Towards a Characterization of Scheduling Task Complexity

John A. Karasinski, John Bresina, Bob Kanefsky, Megan Shyr, and Jessica J. Marquez
IS-18, Human-Automation Interaction I
05-Jan-2022

Motivation

Astronauts currently have their schedules managed by a team of professional planners.

In the face of emergencies, unscheduled maintenance, or other unplanned events, these planners must reschedule as many tasks as possible while still meeting all the complex constraints between both activities and resources.

In future missions, astronauts are envisioned to act with **greater autonomy** and reschedule their own timelines.

This Study

Primary research questions:

1. How do different types and number of constraints impact human performance in the scheduling task?
2. Can we use computational approaches to scheduling difficulty to predict human performance in the scheduling task?

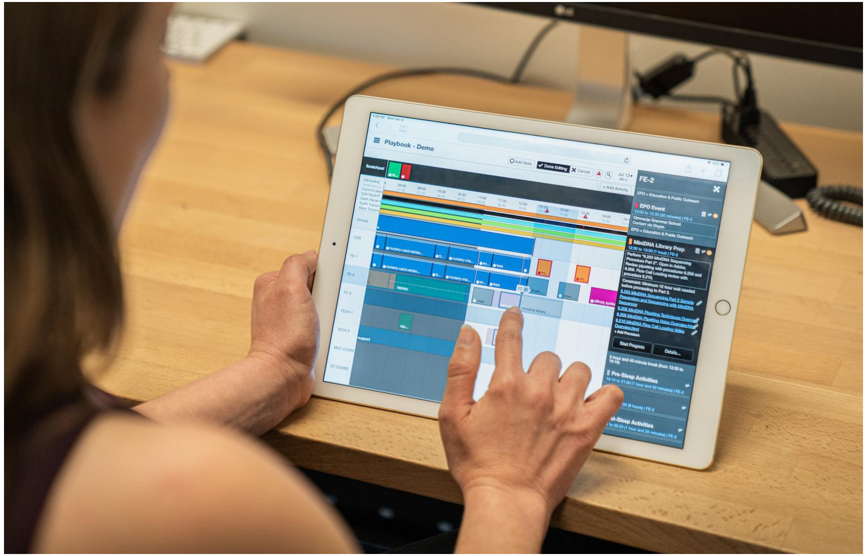
Identifying a **robust computational measure of scheduling difficulty** will allow us to **predict human performance** and identify scheduling problems that novice planners will be able to solve, and provide aids when they cannot.

Playbook

Playbook is a mobile, web-based scheduling software used to enable crew self-scheduling.

The timeline view displays time horizontally with each crewmember having their own row and their own activities.

An activity is displayed as a colored rectangle with the length of the block directly related to the duration of the activity.

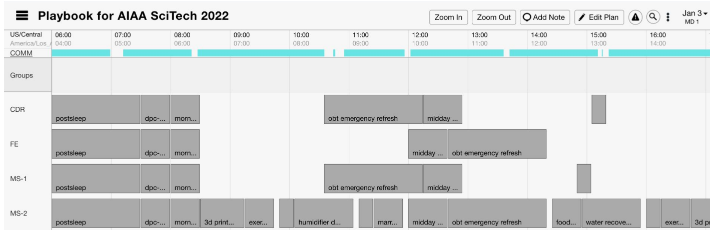


Experiment

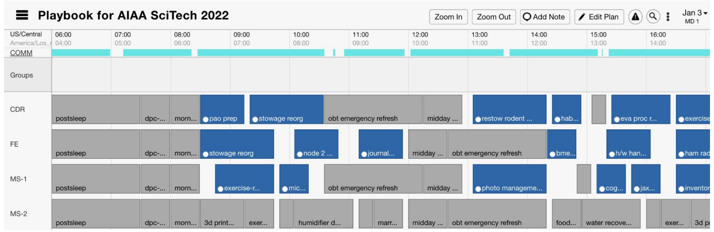
We recruited 31 participants which were split between two groups: scheduling or rescheduling.

Scheduling: initial timeline with some pre-planned, *inflexible* activities in the timeline, and blocks of time available for the participants to plan activities.

Rescheduling: initial timeline with both *inflexible and flexible* activities, and were then asked to reschedule as needed to meet their goals.



Scheduling



Rescheduling

The Task

Participants completed 8 trials with different types and number of constraints.

In each trial, activities had only one type of constraint at one of two levels:

- Low (33% of activities constrained)
- High (66% of activities constrained)

Participants were instructed to schedule the activities in order of priority, starting with high, then medium, and finally low.

