Appears in Proc. Of International Workshop on Planning and Scheduling for Space, Darmstadt, Germany, June 2011. JPL CL# 11-1781

Scheduling Onboard Processing for the Proposed HyspIRI Mission

Steve Chien[1], David Mclaren[1], Gregg Rabideau[1], Daniel Mandl[2], Jerry Hengemihle[3]

[1] - Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 91109-8099, USA
[2] - Goddard Space Flight Center, Greenbelt, MD, 20771, USA
[3] – Microtel, LLC, Greenbelt, MD, 20771, USA

Abstract

The proposed Hyspiri mission is evaluating a X-band Direct Broadcast (DB) capability that would enable data to be delivered to ground stations virtually as it is acquired. However the HyspIRI VSWIR and TIR instruments will produce 1 Gbps data while the DB capability is 15 M bps for a ~60x oversubscription. In order to address this data volume mismatch a DB concept has been developed that determines which data to downlink based on both: 1. The type of surface the spacecraft is overflying and 2. Onboard processing of the data to detect events. For example when the spacecraft is overflying polar regions it might downlink a snow/ice product. Additionally the onboard software will search for thermal signatures indicative of a volcanic event or wild fire and downlink summary information (extent, spectra) when detected. The process of determining which products to generate when, based on request prioritization and onboard processing and downlink constraints is inherently a prioritized scheduling problem – we describe work to develop an automated solution to this problem.

Introduction

Future space missions will produce immense amounts of data. A single image from the HiRise camera on the Mars Reconaissance Orbitter (MRO) spacecraft is 16.4 Gigabits (uncompressed). The future HyspIRI mission under study is proposed to have two instruments - the HyspIRI thermal infrared imager (TIR) instrument producing 1.2 million pixels per second with 8 spectral bands at 4 and 7.5-12 microns per pixel and the HyspIRI visible shortwave infrared (VSWIR) producing 300 thousand pixels per second with 220 spectral bands per pixel in the 0.4-2.5 micron range. Keeping up with these data rates requires efficient algorithms, streamlined data flows and careful systems engineering. HyspIRI is also considering using Direct Broadcast technology to rapidly deliver this data to application users on the ground. However, in order to leverage the existing DB network, this downlink path is limited to approximately 15 million bits per second. The question is – which data to downlink when, in order to maximize the utility of the DB system?

We are studying the desired products and spectral bands required by volcanic, wildfire, flood and ocean/coastal, snow/ice, dust, and vegetation/ecosystem applications to assess onboard processing and band selection strategies for the mission. Three baselines for study are being investigated:

1. downlink of the spectral bands of the MODIS instrument (a commonly used rapid response instrument) over all target areas;
2. downlink of specially selected subsets of the bands based on overflight targets; and
3. onboard development of custom products based on overflight masks.

Volcanic applications include thermal detection and signature analysis as well as plume tracking applications. These volcanic techniques enable spatial subsampling to the areas of interest for dramatic downlink reduction. Onboard (EO-1) detection and ground-based (MODIS, AVHRR) detection algorithms are well understood. Wildfire applications include active fire mapping based on thermal signature (onboard EO-1, ground-based MODIS) as well as development of burned area products (significant heritage with Landsat ETM+, prior work with EO-1/ALI). A significant range of other applications with strong heritage in MODIS, AVHRR, GOES, and other rapid data delivery sensors exist in a range of disciplines. Ocean/coastal applications include products such as such as sea surface temperature and sea color applications such as harmful algal bloom tracking and Chl indices. Snow/Ice applications include trafficability and commerce route safety products as well as science cryosphere uses. Dust applications include aviation hazard assessment and environmental applications. Vegetation applications include plant stress, fire hazard, and disease vector applications based on measures of plan health and species identification.

Portions of this work were performed by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

These applications were derived from existing DB applications, discussions with the HyspIRI working groups, and others working in the relevant areas. Processing algorithm under consideration were assessed for adaptability and heritage from relevant prior sensors including MODIS, AVHRR, ASTER, Hyperion and others. These products are being refined with science and applications inputs and tested on current datasets and missions (such as EO-1) as well as being tested on relevant flight processing hardware and software configurations.

