Science Databases: Contents & Access

- Disclaimer: *I don't know beans about databases*
- Definitions
- Enumeration
- Real-world Considerations

Definitions

- Science Databases

  All of the databases that have more than telemetry information in them (i.e., ≥ Level 1)

- Database

  Specified by the kind of data and the way it will be accessed (vs. “The GLAST Database”)

Enumeration of the Databases

- Event
- Photon
- Cosmic Ray
- Calibration
  - Response Function (HRM)
  - Low-level (JRB)
- Exposure History
- Point Source Detections
- Point Sources
- Transient Alerts/GRBs

Event Database

- Contents: Higher-level than Level 0
  - Has everything, including assignments of the hits and direction/energy estimation
  - Records are variable length, like EGRET 'Primary Database'
- Access:
  - Sequential, i.e., in time order
  - Could be useful to be able to 'point forward' to corresponding entries in the Photon & Cosmic Ray databases

Photon Database

- Contents:
  - Information about each event identified as a *photon* that might be useful in subsequent, higher-level processing
  - Fixed record size (a real-world issue)
  - Equivalent of the EGRET 'Summary Database' files
  - Points back to the Event Database
- Access:
  - Region of the sky, time range
  - Energy range, photon type (e.g., front vs. back conversion)

Cosmic Ray Database

- Contents:
  - Information about each event identified as a *cosmic ray* that might be useful in subsequent, higher-level processing
  - Fixed record size (a real-world issue)
  - Equivalent of the EGRET 'Summary Database' files
  - Points back to the Event Database
- Access:
  - Arrival direction in instrument coordinates, time range
  - Calorimeter crystal hit, Z, lack of nuclear interaction
  - ACD tile hit, Energy

Calibration Database

- Contents:
  - High-level response functions: PSF, A(eff), Energy resolution (HRM talk)
  - Also contains low-level information: pedestals, gains, alignments, channel masks, dead strips, thresholds,...
  - If the calibration files are updated, then need to store calibration history, so events can be processed with correct files
- Access:
  - High-level: Energy, angle(s)
  - Low-level: ? By layer, ladder, crystal, tile number
  - Both: Time or time range

Exposure History Database

- Contents:
  - Records for each ~1 minute time step and 2°×2° grid on the sky angular distance (and azimuth) from GLAST's pointing direction
  - Also has for each time step the mode of GLAST (on/off, etc.) and livetime
  - Generation of exposure maps becomes simple multiplication by A(eff) arrays
  - Tradeoff between database size and speed of generating exposure maps: ~40 Gbyte/year:
- Access:
  - Region of the sky, time range

Detection Database

- Contents:
  - Point source detections from each routine analysis of sky survey data (≥1 orbit)
  - 'Detections' means: position, flux, significance, perhaps hardness, and associated uncertainties
  - Each analysis of the sky is independent; we will be looking for flaring (new) sources more sensitively than inflight
- Access:
  - Region of the sky, time

Source Database

- Contents:
  - Essentially a summary of the Detection Database, derived by figuring out which detections are (probably) associated with others; not a static catalog
  - This is really how flaring sources will be recognized, routine flux histories constructed
  - Cross correlation with other astronomical catalogs (and other techniques) for IDs
- Access:
  - Region of the sky
  - (Questions about appropriate time intervals for Detection &Source Databases, whether to have ‘Upper Limits’ database)

Transient Alerts/GRBs Database

- Contents:
  - A subset of the Source Catalog with alerts that have been generated
  - Potentially has >1 entry per point source
- Access:
  - Time range, region of the sky

Real-World Considerations

- Reprocessing: (e.g., for improved reconstruction or modified background cuts)
  - Probably will require updating the Event Database (in place) then just regenerating the Photon and Cosmic Ray databases, because identifications and directions could change
- Access times:
  - Main databases will be big: 10^9 events/year, 10^8 photons/year
  - What waiting times are acceptable?
- How can the databases be shared ("mirrored")?