Rapid Scheduling: A Constraint Satisfaction Approach to River Usage

Team: 16677

February 13, 2012

Contents

1 Preliminaries

1.1 Introduction

FIRST LAUNCH FINAL EXIT

Floating down a river is easy enough, but imagine scheduling an entire season of rafting trips, for hundreds of people. A *scheduling problem* occurs when there are start and end times and places for a group of people or objects. The schedules of the group must be arranged such that all participants start and stop as desired. Between the start and stop, they must interact in a certain way with each other and their surroundings.

We consider the scheduling problem for a river with visitors who desire to raft down it. The goal is to schedule as many visitors as possible. Visitors, however, are seeking a relaxing and authentic experience, and will become unsatisfied if they do not receive one. While scheduling unsatisfied visitors is allowed, we seek to schedule them in such a way that as many visitors are as satisfied as possible.

1.2 Definitions and Terminology

- Big O: (of an algorithm) Also "time complexity." A function relating the time it takes to solve a problem to the inputs of the problem. For instance, a function whose big O is O(x) computes in linear time.

- Carrying Capacity: The amount of visitors a river can see in a single year, while maintaining the likelihood of successfully scheduling all visitors and keeping them satisfied.

- Constraint Satisfaction Problem (CSP): A type of problem solved by assigning values to variables. These values must be within specified domains. The problem is solved when all variables have values within the specified domains – eg, they have values such that all constraints are satisfied.

- Group: See "visitor."

- Itinerary: The ordered list of campsites a visitor will stay at each night.

- Non-deterministic Polynomial-Time Hard Problem (NP-Hard): A class of problems whose solutions, if they exist, behave unpredictably, and are provable in polynomial time – see "Big O." CSPs are a type of NP-hard problem.

- Visitor: Also "group" or "travel group." A person or a group of people seeking to travel the Big Long River by raft. Their raft could be powered by oar, or it could be motorized. Visitors are very irritable and will become unhappy if they see other visitors, don't get their desired water time each day, or don't get scheduled.

- Satisfaction: (of a visitor) Visitors seeking to raft the Big Long River start with a satisfaction of 100. This number will drop each time the visitor sees another visitor, or doesn't get their desired water time (too much or too little of it) each day. It is possible for a visitor's satisfaction to drop to zero by these two factors alone. It will be zero if the visitor can't be scheduled.

- Season: A six month season is assumed to have 180 days; that is, six months times an average of 30 days per month.

- Tolerance: The amount of campsites a visitor's travel plan can be altered by. For instance, a tolerance of 1 would allow a visitor desiring to move four campsites each day to move between three and five campsites each day. This is done in order to accommodate for other visitors who may be in the desired spot.

- Water Time: Visitors specify how much time they want to spend on the water each day. This value is the same for each day of the trip. A visitors' satisfaction rating will drop if they are forced to spend more or less than this amount of time on the water.

- X: The current number of trips taken by visitors down the Big Long River each year.

- Y: The number of campsites on the Big Long River. Unless otherwise stated, this was assumed to be 128, as provided by literature[2].

1.3 Assumptions and Simplifications

- Season Length: A six month season is assumed to have 180 days; that is, six months times an average of 30 days per month. Let it also be noted that this season is consistent (i.e. there exist no peak visitation months or months of higher flow).

- Trips ranging from "6 to 18 nights of camping on the river, start to finish" are assumed to allow between 7 and 19 days of travel on the river, start to finish.

- Visitors or groups of visitors travel at a constant speed throughout the trip, which is defined only by the type of boat they use. Oar-powered rubber rafts travel at 4 mph, and motorized boats travel at 8 mph.
- Visitors do not separate into more than one group or combine with other groups. Therefore, any group of visitors is scheduled the same as a single visitor.
- "Contact with other groups of boats on the river" is only counted if one group passes another group, or is passed by another group.
- Groups are capable of passing one another at any point (i.e. passing is not restricted by the width of the river or anything else at anytime).
- The satisfaction of a visitor is affected exclusively by the amount of contact with other groups that visitor has, how much water time they have compared to their desired amount, and whether or not they are actually scheduled.
- Visitors specify only their starting date, boat type, and water time, and only valid values for these parameters. Therefore: They may not start before the start of the season, or so close to the end of the season it would be impossible for them to finish; they may not have a water time such that their trip takes shorter than 7 or longer than 19 days; they may only choose an oar-powered rubber raft or a motorized boat.
- Visitors are willing to follow the schedule given to them exactly: no backtracking, skipping, or slowing down.