Constraints

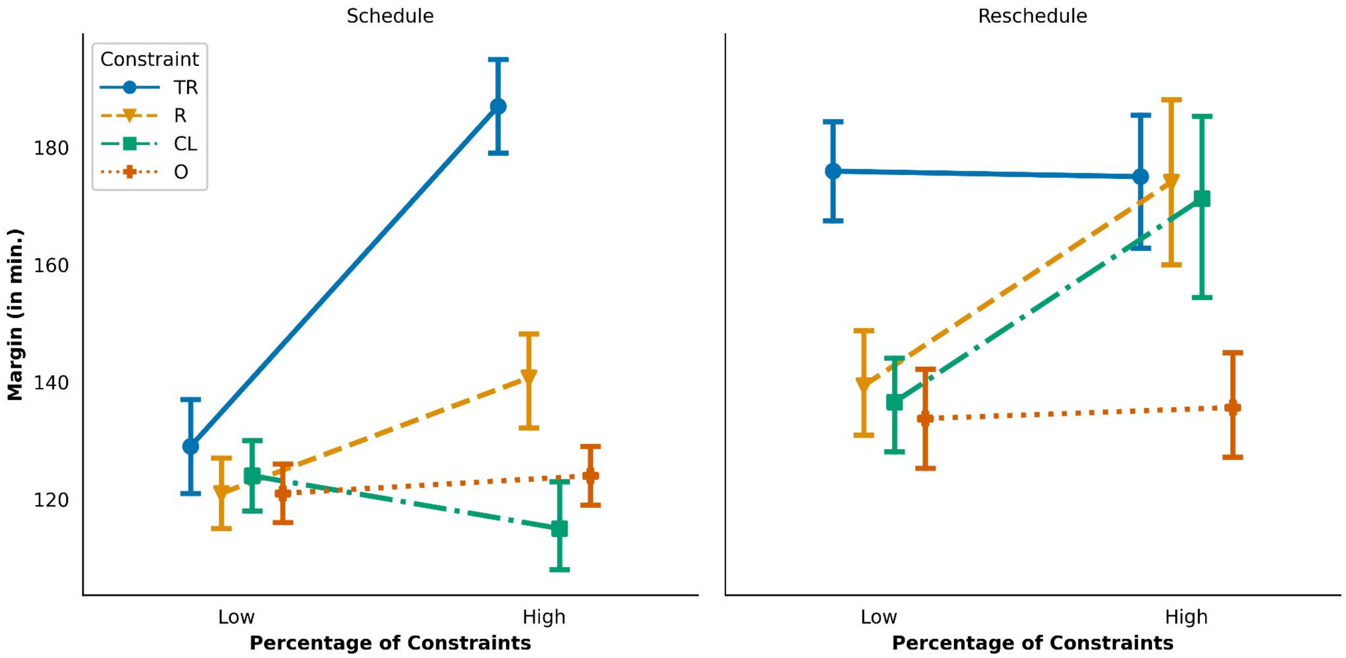
Time Range (TR) limits the time of day an activity can be scheduled.

Requires (R) states that the activity needs to have a particular resource available.

Claim (CL) describes a specific piece of equipment required for a particular activity.

Ordering (O) describes when an activity should be scheduled in relation to another activity.

Experiment Results



Modeling Approach

We designed an iterative sampling algorithm which mimics the actions that a human planner takes to solve the scheduling task.

Given an initial, partially scheduled timeline that consists of entirely inflexible activities, the algorithm takes random actions in an attempt to schedule all of the new activities in order of priority.

An activity is chosen from the set of activities to be scheduled (high, then medium, then low priority), a valid slot which the activity can be scheduled in chosen from the set of valid slots, and the schedule is updated to accommodate the activity.

Solution Quality

We defined the proportional solution quality

$$SQ_p = 10^2 * \frac{U_H}{T_H} + 10 * \frac{U_M}{T_M} + \frac{U_L}{T_L}$$

where $U_{H,M,L}$ and $T_{H,M,L}$ are the number of activities of unscheduled and total activities of high, medium, and low priority.

We collected 10,000 samples in each condition for this analysis.

Task	Constraint	Percentage	Min	Mean	Median
Schedule	CL	High	0.375	1.225	1.500
		Low	0.500	0.805	0.500
	O	High	0.250	5.897	1.625
		Low	0.250	3.000	0.500
	R	High	0.250	0.435	0.375
		Low	0.250	0.546	0.500
	TR	High	0.375	2.228	1.875
		Low	0.250	0.834	0.500
	Baseline	—	0.250	0.944	0.500

Ranking

We included **six measures of efficiency, accuracy, and workload** in order to capture the "difficulty" of the experimental trials, and **scheduling quality and margin** were used for the model results.

For the experiment and modeling metrics

- Data was separated between the two groups (scheduling, rescheduling),
- Converted to z-scores,
- Averaged and summed for each condition to create a final score,
- Ranked in order of relative difficulty.

Rank Comparisons for the Scheduling Task

(a) Experimental Data

Constraint	Percentage	Rank	Score
TR	High	1	1.36
CL	High	2	0.26
O	High	3	0.11
CL	Low	4	-0.06
R	High	5	-0.07
TR	Low	6	-0.08
O	Low	7	-0.34
Baseline	—	8	-0.41
R	Low	9	-0.68

(b) Algorithm Result

Constraint	Percentage	Rank	Score
TR	High	1	0.67
O	High	2	0.38
CL	High	3	0.25
CL	Low	4	0.15
O	Low	5	-0.12
TR	Low	6	-0.24
Baseline	—	7	-0.25
R	Low	8	-0.25
R	High	9	-0.59

Future Work

While we were able to create rankings which compared well between the experimental and algorithm results for the scheduling task, the rescheduling task proved more difficult to estimate.

The strategy taken by the rescheduling algorithm—to attempt to schedule high priority, then medium, then low—may not be reflective of the participants actual behavior.

Future work can aim to better identify the strategies that the rescheduling participants were taking and attempt to integrate these into a new rescheduling algorithm.