

# So You Think You Can Graduate

Julie Jiang

December 15, 2016

## Abstract

College can be overwhelming, especially when it comes to trying to meet all the graduation requirements, which are quite interdependent. For example, a Math major student at Tufts must complete 13 classes in the Math department. One of these classes is Math 87 Math Modeling, which requires the completion of Math 34 or 39 and Math 70. A student's time is also limited – one can only take 5 or 6 classes at most per semester, for a total of 8 semesters.

Given the constraints, what are the possible schedules of classes to take each semester so that a Math major can graduate on time? In this project, we will explore the different sequence of classes that a Math major must take in order to graduate. Our interest is to optimize the class scheduling problem. Like other scheduling problems, we want to aim to reduce the total time (or the number of semesters) spent fulfilling the requirements and maximize the utility of each class that will be taken. The utility of a class is defined by how much it contributes to fulfilling graduation requirements.

We will use a depth-first backtracking algorithm to generate all the possible sequence of classes that satisfy the Math major requirements. We will then analyze the “best” sequences, which are the ones that complete the requirement in the shortest amount of time and/or smallest number of classes. We will then use a Gantt chart to illustrate the findings.

## 1 Introduction and Assumptions

All class prerequisites are taken from either the [course descriptions](#) listed on the Tufts Math Department webpage, or the [Tufts Student Information System](#) webpage. For this model, we consider only the classes that are offered in the semesters of Fall 2016 and Spring 2017, which means some of the optional classes that normally can be used to fulfill the requirements will not be considered in this model if they are not offered in the year 16-17 (such as Math 72 Abstract Linear Algebra). This is because information on the enrollment prerequisites on these classes either can't be found online or is outdated.

Without loss of generality, we also neglect to consider classes that require nothing but ‘departmental consent’ (such as Math 104 Math Education Change and Invariance). Classes that can be used to fulfill the Math major requirement but are not listed in the math department are also ignored, since they may or may not encompass a whole different set of prerequisites.

In addition to the class prerequisites, we also factor in the semester of which a class is offered. While most classes are offered both semesters, some classes, such as Math 135

Real Analysis I, is offered only in the Fall, and some classes, such as Math 136 Real Analysis II, is offered only in the Spring. To account for this, we make this distinction between Fall and Spring semesters.

For the remainder of this paper, we will refer to the classes by their course number.

## 2 Math Concentration Checklist

The required classes for a math major are

1. Math 42 or Math 44
2. Math 70
3. Math 135
4. Math 145
5. Math 136 or Math 146
6. Three additional Math classes with course number above 50
7. Three additional Math classes with course number above 100

In this model, we assume that the first five core classes have the following prerequisites and semester requirements

- Math 42 requires the completion of either Math 34 or Math 39 and is offered both semesters.
- Math 44 requires the completion of either Math 39 or (Math 34 and Math 70) and is offered both semesters.
- Math 70 requires the completion of either Math 39 or Math and is offered both semesters.
- Math 135 requires the completion of (Math 34 or Math 39) and Math 70 and is offered only in the Fall.
- Math 136 requires the completion of Math 135 and is offered only in the Spring.
- Math 145 requires the completion of Math 70 and is offered only in the Fall.
- Math 146 requires the completion of Math and is offered only in the Spring.

## 3 Algorithm

The algorithm used to generate all possible sequence of classes is a depth-first, backtracking algorithm. For our purposes, a class may have more than one set of classes that can satisfy its prerequisites. For example, Math 87 requires Math 32 or Math 39 *and* Math 70. In logical notation, this means that

$$\text{Requirement}(\text{Math } 87) = (\text{Math } 34 \vee \text{Math } 39) \wedge \text{Math } 70.$$

To fulfill a prerequisite for a class, we only need to satisfy one set of requirements.

Below is a abstracted pseudocode of the algorithm.

```

SCHEDULE(class, sequence)
Initialize an empty list of new-sequences
foreach set of prerequisite that satisfy this class:
    foreach pre-class  $\in$  prerequisite:
        if pre-class is not found in sequence:
            new-sequences = new-sequences + SCHEDULE(pre-class, sequence)
return the shortest sequences in new-sequences

```

The full source code can be found on my [Github Repository](#).

## 4 The Modeling Problem

Preliminary explorations suggest that given all the constraints of the Math major requirement, a Math student has to take 11 courses in total to satisfy the requirement. This includes one calculus II class that is not required as a major concentration but is required as a prerequisite for many other classes.

One important constraint in choosing a sequence of classes to take is that a student typically takes between 3 to 6 classes every semester. There are, of course, other classes that a student has to take that lies outside of the Math requirement, but we won't consider those because we want to focus more on investigating how Math major requirements specifically affect graduation.