2 The Solution

Our model is structured as a constraint satisfaction problem solver. The general definition for such a problem is that - given variables (X), values (D), and constraints (C) - find an assignment of value(s) in D for each X such that all constraints in C are satisfied. In general then, a CSP may be defined as a set (X, D, C). Good examples of this problem set are 2-SAT and coloring problems[1].

We view the problem of scheduling rafting trips as a progressive and dynamic CSP: each day being a constraint satisfaction problem dependent upon the solution to the previous day's problem. By limiting the values any variable may take on and applying efficient algorithms, the CSP becomes tractable and we may successfully schedule river trips without the need of unreasonable computing power.

2.1 Constraint Satisfaction

In general, our approach is to first consider that there are exponentially many combinations of visitors to campsites during their trip. Many of these combinations, however, are invalid given the constraint of no two visitors being allowed to stay at the same campsite. This is perhaps the greatest limiting factor in the problem as it applies to all campers passing through the same territory on a given day. This is an $n-ary$ constraint: the constraint relating n variables. Such constraints are computationally challenging for many problems, but our problem space is limited enough that the computations are relatively painless ($O(n^2)$ with n as number of visitors) to pair all campers and detect collisions at any campsite. There were few other constraints involved in solving the problem as we were mostly interested in solutions to our schedule.

As we have been discussing, each visitor was considered to be a variable in the CSP and campsites were considered the values. The domain of these values (for each visitor) varied day to day depending upon the distance each visitor was capable of travelling. This travel distance was derived from a visitor's average speed of 4 or 8 mph, v, multiplied with a uniformly distributed random variable for desired daily time spent on the water, t. This variable is such that the travel distance per day would allow the visitor to complete the trip within the allotted time of 7-19 days. The following few equations explicitly describe these relationships:

$$v * t = OptimalDayDistance$$

$$v * t * 7 \geq L$$

So in generating any visitor, we ensure the visitor is reasonably parameterized for finishing a trip on time. Each visitor is also randomly assigned, uniformly distributed, a day for departure (dDay) such that

$$v * t * 7 + dDay \leq S$$

That is, a visitor's maximum travel speed will allow them to finish before the close of the season.

Visitors are generated en masse using a varitey of uniformly distributed variables. It is possible to load in a file containing visitor information so that the program will work on problems of corporeal, rather than testing and verification, import. Once visitors are generated and basic parameters are set (discussed in **Section 2.2**) the program begins attempting to produce a solution through our CSP solver.

2.1.1 Backtracking Search

In order to solve a CSP, one must navigate the vast multitude of combinations for variables and values. We utilize a backtracking search algorithm to assign campsites to our visitors on a daily basis[4]. On any given day d, the solver looks to see which visitors are set to depart on d and subsequently adds them to a list, *toDepart*. All visitors still on the river from the previous day are added to *toDepart*, creating a list of all visitors requiring some movement for day d. In order to determine appropriate movements, the visitors are assigned campsites (analagous to distance traveled for the day) one-by-one. Following a pairing, the full day's assignment is checked to see if this pairing is consistent with current progress. If the pairing doesn't violate any other existing pairings, it is added to the day's list of assignments. We then proceed to the next visitor.

