Scientific Programmer

Overarching Objectives:
To develop algorithms and software pipelines to implement statistical and computational methods for efficient and reproducible cancer research.

Principles:
Code produced by this group will be shared with the research community under a public license (typically GPLv3) and will typically be in the form of packages equipped with help files and vignettes. The candidate will identify and critically review existing hardware and software resources to assess their application as a solution to the problem at hand. Successful applicants will be expected to have experience with and to promote good programming practices, including using version control, conducting literate programming, performing unit testing and providing clear documentation. Requisite mentoring to write software papers, and will have the opportunity to contribute to scientific manuscripts. The position is housed in the Department of Biostatistics and Bioinformatics, and is funded by a grant from the National Cancer Institute.

Environment:
The group's programming environment primarily consists of the following four languages/environments: R, C/C++, Python and CUDA/OpenCL. The group is actively moving towards employing tools for distributed computing (e.g., Apache/Spark). The candidate will be expected to already have, or demonstrate the aptitude and enthusiasm to develop in a short time, the requisite expertise to effectively use these tools. The group uses the mercurial source code management system. Manuscripts and slides are typeset primarily using LaTeX. The group extensively uses the knitr extension package for producing reproducible reports. The use of other computing paradigms (e.g., functional programming) or languages (e.g., scala) are under constant review. The candidate needs to be able to conduct critical evaluations of alternate computing paradigms and languages and be able to effectively contribute to group decisions regarding their use and to ultimately employ these when appropriate.

Additional expectations:
While the emphasis of the position is on algorithm and pipeline development, along with programming, other duties include attending meetings and writing progress reports. The candidate will also manage the group's web presence, occasionally provide hardware (e.g., replace desktop or laptop hard drive) and software (e.g., compiling and installing software) support, and facilitate interactions with collaborators, investigators, and system and server administrators. The successful applicant will also contribute material for manuscripts, posters and presentations, including tables and figures (e.g., produced by R, tikz, graphviz).

Minimum Qualifications:
The candidate should hold a bachelor degree in computer science or equivalent quantitative/mathematical science (e.g., Mathematics, Statistics, Computational Biology, Engineering and Physics) with 1 – 3 years in scientific programming. The candidate should have had adequate training in basic mathematics, including calculus, linear algebra, discrete mathematics and numerical analysis (e.g., integration, optimization, linear algebra) along with basic training in statistics. The group utilizes the GNU/Linux operating system on its laptops, workstations and servers. The candidate should be comfortable working in this environment.

Interested candidates should apply at http://www.hr.duke.edu/jobs/ to requisition # 400992227.