We extract the best possible sequence of classes for a given *quota* per semester, where quota is the maximum number of Math classes a student can take every semester. The best sequence of classes are defined to be the ones that require the least amount of classes taken in total, and spans the least amount of semesters.

An example of one of the best sequence of classes given a quota of 3 is the following:

- 1st year fall: Math 34
- 1st year spring: Math 42, Math 70, Math 112
- 2nd year fall: Math 135, Math 145, Math 87
- 2nd year spring: Math 136, Math 158, Math 146
- 3rd year fall: Math 161

The number of possible permutations grows with the quota, as to be expected, since larger quota introduces larger variation. However, it is interesting to see that there is a lower bound for the number of semesters needed. No matter how many classes a student is allowed to take each semester, he or she needs at least 4 semesters to complete all requirements. This is due to the constraints placed on some of the core classes in terms of their prerequisites.

We also observe that if a student takes only 1 Math class every semester, then he or she will not be able to graduate on time (in 8 semesters).

Quota per Semester	Number of Semesters Needed	Number of Possibilities
1	13	5109
2	6	5850
3	5	5204
4	4	7280
5	4	10910
6	4	11158

## 4.1 Sensitivity

We want to find how sensitive each class is to graduation. To do this, we will calculate the free float of each class. Free float is defined to be the number of semesters a course can be delayed taking without delaying any other courses.

To do this, for each possible sequence obtained above, we backtrack from the last course that is scheduled to be taken to find the last possible semester it can be taken while still maintaining the total number of semesters.

In the example above, Math 87 has a free float of 2 semesters, and the rest have no free floats. Which means that the rest of the classes cannot be delayed – they must be taken on time. Math 87 has large total floats because it is offered both semester and and does not require as many prerequisites.

Below is the same example of sequences above, with Math 87 moved two semesters after:

- 1st year fall: Math 34
- 1st year spring: Math 42, Math 70, Math 112
- 2nd year fall: Math 135, Math 145 , Math 87
- 2nd year spring: Math 136, Math 158, Math 146
- 3rd year fall: Math 161, *Math 87*

It is easy to see that free float is dependent on the quota, or the maximum number of Math classes that one can take every semester. The more space. So we will repeat this process for each quota, and avearge the free floats for each class. Most classes have free floats of 0, meaning they are critical to the sequence to gradaution. The classes that have nonzero free floats are listed below.

Course Name	Avg free float			
	Quota = 3	Quota = 4	Quota = 5	Quota = 6
Math 42	1.33	0.09	0.14	0.15
Math 44	0.21	0.03	0.08	0.10
Math 51	0.68	0.42	0.97	1.00
Math 87	1.92	0.41	0.97	1.00
Math 112	0.52	1.14	1.92	2.00
Math 135	0.37	0.00	0.00	0.00
Math 145	0.37	0.00	0.00	0.00
Math 150 Poverty	0.44	0.00	0.00	0.00
Math 150 Sci Comp	0.62	0.00	0.00	0.00
Math 161	0.29	0.00	0.00	0.00

About half of all the classes have a consistent free float of 0, or no total float, and are therefore not shown. The majority of those that do have total floats have them smaller than 1, which means that on average they either can't be delayed or are only delayed for a semester.

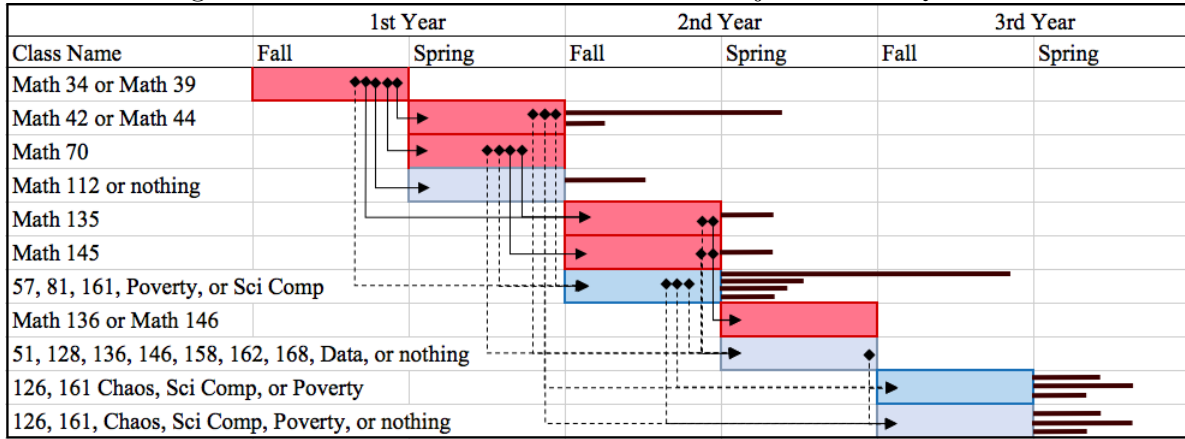
## 4.2 Gantt Charts

To help us visualize the data, we can build a Gantt chart, a type of bar chart that illustrates a project schedule. A Gantt chart is particularly insightful as it clearly encodes the relation of activity dependencies to time [1]. The following figure is the Gantt Chart for the best possible sequence of classes when the quota is 3. The cells in red mark the core classes, and the cells in blue mark the electives. The arrows indicate dependency relations. For example, The arrow that originates from the far right of the first cell means that Math 42 or Math 44 depends on Math 34 or Math 39. The dotted arrows indicate that the dependency may or may not exist, given the possibility of classes that cell represents.