Encountering failed pairings causes the algorithm to attempt another pairing until success is achieved or failure must be returned. If failure occurs, the algorithm returns to the most recent pairing, undoes this pairing, and then attempts a new pairing. The entire process is structured as a depth-first-search. When a day's assignment is complete, the algorithm creates a new CSP for the next day and executes again, deepening the recursing into another day. **Figure 1** highlights the major steps involved in this process and provides a simple example. The example is simplified so that only two visitors are being assigned a possibility of two campsites. The algorithm begins by assigning site 1 to visitor A, denoted by the diamond below and to the left of the tree-top. Then assigning site 1 to B is tried but found to fail, so the algorithm assigns site 2 to B (rightmost circle). At this point, a the day would "end" and the algorithm would begin attempting to schedule the new day. In the event of a failure on a new day, the algorithm continues jumping up a level, as seen in the example, even to the extent of needing to reschedule a previous day.

Our solver is optimized to attempt maximal satisfaction of the travelers. That is, the possible campsites a visitor may travel to are presented to the solver in order of ideal travel distance. While not every visitor is guaranteed to always travel their desired distance, the algorithm attempts to satisfy this "fuzzy constraint". We also calculate the number of encounters between visitors. After the itineraries are created, each visitor's campsite listing is checked to determine how many visitors passed and were passed by this specific visitor. By combining the amount of erring from ideal travel distance and encounters with other visitors, we obtain the aforementioned effect on satisfaction. This effect is calculated post-solution. Throughout the course of operation, the solver relies

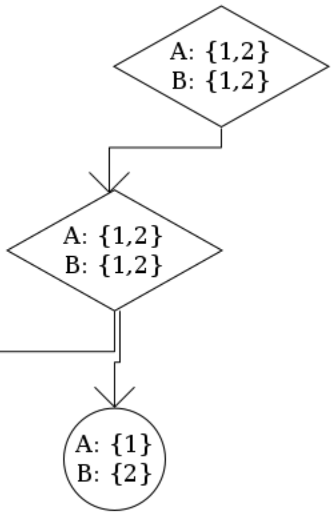

Figure 1: An example of the backtracking algorithm's attempt to schedule visitors A and B with campsites 1 and 2.

heavily on several parameters to relax the problem and make solutions more attainable.

2.2 Usage of High-Level Parameters

Given some of the information in the prompt and our findings, as discussed in **Sections 1.2** and **1.3**, our program is designed to adjust its operation depending on various conditions. The parameter of greatest variability, number of visitors, effects how many visitors we generate at the beginning, the load distribution on portions of the season, and some of the satisfaction scores for other visitors. Varying this value was crucial in analyzing the model, discussed in **Section 3.2**.

Values that tend to be static include the river length, number of days in the season, and number of campsites. While each of these is capable of drastically effecting the outcome of any given attempt, the values have a well-defined range (see **Section 1.3**). The problem could be generalized beyond these variables such that the CSP solver attempts to find optimal values for these as well, but this would likely place our problem into the realm of truly NP-hard problems.

2.2.1 Tolerance

Built into the algorithm is a *tolerance* variable: designed to create a range for the distance a visitor may cross in a given day. This variable is straightforward at the outset but provides several neat nuances for additional complexity in the solution. The introduction of a range to the CSP allows for a larger domain of solutions and, therefore, variability. In our case, we desire such flexibility, to account for the real-world implications of directing human beings. Tolerance extends within the model as a sort of aggregate for many of the various instances where we might expect variability — such as inclement weather, cancellations, or slow visitors—, without having to introduce variability into every aspect. This may seem like a stretch of the model's design, but any fuzziness for any aspect introduces variability into the solutions. We are satisfied that tolerance is an opportunistic variable for allowing more visitors to be scheduled and thus for the rafting season to be enjoyed by more individuals.

3 Predictions and Analysis

3.1 General Expectations

Before creating the model we developed a few "common sense" expectations of the results. For instance, it makes sense that increasing the visitor size would make it harder to schedule each visitor at their desired time. The more visitors one is trying to schedule, the faster the schedule fills up. However, as the number of visitors on the river during a season increases, it would become more and more likely that visitors would pass each other. Passing would even be mandatory if the visitors are travelling at different daily speeds; consider a visitor with a daily speed of 8 mph leaving a few days after a visitor travelling at 4 mph.