In the remainder of this paper we describe the operations concept being developed for the direct broadcast option with an emphasis on the scheduling problem for HyspIRI onboard product generation.

A Direct Broadcast Operations Concept for HyspIRI

The HyspIRI DB operations concept key drivers are:

1. Low or no sustaining operations costs
2. Low or no system development costs
3. Maximize utility of returned data
4. Graceful degradation/ high reliability of operations concept
5. Low risk, high heritage

With these drivers in mind, we have developed a highly automated operations flow for the DB component on HyspIRI consisting of the following steps. In order to reduce cost and risk we have used mature software systems.

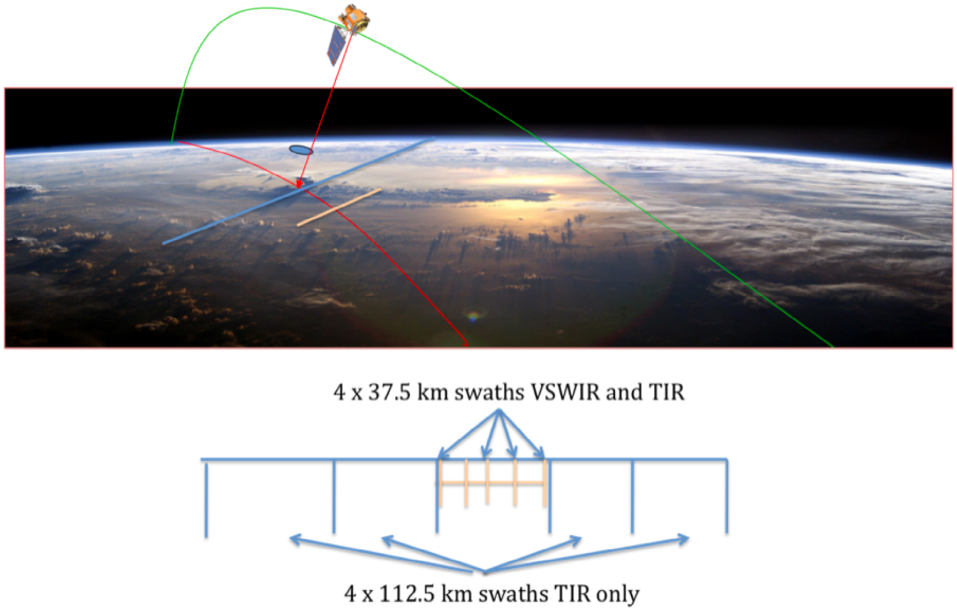

Figure 1: Projected Instrument Swaths

1. Specification of geographical regions of interest (ROIs) by the DB applications team. In this step, the applications team has a set of geographical regions (in essence polygons on a map of the Earth). For each polygon there is an algorithm and a priority. These polygons may also be seasonal (e.g. January to March of each year) and may be derived based on external information (e.g. reports of flooding, or rainfall, or the National Interagency Fire Center (NIFC) fire reports.
2. The spacecraft operations team provides the current best projection of the orbit of the spacecraft (e.g. a ground track file).
3. The spacecraft orbit is combined with the ROIs using knowledge of the spacecraft instrument swaths (150km wide for VSWIR and 600km wide for TIR) using the CLASP coverage planner [Knight 2009]. This produces a timeline of overflights for each of the 8 instrument swaths. CLASP is then used to determine the top priority products/spectral bands to process for each timestep, respecting the product priorities as specified by the applications team. CLASP produces an activity plan/sequence for the onboard processing module.

