma gramática formal e, assim, existe também a possibilidade de compilá-las a partir do código-fonte e executar a avaliação sobre os dados de treinamento. Uma estratégia de compilação foi proposta, na qual a cada iteração do algoritmo, somente uma compilação é realizada, englobando toda a população de indivíduos.

Experimentos mostraram que o desempenho computacional da estratégia de compilar os programas candidatos proposta é superior ao de interpretá-los para bases de dados com (i) grandes quantidades de registros ou (ii) alta dimensionalidade.