Algorithm Research and Development Group (ARDG)

As proposed 11 November 2008

Background & Context

The NRAO is about to deploy two major new instruments - ALMA and the EVLA. The promised scientific capabilities of these instruments cannot be met by the data processing procedures and algorithms in any current software package (see, for example, EVLA memo #122: "EVLA Requirements for Post-processing Algorithms and Computing" by Frazer Owen).

Goals

The Algorithm R&D Development Group (ARDG) will provide a forum to coordinate algorithms needed for deployment of new instruments and for enhanced operation of existing instruments. The primary goal will be to produce internal memos and formal publications of new algorithms, not the writing of production software. The development of prototypes will be an essential step for testing by scientists throughout the observatory.

Coordinator

The Coordinator will be responsible to EVLA Operations and OEO. He will solicit priorities from management and from designated scientists in the major NRAO projects. He will be responsible for group budgeting, policies, work processes, and the search for external funding for these activities. He will identify the need for external assistance (e.g., consultants, students, etc.) to facilitate the development process. He will seek additional funding internally and externally is support of this effort. He will coordinate with other external radio astronomy facilities with similar needs.

The Coordinator will be responsible for producing periodic status reports to the various project management teams.

Project Scientist

The Project Scientist will be the primary liaison with the scientists specializing on the different instruments, not only within the NRAO but in the world community.

The project scientist will be responsible for arranging for scientific testers of code that the Developers have made available for testing. It must be recognized that the resources available for this are limited; most NRAO support scientists are fully subscribed for their service time.

Project scope

Since resources will initially be not much greater than the individual efforts already being expended on algorithms and associated efforts, the project will need to begin as a relatively small effort.

We will initially establish an organizational framework with the available resources. When additional resources become available, we will then be in a position to expand the scope of the effort.

Requirements and priorities

It will be the responsibility of the Coordinator working with and through the Project Scientist to assemble a list of problems without readily available solutions. These are available in many individual project documents, so this should be a matter of assembly. A list of algorithms to address these problems will be constructed. Periodic reviews of the requirements will be essential.

Similarly, the two will work with the support scientists and management of the individual instruments to define the relative priority of the required algorithms.

The Product

The products produced by this effort will be algorithms, documented by memos or formal publications, to the point where they are ready to be incorporated into publicly available software packages.

In order to centralize this information, we will create a web site which will contain or have links to requirements, work-in-progress memos, and completed memos and publications. To track these simply, we will create a new memo series. The website will also contain or have links to useful external relevant references.

Algorithm Development

It must be recognized that algorithm development is research. Although time estimates may be possible in some instances, it must be recognized that in others these may be difficult or impossible to determine a priori. Every attempt will be made to provide realistic estimates as appropriate.

Algorithm Types

In broad outline, there are two types of algorithms needed for data processing. The first is the development of ideas that are completely new; the second is the application of existing algorithms to modern computing environments (single and distributed clusters).

Computer Platforms

Algorithm developers have their own favorite development platform. It is essential that they be permitted to develop on the platform with which they are most comfortable.

Software Packages

It will be necessary to have the packages used for development available widely to the scientific testers within the NRAO. It will be necessary to provide the packages on the commonly used computer platforms (Linux and Mac OS/X) running the versions of the operating systems supported by the IT groups.

As mentioned, it is not required that the ARDG provide operational software. It is expected that several different packages may be needed to provide all of the operational software to reduce different observations on a given instrument.

Line Management

ARDG will function as a Science Working Group. The only difference is that key OEO personnel (Cotton, Fomalont, Hunt) will have this as their highest functional priority.

Scientific staff involved in other activities as a part of their functional duties will continue to report to their present line management (e.g., to NRAO-NM Ops for CASA software development).

Other scientific staff members will agree with their functional management concerning the amount of non-research time to dedicate to algorithms.

Software engineers will continue to report to their present management.

Developers

Developers are staff members who are willing to develop algorithms as agreed to by their Line Management. They will select appropriate development projects in consultation with the Coordinator and the Project Scientist.

They will be free to develop on platforms of their choice. Prototypes for review and testing will probably be a common intermediate product. Periodic (quarterly?) progress reports will be needed. Otherwise reporting will be via an internal memo or a journal publication.

Working FTEs

OEO will guarantee that approximately 1.5 FTEs will be available for the ARDG.

Algorithm work will also make up about 1 FTE of CASA development for the next development cycles.

The EVLA (through NRAO-NM Ops contributed effort) will provide approximately 0.2 FTE, although this may have to increase if we are to have EVLA staff scientists test new algorithms.

Involvement from other groups (e.g., AIPS) will be welcomed.

Group Philosophy

To encourage participation, we need to provide the necessary support and resources. By its very nature, this will be a heterogeneous group, using a heterogeneous set of tools. We want to create a collegial forum so that they feel part of a productive community.

We believe that we can demonstrate that we can produce techniques that are applicable to the needs of our operating instruments.

Group Meetings

This will need discussion after the group is founded. Several different forums will be needed:

- Developers actually developing doing new algorithms (monthly?)
- Scientific Coordination and testing (monthly?)
- Progress reviews (?) (quarterly?)

Oversight

We propose that an oversight committee be appointed. This will comprise the ARDG coordinator (Hunt) and project scientist (Owen) plus representatives from the computing groups of ALMA (Glendenning) and EVLA (Butler). We propose that meetings of this group will take place quarterly.