Thus the spacecraft orbit determines the type of terrain that will be overflown (e.g. land, ice, coastal, ocean, etc.). The TIR instrument has a 600km swath under the spacecraft and the VSWIR a 150km wide swath. In order to satisfy the high data rate from the instruments there are four interfaces from the instrument to the onboard processing. The VSWIR data is divided into four across track swaths of 37.5 km each. The TIR data includes four swaths matching the VSWIR swaths with the remaining 450km with of TIR only data divided into another four data paths (see Figure 1). Therefore each of the interface paths receives one 37.5km swath of VSWIR and TIR data and one 112.5k swath of TIR only data.

Each of the terrain masks implies a set of requested modes and priorities and is evaluated based on the eight swaths from the instruments. For example, when overflying polar or mountainous regions, producing snow and ice coverage maps can provide valuable science data. Additionally, the science team can adjust these priorities based on additional information (e.g. external information that a volcano is active, knowledge of a flooded area, an active wildfire, or a harmful algal bloom). The mission planning tool accepts all of these requests and priorities, and determines which onboard processing algorithms will be

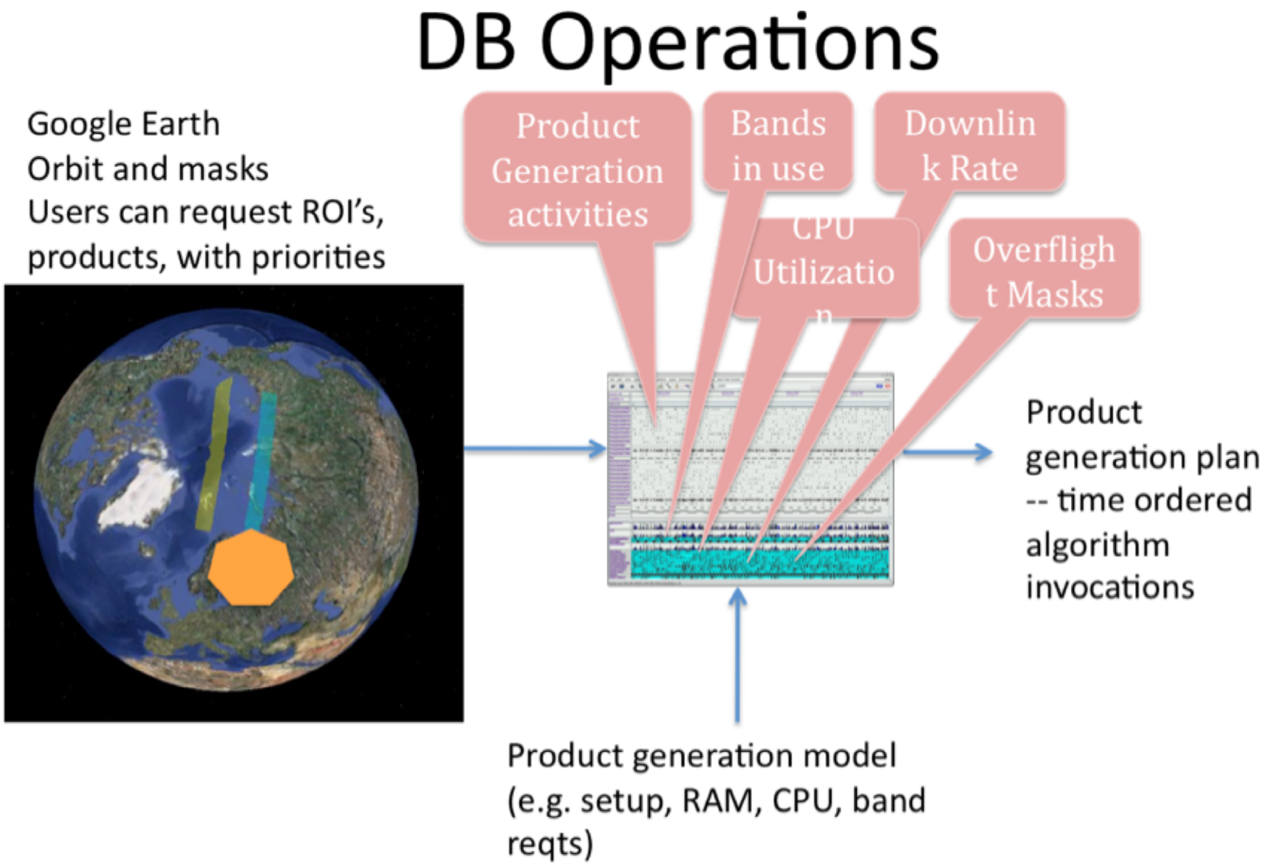

Figure 2: Direct Broadcast Operations Flow

active by selecting the highest priority requests that will fit within the onboard processing CPU resources, band processing limitations, and downlink bandwidth.

The automated planning model tracks the limited spacecraft resources that in this case include: # of bands processed, onboard CPU (each algorithms places a different load on the CPU), and downlink bandwidth. These operations constraints represent the onboard restrictions that: 1) only a limited number of bands of the instrument data can be processed onboard (for example, on EO-1 we can only process 12 of the bands per image), 2) that we have limited CPU processing capability onboard and this may limit the products we can generate at any one time, and 3) that the downlink transmission rate is limited to 15 Megabits per second. Accounting for these operations constraints, the mission planning system chooses the highest priority products that can be produced. Figure 3 shows the instrument processing swaths and Figure 4 shows a sample mission plan generated based on CPU and downlink resources.