We also expected that as the tolerance for which campsite visitors stayed at each night increased, so would the ease of scheduling each visitor. The greater the tolerance, the greater the likelihood of finding a site to stay at if other visitors are also in the area. It also "made sense" to us, however, that an increase in tolerance would cause a decrease in overall happiness. An increase in tolerance would allow visitors to travel less than or more than their desired time each day. This would decrease their happiness.

It was expected that assigning more visitors to motor boats (those that travel at 8mph) would increase the success rate. Visitors that travel faster would spend fewer days on the river, freeing up space for other campers faster, and allowing more visitors to travel on the river each season. We also predicted that seasons with more uniform visitor speeds (e.g. seasons where all visitors travel at 4mph, or where all visitors travel at 8mph) would be happier. This is because visitors travelling at the same speed are less likely to pass each other, or to want to stay at the same campsite on the same night.

The greater the variation in visitor speed, the more likely it is faster visitors would have to pass slower visitors.

3.2 Results

Generally speaking, the results of the model were in line with our expectations. For seasons with a 2-site camping tolerance, the success rate was seen to decrease as the number of visitors increased (**Figure 2 Spark 1**).

The success rate was one-hundred percent up for seasons with 5, 50, and 100 visitors. Once the number of visitors per season increased past 100, the success rate began decreasing at an increasing rate. The success rate dropped from 100 percent to 91 percent between 100 and 300 campers, then from 91 percent to 25 percent between 300 and 500 campers. For seasons with more than 500 campers, the success rate leveled off, appearing to have an asymptote just above zero percent. The success rate for a season with 600 visitors was five percent.

Similarly, the average happiness of each of the visitors decreased as the number of visitors increased(**Figure 2 Spark 2**). Once the number of visitors per season increased past 100, the happiness rate began decreasing at an increasing rate. The happiness dropped from 97.9 percent to 89.4 percent between 100 and 300 campers, then from 89.4 percent to 24.5 percent between 300 and 500 campers. For seasons with more than 500 campers, the happiness rate leveled off, appearing to have an asymptote just above zero percent. The happiness for a season with 600 visitors was 4.9 percent.

For seasons with a 1-site tolerance, the effect of visitor size was seen to increase(**Figure 2 Spark 3**). Both the success rate and happiness of visitors began dropping with a visitor size of only 50 people. The success rate dropped from 100 percent to 96 percent when visitor sized increased from 5 to 50 visitors per season. By just 250 visitors, both the success rate and the happiness were zero percent.

The tolerance for which campsite visitors stayed at each night had a positive effect on success rate of the scheduler and happiness of the visitors (**Figure 2 Spark 4**). For seasons with 200 visitors, a tolerance of zero campsites had a success rate of 0 percent

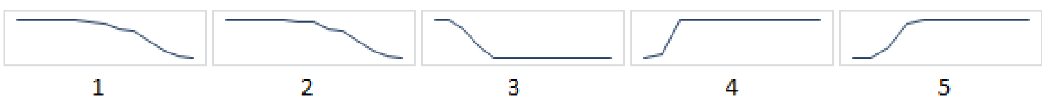

Figure 2: 1: Success Rate vs. Visitors 2: Satisfaction vs. Visitors 3: Satisfaction vs. Visitors 4: Tolerange vs. Success Rate and Satisfaction 5: Tolerance vs. Success Rate and Satisfaction

and happiness of zero percent. Increasing the tolerance to one campsite resulted in a success rate of 52 percent and a happiness of 4.9 percent. Increasing the tolerance to two campsites had the greatest effect, raising success rate to 97.8 percent and happiness to 95.9 percent. With any greater tolerances, the success rate was one-hundred percent, and the happiness leveled off at 97.8 percent.

For seasons with 500 visitors, a greater tolerance was needed to achieve similar results (**Figure 2 Spark 5**). A tolerance of 0 campsites had a success rate of 0; a tolerance of 1 campsite had a success rate of only 2.6 percent; 2 campsites, 28 percent; 3 campsites, 90 percent. The success rate was only 100 percent with a tolerance of greater than 4 campsites, and the happiness was 98 percent.

