Report on the Analysis Tools Breakout Session

David Band and Seth Digel, Conveners
LAT Science Analysis Tools Workshop
June 14, 2002

The session was attended by a large number of scientists; unfortunately their names were not recorded.

Most of the session reviewed the analysis tools, and the anticipated difficulties in developing them. The following are the results of discussions about the four groups of tools that were identified.

1. Pulsars—we should be able to copy existing tools and therefore the development of these tools will not be a major research effort. Can these tools be applied to analyzing binaries?
2. GRBs—the tools should be applicable to both GBM & LAT data. The current concept is that the LAT data can be treated in the limit where the burst is a transient signal above the background of constant surrounding sources; this type of analysis is well-developed. These tools will be used for both temporal and spectral analysis, and will provide the user with the option to use a variety of statistical techniques.
3. Catalog comparison—the correlation of the positions of LAT sources with source catalogs in other wavebands must quantify the confidence of any possible identification. Catalog comparison is a mature field, and the GLAST point source catalog will not be a particularly large catalog by modern standards.
4. Likelihood tool—This will be the "Swiss Army Knife" of LAT analysis since the tool will be used for fitting source parameters, detecting sources, setting upper limits, comparing models, etc. Although the basic methodology is straightforward, the actual implementation will be difficult and will probably require various compromises so that the analysis will be computationally tractable. Many agreed that the likelihood tool should consider time dependence, both because GLAST will make such analysis routine, and also because time dependence can assist in the detection of weak sources. By adding to the LAT likelihood function the likelihood of other datasets (e.g., INTEGRAL), joint fits may be feasible.

In the discussion it was concluded that many tools should be split into smaller, simpler tools. In particular:

1. Verifying a candidate pulsar ephemeris and searching for the ephemeris of a candidate pulsar use different methodologies, and should be performed by different tools.
2. Barycentric corrections and phase assignments should be separated.
3. Whether the analysis of binary and pulsar periods is sufficiently different that separate tool sets should be created will have to be investigated.

The testing of the analysis tools was discussed briefly. There was a general consensus that the tools should be challenged with simulated data, particularly data created with the same simplified response functions used in the analysis. The tools

should also be applied to EGRET data. In addition to the informal tests performed while developing the tools, there should be ~2 formal data challenges held by the development team. It is anticipated that the first such data challenge will be catastrophic, while the software should perform nearly flawlessly for the last (second?) challenge. Finally, the testing should be documented.

The consensus was that the requirements need to be refined for the September review. In particular, the interfaces need to be defined clearly. Unfortunately, there was little detailed discussion of specifics in the current requirements documents.

Editors were assigned the task of revising the requirements and related use cases for groups of tools:

1. The likelihood tool—Pat Nolan
2. Pulsars—Masa Hirayama
3. Catalog comparison—Isabelle Grenier
4. GRBs—David Band

The group did not estimate the time required to develop each set of tools, although the consensus was that the likelihood tool will be the hardest to develop, and no schedules were established. A survey identified the estimated resources that each group planned to apply to the analysis tools development effort. (The table below does not differentiate between programmers, scientist/programmers, and scientists. Also, we may expect that some of the development effort included in the estimates will be devoted to utilities and databases.)

SU	1 FTE now, rising to 2-3 FTEs by 2004
SLAC	½ FTE now, rising to 2-3 FTEs by 2004
SSC	3-4 FTEs by September
CEA	½ FTE near term, 3+ from 2003
INFN	2-3 FTE
IN2P3	LLR maybe in future, Bordeaux net ~1 FTE
Bochum	1-2 FTE mostly on interstellar emission model
Japan GLAST Collaboration	Contribution to the standard analysis environment is unclear; their current focus is on background simulations, fast MC
Total	~8 FTE near term, ~17 in 2004 time frame

In many cases groups will work on tools that complement their research interests.