Onboard processing algorithms can use a wide range of techniques. Past algorithms have consisted of: expert derived decision tree classifiers, machine learned classifiers such as Support Vector Machines (SVM) classifiers and regressions, classification and regression trees (CART), Bayesian maximum likelihood classifiers, spectral angle mappers, and direct implementations of spectral band indices and science products.

For example, SVM's have been applied to learn to classify EO-1 Hyperion images into Snow, Water, Ice, Land, and Cloud pixels [Castano et al. 2005].

CART techniques [Breiman et al. 1984] have been applied to a wide range of classification problems including remote sensing [Castano et al. 2006]. CART techniques recursively split the decision classification or estimation problem until a stopping criterion of goodness of fit is met. Maximum likelihood classifiers have also been applied to classification of remote sensing imagery (e.g., [Goodenough et al. 2003]). Given a presumed parametric probability distribution, these techniques find the parameters that maximize the likelihood of the observed training set. Spectral angle mapping (SAM) is an instance-based classification technique. If one considers each pixel as an n dimensional vector if the remote sensing imagery has n spectral measurements, SAM selects as the matched class the one whose prototype spectral vector is closest in angular distance to the vector of the pixel in question.

In other cases the science disciplines have already developed science products (e.g. measures) to track a physical phenomenon. In oceanography, Flourescent Line Height and Maximum Chlorophyll Index are indicative of biological activity such as algal bloom. In vegetation and ecosystem monitoring Normalized Difference Vegetation Index and Photochemical Reflectance Index (PRI) are indicative of plant health.