Since there are more than 5000 possible combination of sequence of classes, it is simply not feasible to draw each and every one of them out. The following graph is a simplified version. Without loss of generality we can group Math 34 and 39 together, since they are both calculus II classes (one is standard level calculus and the other is Honors level), and for the same reason Math 42 and 44 are grouped together.

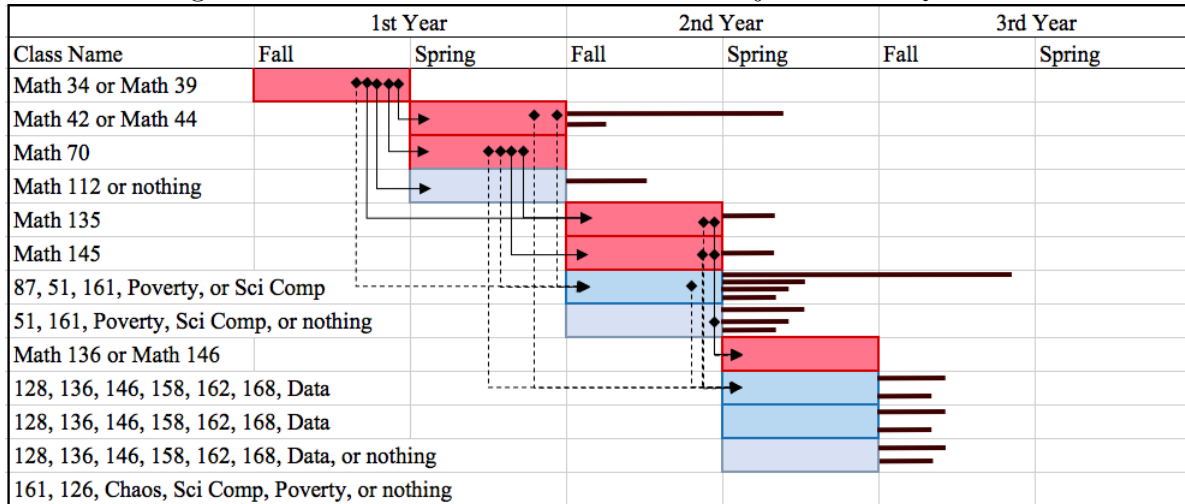
The cells that have a tail of black line connected to it indicate the slackness of the class, or its average free float.

Figure 1: Path to Graduation for Math Majors When Quota = 3



Below is another Gantt chart when quota is 4. Some of the dotted lines are removed to enhance clarity.

Figure 2: Path to Graduation for Math Majors When Quota = 4



From these charts, it is clear that the number of semesters in order to fulfill graduation requirement is at least 4, regardless of the quota. We also observe that Math 112 is a highly selected elective. This is in part due to the fact that Math 112 requires only the completion of Math 34 or Math 39, which is a lot less than what a lot other classes above the hundred level require. Therefore, Math 112 can be taken a lot earlier.

## 5 Discussion

In this model, we assume that a student wants to take as many Math classes each semester as the quota allows. We also assume that a student has no desire to take an additional Math class that may not fulfill graduation requirement but may bring about personal satisfaction. This is due to the fact these variations cannot be very mathematically well defined.

We also highly abstracted the complexity of graduation requirements. For example, the distribution requirements required by the School of Arts and Sciences is entirely neglected. This is due to the fact that many of these requirements do not have prerequisites, and thus will not add to the dimensionality of our scheduling problem, but will exponentially increase the possible permutations.

That said, including the full package of the graduation requirement will improve the accuracy of the model. Currently, all the free floats are calculated as the number semesters a class can be delayed without delaying any subsequent class in the sequence, but the sequences are designed to complete the math requirement as fast as possible, with no consideration whatsoever of the fact that a student has eight semesters in total.

As another room for improvement, we could also introduce probability to our model. Currently, this model is pure deterministic, but life is not. A student might not be able to take, for one reason or another, a class that he or she initially scheduled to take.

## 6 Sources

[1] Clark, Wallace, Walter Nicholas Polakov, and Frank W. Trabold. The Gantt chart: A working tool of management. Ronald Press Company, 1922.