Chang'E-3 data pre-processing system based on scientific workflow

xu tan, jianjun liu, yuanyuan wang, wei yan, xiaoxia zhang, and chunlai li

National Astronomical Observatories, Beijing, China (tanx@nao.cas.cn)

The Chang'E-3(CE3) mission have obtained a huge amount of lunar scientific data. Data pre-processing is an important segment of CE3 ground research and application system. With a dramatic increase in the demand of data research and application, Chang'E-3 data pre-processing system(CEDPS) based on scientific workflow is proposed for the purpose of making scientists more flexible and productive by automating data-driven.

The system should allow the planning, conduct and control of the data processing procedure with the following possibilities:

• describe a data processing task, include:1)define input data/output data, 2)define the data relationship, 3)define the sequence of tasks,4)define the communication between tasks,5)define mathematical formula, 6)define the relationship between task and data.

• automatic processing of tasks.

Accordingly, Describing a task is the key point whether the system is flexible. We design a workflow designer which is a visual environment for capturing processes as workflows, the three-level model for the workflow designer is discussed:1) The data relationship is established through product tree.2)The process model is constructed based on directed acyclic graph(DAG). Especially, a set of process workflow constructs, including Sequence, Loop, Merge, Fork are compositional one with another.3)To reduce the modeling complexity of the mathematical formulas using DAG, semantic modeling based on MathML is approached. On top of that, we will present how processed the CE3 data with CEDPS.