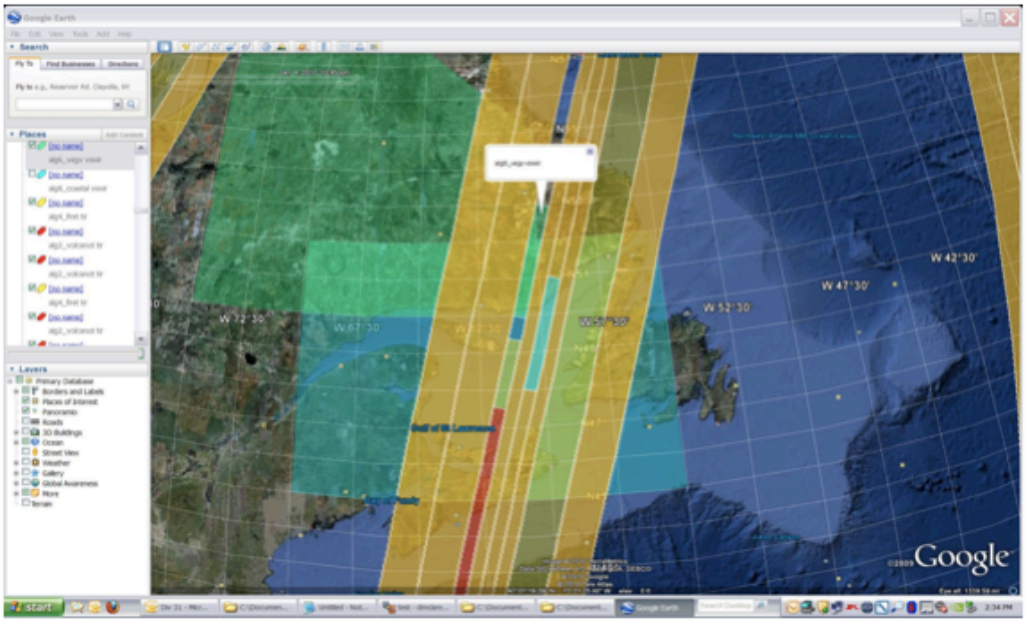

Figure 3: Instrument swaths overlayed against science regions of interest.

Spacecraft Operations Modeling

In this section we describe the range of spacecraft operations constraints present in the HyspIRI onboard processing scheduling problem. We begin by describing constraints that are easily modeled in automated planning/scheduling systems and then discuss more