Assigning more visitors to motorboats (those that travel at 8 mph) was seen to increase the success rate. For a homogeneous season of 8 mph visitors, the success rate was 98.2 percent and the happiness 74 percent. At the number of 4 mph boats increased, both success rate and happiness steadily decreased. With half rowboats and half motorboats, the success rate was 92 percent and the happiness 72 percent. With a 9:1 ratio of rowboats to motorboats, the success rate of the season was only 89.9 percent. However, there is a jump in success rate to 99 percent once all visitors are travelling at 4 mph. This jump in successes is not accompanied by a jump in happiness; with all rowboats, the happiness is only 70 percent.

3.3 Analysis

As defined in our definitions section, CARRYING CAPACITY IS []. With a tolerance of 2 campsites (**Figure 3**), we found the river could manage 100 visitors per season with a 100 percent success rate and a happiness of at least 97 percent. It is worth noting, however, that doubling the capacity to 200 visitors per seasons yields a success rate of 99.5 percent, and a happiness still greater than 97 percent. Once the number of visitors per season increases past 200, both success rate and tolerance begin to drop rapidly. With a tolerance of 1 campsite, we found the river could only manage about five visitors per season with a success rating of 100 percent. The average happiness of these visitors is 97.7 percent. After the number of visitors per season increases above 50, both happiness and success rate begin dropping rapidly.

OUR CARRYING CAPACITY - X IS THE ANSWER

SANITY CHECK USING NUMBER OF PEOPLE EVERY YEAR AND SUCH

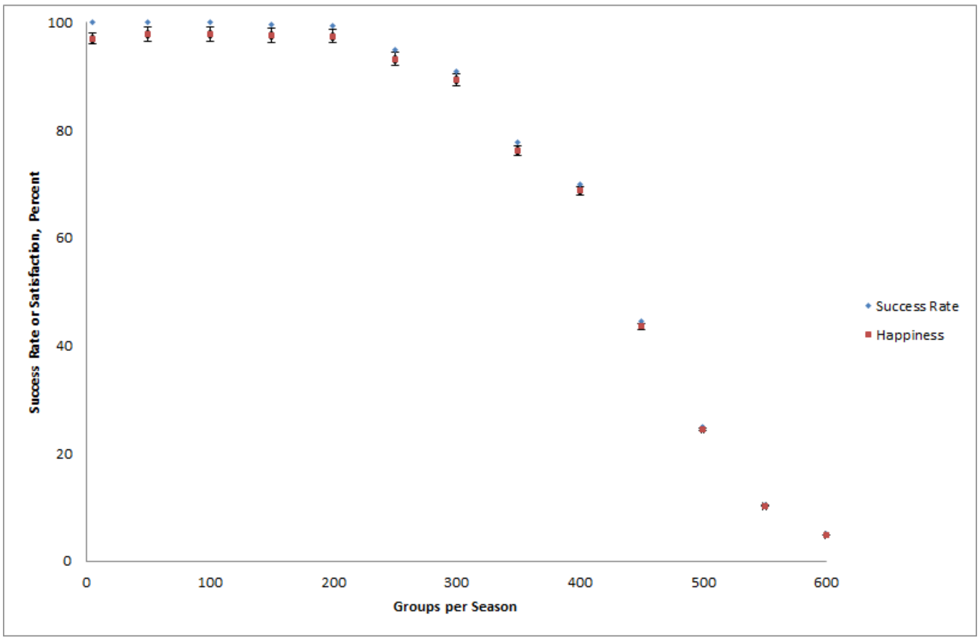

Figure 3: Groups per Season vs. Success Rate and Satisfaction with Tolerance 2

4 Final Thoughts

4.1 Advantages to the Approach

Throughout, we have seen that a constraint satisfaction approach to scheduling visitors to a river is a natural fit. This approach allows for a logical approach to the scheduling process, and for fairly quick solutions. In most cases, solutions were found on mid-grade hardware in less than ten minutes. We also found that the structure of our solution in both code and at a higher level lent itself well to the adjustment of any number of parameters. Moreover, there exists a great deal of flexibility once one begins to define "fuzzy values" that we can assign to the variables. Given more time, extending the model with these fuzzy values would likely lead to better assignments and even more interesting results.

