Commentary on
"Explanations and Recommendations for Temporal Inconsistencies"

John L. Bresina and Paul H. Morris
NASA Ames Research Center
M/S 269-2, Moffett Field, CA 94043

Commentary by:

Robert Hawkins

Space Telescope Science Institute
3700 San Martin Drive, Baltimore, MD 21218
rhawkins@stsci.edu

Introduction

This focal research in this paper is on the addition of an explanatory module to the MAPGEN system, which is currently operationally deployed for both the Spirit and Opportunity rovers of the MER mission. MAPGEN is a mixed-initiative constraint based system for laying out the day-to-day plans for the rovers. The additions described in this paper work toward providing explanations and recommendations for the resolution of nogoods in generated plans. The work was conceptualized in (Bresina Et. Al. 2005a), and has been grounded and extended in this work.

MAPGEN uses the Simple Temporal Network (STN) framework of its constraint-propagation component, EUROPA, to manage constraints. There are 3 types of constraints in the MER mission - *model*, *problem-specific* and *expedient*. Model and problem-specific constraints are domain constraints having to do with either flight rules for the rovers and science objectives, respectively. Expedient constraints are a necessary addition to handle disjunctive model constraints given the underlying STN framework from EUROPA.

MAPGEN Planning

The primary users of MAPGEN are the Tactical Activity Planners (TAPs). The TAPs use the MAPGEN GUI to plan the activities for each Sol for a rover. The process is highly interactive and time critical, so the system needs to be responsive to the users. The TAPs have several levels of interactivity available to them when collaborating with MAPGEN, ranging from the *plan-all* operation where MAPGEN performs the entire planning process to the *super-move*, wherein the TAP maintains complete control, overriding the underlying constraints.

Because of the time critical and interactive nature of the MER operation planning process, design and problem solving "compromises" must be made to enable the planner utilize the tool in a responsive, interactive fashion.

The authors mention that the MAPGEN backtracking algorithm, for example, will reject activities that are taking too long to plan, striking a careful balance between optimality and responsiveness.

The use of a Disjunctive Temporal Network (DTN) in place of the STN would remove the need for the expedient constraints that are the focus of this research, but the cost in performance could potentially be too great. It is my opinion that this work helps bridge the gap caused by this "compromise" by enabling, in an interactive fashion, TAP resolution of expedient-caused nogoods.

Generating Explanations

The authors describe 3 contexts within which temporal inconsistencies can arise during the planning process. One is associated with the TAP-controlled *place-selected* operation. Another is associated with the more automated *plan-selected* operation. The context focused on within this research is the general context in which the activity is inconsistent with the current plan prior to being inserted.

The explanation generation process involves taking raw, lengthy MAPGEN nogood descriptions and summarizing and explaining them so that they will be tractable for a time-pressured TAP to resolve. There are several steps in this process.

The first step in the process of summarizing nogoods is to characterize the type of each constraint that is in the nogood. The second step is to compress the nogood, removing layers of expansion that aren't relevant to the explanation or resolution of nogoods.

As to the explanation, the authors describe 6 abstract examples of nogoods, but, for explanatory purposes, they all boil down to one basic issue – that there isn't enough room between two activities A and C to fit a third, B. The authors then break the explanation down into an intuitive two 2 steps – why A needs to move relative to C to accommodate B and why it can't.

I feel that these explanations should provide the TAP with the necessary background to understand why the insertion of a new activity and its associated set of constraints has caused a nogood. The method of resolving nogoods in the operational version of MAPGEN appears to be more or less taking a guess, presumably using the graphical interface to the timeline and general TAP knowledge and experience, as to which current activities are in need of unplanning to get the new activity into the plan (Author communication).

Given that, it is quite clear to me that the results described in this work would provide an significant improvement. Obviously, as the authors describe, there need to be a testing and evaluation process.

Recommending Solutions

Due to the nature of the problem, and the a priori constraint validation in the Constraint Editor (CE) the solution recommendation process is, of necessity, focused on the expedient constraints within the nogood.

The process for determining the recommendation involves 3 steps – determining all of the mutex activities in the nogood, replacing low-level activities in those pairs with the high-level TAP plannable activities, and finding a minimal set cover within this set.

Using this information, the system then provides the TAP with a set of steps for resolving the nogood and a backup set for the case when the first fails.

It isn't clear to me how the system handles the failure so that the backup recommendation can still be used – i.e. is there an undo feature? If not, I would think that it could be the case that after attempting the first recommendation, the second would no longer valid, but perhaps the backup (drastic) recommendations are general purposes enough that the user's previous actions won't affect this.

Future Work

The authors highlight a couple of paths for quantifying the results of this work, including finding more examples of nogoods to resolve, and evaluating the quality of the explanations with a user study of the system with the TAPs.

I agree with the authors that live deployment, or at least active testing with users is key to understanding the utility of this approach, and to guiding future development. It is understandably (and unfortunately) difficult to deploy large-scale updates such as this to an operational mission such as MER, so the authors' suggested path of getting this work in at the ground level for MSL seems to be a viable plan for the future.

It would be nice to see a quantified test of TAP performance with and without the explanation facilities, to give an idea of how this work aids in TAP performance.

The authors highlight a couple of paths for future work. One is the combination of multiple inconsistencies in the plan into one explanation and recommendation pair. The other is the potential of automated validation of the recommendations. I agree with the authors that these paths are likely to be fruitful.

It might also be useful to the potential users to have a mixed initiative theme to the future work on providing explanations and recommendations – such as guided recommendations or active explanation interfaces in the GUI. For example, the UI for recommendations could be an interactive UI like the "hopper", allowing the TAP to click on each step in the recommendation and be presented with the command dialog to enact that step, or the explanation dialog might be capable of highlighting in the timeline the activities or constraints that it is listing.

Conclusion

The authors present a useful extension to the MAPGEN system. This work continues and expands on the intuitions that were outlined in the referenced paper (Bresina Et. Al. 2005a), providing not only an implementation of the described compression concept but also derivation of useful explanations and recommendations for nogoods. Though replacing the DTN with an STN should resolve the need for the expedient constraints that are the causes of the nogoods addressed by this work, it isn't at all clear that that solution would be satisfactory in a sufficiently complex mixed initiative domain, such as that of MER. I hope to see this work continued in the MSL mission, and look forward to further results and evaluation of the current system.