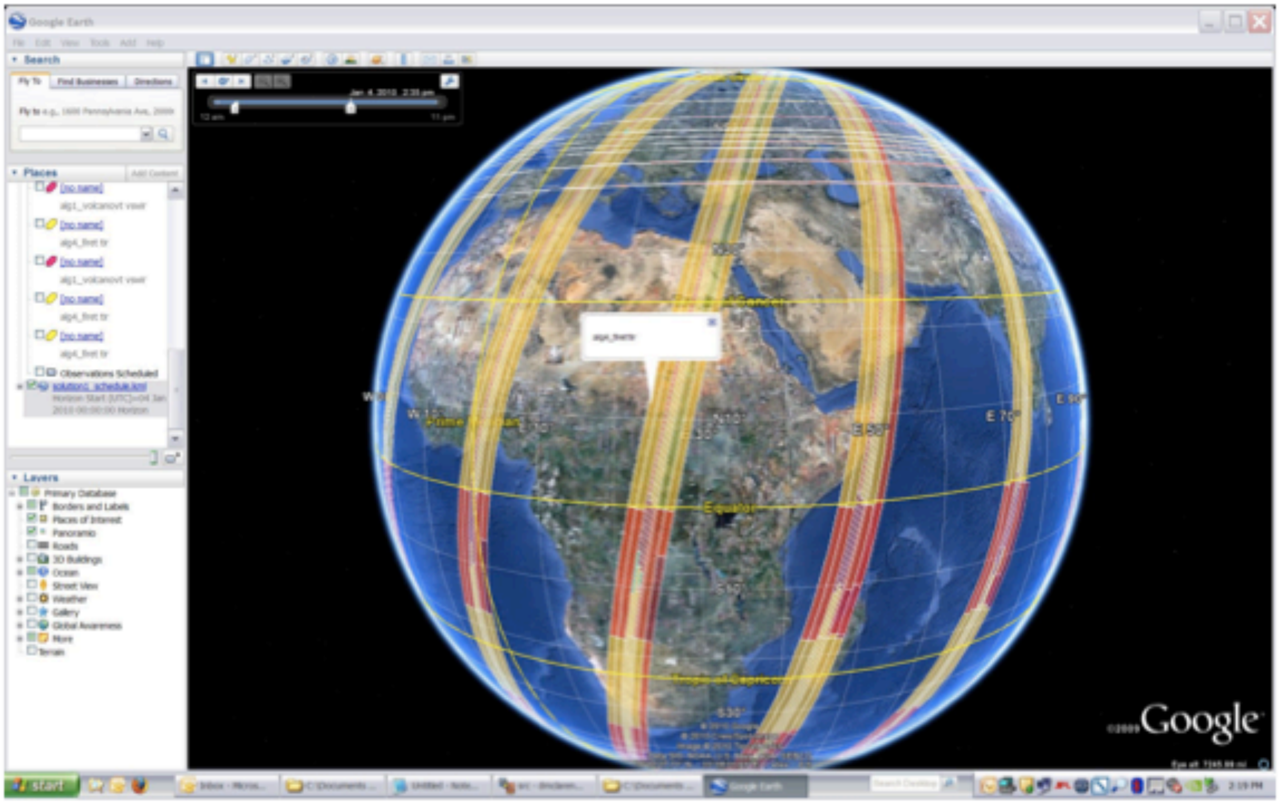

Figure 4: Sample processing plan.

challenging operations constraints.

In our discussion, we assume the context of a timeline based modeling framework, in which state and resource values are represented by a sequence of values. Furthermore, in our framework, value transitions are grounded in time (e.g. not flexible). For our specific class of scheduling problems this is a reasonable restriction because the overflight times or scene targets and data up/downlink times are fixed (therefore fixing most events of interest).

Naturally modeled spacecraft operations constraints

The HyspIRI product generation scheduling problem has a wide range of constraints that can be naturally represented in common planning & scheduling system modeling constructs.

Integer capacity – non depletable – this is an integer capacity of a resource that is used only for the duration of the activity. For the HyspIRI model onboard processing (CPU) and downlink bandwidth are non depletable resources. At any point of time the CPU cannot be subscribed more than 100% and the data volume being generated downlink bandwidth cannot exceed the available rate (10 x 10^6 bits per second). Note that in a later version of the model we are considering augmenting the model to allow for slight buffering which will make this a much more complicated depletable resource.

Integer capacity – depletable – this is an integer capacity resource reserved by one activity making a portion of the resource unavailable until it is freed by another activity. In contrast, non-depletable resources are used only for the duration of the activitiy (e.g. power).

Discrete states – there are numerous discrete state constraints. These both represent transition constraints and state constraints. These constraints are not yet incorporated into the HyspIRI model however we believe that they will not impact schedulable operations but will enable generation of an executable command sequence.

Decomposition – often a high level activity consists of several lower level activities. These are represented as Hierarchical Task Network planning decompositions. In the current HyspIRI model decompositions are also not modeled but we believe that they will need to be modeled to generate executable plans but again will not influence schedulable events.

Other operations constraints

In this section we describe how priority, short duration opportunities, and overflight are modeled.

One key aspect of HyspIRI product generation is prioritization of products. HyspIRI incorporates a simple, strict prioritization model. In this model a higher priority product strictly dominates any number of lower priority products. This is handled in the scheduling algorithm described below.

Short duration opportunities are also handled in a special case fashion. We schedule in a 10s temporal resolution so this would be the minimum length for any product being generated. Additionally, any two segments separated by 60 seconds or less that can be merged without conflicting a higher priority observation are combined.

Overflight is the primary operations constraint for HyspIRI product generation. Within the CLASP framework (see [Rabideau et al. 2010] for further details), overflight is modeled as described in the following sections.

Swath Generation

In swath generation, we first ingest the HyspIRI Earth ground track (projected for the proposed HyspIRI mission). This ground track shows the nadir points for the spacecraft over time for the scheduling period of interest.

Next, polygons are created from ground track points representing the area on the surface of the Earth that is viewable by the instrument. As discussed above, these represent the TIR and VSWIR instrument swaths underneath the spacecraft. A separate swath is generated for each instrument with the VSWIR being represented by four (4) narrow swaths and the TIR by four (4) wide swaths and also the four narrow swaths for VSWIR.