4.2 Detractors from the Approach

river variation perfection of visitors vary start date In our assumptions (**Section 1.3**) we note that the visitors are basically perfect. This, of course, is not true. Visitors will probably vaguely follow the schedule given them: due either to inclement weather, rudeness, or physical constraints. It is not immediately apparent how to deal with

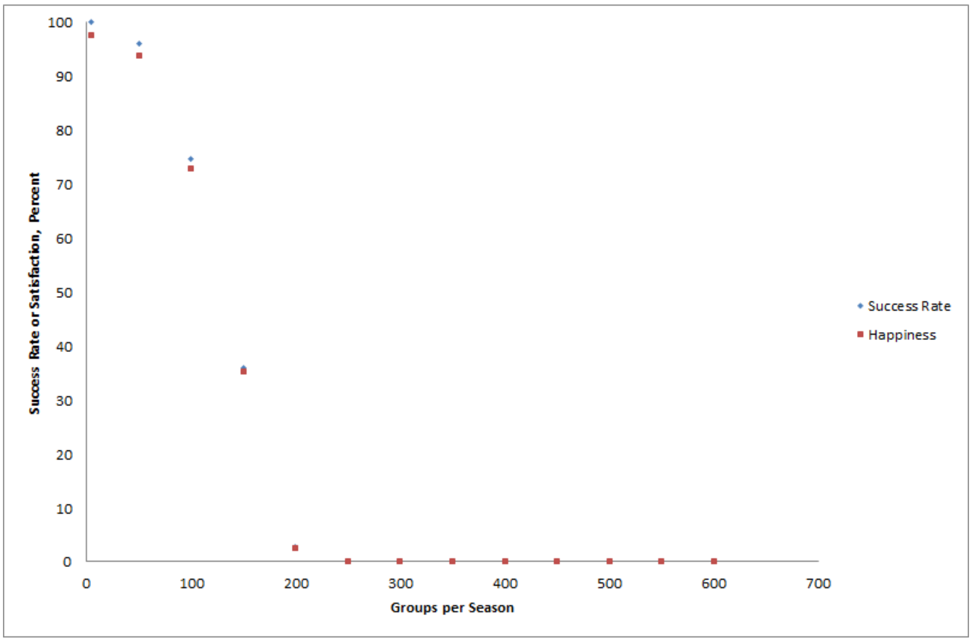

Figure 4: Groups per Season vs. Success Rate and Satisfaction with Tolerance 1

this shortcoming other than simply accepting that not all variables can be accounted for. Padding the schedules with "free days" on which any amount of travel lag may be recuperated.

The use of a consistent river in our model is less detrimental to the overall performance and accuracy. A realistic river would include areas where passing other groups is impossible, visitors travel faster or slower, and the potential for divisions of the river. These could all be built in as constraints on the campsites and their assignment to visitors. Should we have more time, this would likely be incorporated into the model.

4.3 Future Extensions

Given more time there are several additions that would improve the technical and practical performance of the model:

- The addition of a filtering technique to detect early failure in the backtracking search. This is often performed as what is called "arc-consistency". Regretfully, it is difficult to implement for our n-ary constraints.
- Utilizing a fuzzy constraint for departure days would allow groups to have a window in which they might leave. This is likely to increase the number of possible visitors, considering that it is nearly a direct analog to tolerance.

- The schedule might be modified to include days of rest for campsites on a rotating basis. This would not aid in adding more visitors (in fact it would decrease the success rate), but it would assist in land management practices and encourage a healthy use of the river ecosystem.

4.4 Final Recommendations

References

[1] R. Dechter and A. Dechter, *Belief maintenance in dynamic constraint networks in proc. of aaai-88*, 1988.

[2] Will Hansen, *Whitewater Campsites colorado–grand canyon*, 2009.

[3] Catherine A. Roberts Randy Gimblett, *Integrating gis and agent based modeling techniques for understanding social and ecological processes*, Oxford Press, 2000.

[4] Stuart J. Russell and Peter Norvig, *Artificial intelligence: A modern approach*, Prentice Hall, 2010.