Additionally, certain instrument-mode combinations are not desired. For example, acquiring VSWIR images during the night would not generate useful data for most campaigns (exceptions being emissive campaigns such as volcanoes or forest fires). Certain other overflight-specific viewing constraints are also important to the scientists. These constraints include: day versus night, restrictions on solar zenith angle (how high in the sky the sun is at the point being observed). Because these constraints depend on the time of the observation we construct additional special instrument swaths for these potential observations.

These instrument coverage polygons and their time tags are combined to make the “Instrument swaths” that are passed as input to observation selection and product generation process (see below).

Product Generation Requests

Product generation requests are input as a tuples of <region, priority, algorithm> where “region” is a polygonal area on the surface of the Earth, “priority” is the priority of the product and “algorithm” specifies the algorithm to generate the product – in effect specifying the scheduling constraints such as data volume generated, CPU utilization, etc. These product generation requests form the “Regions of Interest” discussed below.

Opportunity Analysis

CLASP utilizes a spatial grid representation for its coverage reasoning. Therefore, the earth's surface is approximated as a grid of points with regions represented by the grid points within their polygonal boundaries. For a given grid separation S, there are a series of constant latitudinal bands with each band being S distance apart and each point in the band being S distance apart. For the HyspIRI model we use a grid separation of 12.5 km.

Opportunity analysis consists of, for each timepoint (recall that for HyspIRI we schedule at the 10s resolution), computing the instrument overflight polygons and intersecting these with the product generation request polygons. With the gridded representation the polygons correspond to sets of grid points and the intersection corresponds to set intersection on grid points. Each such non-null intersection represents an observation-product-generation opportunity or an "Observation Record" as discussed below.

Scheduling Algorithm

Scheduling the HyspIRI product generation therefore is selecting a subset of the observation records computed above, to maximize priority score, subject to operations constraints.

Given
a set of potential observation records $O = \{o_1 \ldots o_n\}$
a set of regions of interest $R = \{r_1 \ldots r_n\}$
a set of instrument swaths $I = \{i_1 \ldots i_n\}$
Where $\forall\, o_i \in O\ \exists (r_i, i_i)\ grid(o_i) \in grid(r_i) \wedge grid(o_i) \in grid(i_i)$
a scoring function $U(r_i)$ -> real
a constraint function $C(S) \rightarrow$ True, False
where $S \subseteq O$ and C is True if S satisfies spacecraft constraints

Select a set of observations A
To maximize $\Sigma_{a \in A} U(a)$ subject to $C(A) \rightarrow$ True

Because the current HyspIRI model has only local constraints (e.g. observing at time T only affects resources at time T) we currently schedule using a priority first greedy algorithm. In this approach, we first sort the observation records by highest priority score. We next walk through this sorted list adding all observation records that do not cause a conflict. This algorithm is believed to be optimal except for tie breaks within the same priority.

Current Status and Future plans

This onboard processing concept, mission operations concept, and associated automated planning system has been studied over the past 2 years for the HyspIRI mission. As part of this study, a prototype planning system was demonstrated on the current best predicted HyspIRI mission orbit, and proposed science and application products for onboard generation. This demonstration was realistic enough that operations of Direct Broadcast capability no longer considered significant risk item therefore are no longer under study. Related ongoing work includes a potential demonstration using the SpaceCube [Flatley 2010] hardware platform based on the Vertex 5 Pro chipset in both an airborne real-time processing demonstration and on a cubesat flight called IPEX.

Discussion, Related Work, and Conclusions

AI scheduling has been applied to a wide range of space operations problems including: Hubble observatory scheduling [Johnston et al. 1993], space shuttle refurbishment scheduling [Deale et al. 1994], controlling the Deep Space One spacecraft for 48 hours [Muscettola et al. 1998], radar campaign scheduling [Smith et al. 2002], coordination of multiple space assets to observe transient phenomena in a terrestrial sensorweb [Chien et al. 2005a], onboard replanning to enable capture of transient science events [Chien et al. 2005b], ground-based surface operations of the MER rovers [Bresina et al. 2005], Mars Express operations [Cesta et al. 2007], and Orbital Express operations [Chouinard et al. 2008]. To our knowledge the HyspIRI application is the first application of AI scheduling to overflight-based science product generation.

This paper has described the application of AI-based scheduling techniques to determining which science products to generate onboard the proposed HyspIRI mission. This problem is an oversubscribed scheduling problem which as currently modeled only involves local constraints (from a temporal perspective) enabling near optimal solution. As work continues to mature the concept it is expected that this problem property may no longer hold, requiring more heuristic solution of the scheduling problem.

References

J. L. Bresina, A. K. Jonsson, P. H. Morris, K. Rajan, Activity planning for the mars exploration rovers, Proc 15th Intl Conf Auto Plan & Sched, 2005, Monterey, CA.

Cesta, A., Cortellessa, G., Fratini, S., Oddi, A. and Policella. N., An innovative product for space mission planning: an a posteriori evaluation. Proc Intl Conf Automated Planning & Scheduling, 2007, Providence, RI.

S. Chien, G. Rabideau, J. Willis, T. Mann, "Automating Planning and Scheduling of Shuttle Payload Operations," Artificial Intelligence Journal 114 (1999) 239-255.

S. Chien, B. Cichy, A. Davies, D. Tran, G. Rabideau, R. Castano, R. Sherwood, D. Mandl, S. Frye, S. Shulman, J. Jones, S. Grosvenor , "An Autonomous Earth

Observing Sensorweb," IEEE Intelligent Systems, May-Jun 2005, pp. 16-24.

S. Chien R. Sherwood, D. Tran, B. Cichy, G. Rabideau, R. Castano, A. Davies, D. Mandl, S. Frye, B. Trout, S. Shulman, D. Boyer . , "Using Autonomy Flight Software to Improve Science Return on Earth Observing One, Journal Aerospace Computing, Information, & Communication, April 2005, AIAA.

S. Chien, D. Silverman, A. Davies, D. Mandl, Onboard Science Processing Concepts for the HyspIRI Mission, IEEE Intelligent Systems, Vol 24. No. 6, November/December 2009.

C. Chouinard, R. Knight, G. Jones, D. Tran, D. Koblick, Automated & Adaptive Mission Planning for Orbital Express, Space Operations, Heidelberg, Germany, 2008.

M. Deale, M. Yvanovich, D. Schnitzius, D. Kautz, M. Car- penter, M. Zweben, G. Davis, and B. Daun, "The Space Shuttle Ground Processing Scheduling System," in Intelligent Scheduling, Zweben & Fox (eds.), Morgan Kaufman, San Francisco, 1994.

R. Knight, S. Chien. "Producing Large Observation Campaigns Using Compressed Problem Representations," Proceedings of the Fifth International Workshop on Planning and Scheduling for Space, Baltimore, MD, October 2006.

R. Knight, S. Hu. "Compressed Large-scale Activity Scheduling and Planning (CLASP) applied to DESDynI," Proceedings of the Sixth International Workshop in Planning and Scheduling for Space, Pasadena, CA, July 2009.

N. Muscettola, P. Nayak, B. Pell, B. C. Williams, "Remote Agent: to boldly go where no AI system has gone before," Artificial Intelligence v103, # 1-2, pp. 5–47, 1998.

M. D. Johnston, R. Henry, A. Gerb, M. Giuliano, B. Ross, N. Sanidas, S. Wissler and J. Mainar, "Improving the Observing Efficiency of Hubble Space Telescope," AIAA Aerospace, AIAA Press, 1993.

G. Rabideau, , S. Chien, D. Mclaren, R. Knight, S. Anwar, G. Nehall, P. Christensen, "Scheduling targeted and mapping observations for the THEMIS Instrument onboard Mars Odyssey," Scheduling and Planning Applications Workshop, ICAPS 2010, Toronto, Canada, May 2010.

B.D. Smith, B.E. Engelhardt, and D.H. Mutz,"The RADARSAT-MAMM Automated Mission Planner," AI Magazine, v. 23, # 2, 2002. pp. 25–36.