CS 2800 Midterm Project Report

CLASS SCHEDULER USING A SAT SOLVER

Robert Sanders
Ohm Nabar

General Problem Area

Our objective is to create a program that students could use to input their preferences on the classes they would like to take in a semester. The program would consider these preferences and some built-in constraints and determine if a schedule matching these preferences exists using a SAT solver. This is an interesting problem for us to approach, as we have spent significant time in the past trying to create our own schedules while determining how to slot in certain requirements. For the purpose of this project, we consider many factors such as the time of each class section, required courses, required subjects, and gen-ed requirement categories. This is a really interesting project, as every student at Northeastern selects their own schedule manually. They may be able to guess what sections to take, but often students struggle to find a schedule meeting all of their needs. Our program may offer a new perspective on how to handle this issue.

Approach

Improvements on Traditional Approaches

Traditional approaches to finding a satisfiable schedule include ranking preference of classes and attempting to find combinations of sections that fit in one's limited schedule. This randomized brute force section selection could lead to many invalid schedules being selected. This approach also includes pre-determining which classes to take in the semester. Our approach will alleviate much of the need for manual work on the user's side.

Our Approach

We encode our problem as an input to a boolean satisfiability solver. Our program is coded in Python and uses a Python library for the SAT solver itself. Assuming that the constraints we handle are SAT-encodable, our program works because the SAT solver has the capability to determine a satisfiable assignment of boolean variables, if one exists. For our purposes, this corresponds to an assignment of sections.

Goals

Our goals for this project were to gain more knowledge in the field of SAT solving by going through the process of translating a real-life problem into a SAT problem and learning how to use a real SAT solver. In addition, because we are currently college students who have to schedule classes every semester, this topic matter is of interest to us and depending on where this project takes us, it may potentially have real world applications to how we schedule classes in the future. To be of the most use to us, we intend this project to sufficiently outline a satisfiable schedule based on certain subject, class, and time constraints. Overall, this project is an interesting foray into the world of SAT solving.

Methodology

Data Sample

First, we have to generate mock class and section data for our program. We represent both of these as Python objects. Classes have a subject, a class number, and an associated NUPath requirement. Sections have a section id, a class name - which is a combination of a class's subject and number (e.g. "CS 2800"), and a start time. A class name can uniquely identify a class, and a section id can uniquely identify a section. The section ids are positive integers, which will allow them to be used as boolean variables in the SAT solver, where the variable "s" means the section with id "s" is being taken and the variable "-s" means the section with id "s" is not being taken. For the sake of simplicity, we choose to have all classes starting on the hour and lasting less than an hour so that one can tell if two classes overlap simply by comparing their start times. We create csv files representing 22 classes and 82 sections and read from those files in Python to create our data. We store the sections and classes in dictionaries whose keys are the identifying attributes for their values (class name for classes and section id for sections).

Schedule Constraints

Next, we must encode our built-in schedule constraints. A student may not take more than 4 sections, more than 1 section of the same class, or more than 1 section at the same time. These kinds of constraints are a particular type of "cardinality constraint", which is a constraint on the number of true variables from a list of variables. In this case, the relationship between the cardinality and the variables is <=, as in the number of true variables must be <= the given cardinality. Let us examine the naïve approach to this problem with our most complicated constraint, which limits the total number of sections to 4. This would require clauses (not converted to CNF for simplicity) like "$(s_1 \wedge s_2 \wedge s_3 \wedge s_4 \wedge -s_5 \wedge -s_6 \wedge ...) \vee (-s_1 \wedge s_2 \wedge s_3 \wedge s_4 \wedge s_5 \wedge -s_6 \wedge \dots) \vee \dots$". One can see that this would take $\binom{n}{4}$, or quartic, time. More troublingly, this would generate a quartic amount of clauses, which could slow down the SAT solver if the number of sections is high.

We instead decide to use sequential counter encoding to enforce cardinality constraints. Given a set of variables $(p_1, p_2, \dots, p_n)$ and a cardinality k, we generate nk new auxiliary variables such that variable c_{ij} represents whether the total number, or count, of true variables has reached j by p_i. We require k clauses for our base case of p_1. The first is $p_1 -> c_{11} (CNF: \; p_1 v s_{11})$, which represents the fact that if p1 is true, the count has reached 1 by p_1. For all j such that 1 < j <= k, a clause s_{1j} is created, which represents the fact that the count cannot be greater than 1 at the first variable. For all p_i where i > 1, we require 2k + 1 clauses. If the count has reached 1 at the previous variable, or the current variable is true, then the count must have reached 1 at the current variable. This is represented by $p_i \vee (c_{(i-1)1}) -> c_{i1} (CNF: (\; p_i \vee c_{11}) \wedge (\; c_{(i-1)1} \vee c_{i1}))$. Additionally, if the count has reached k by the previous variable, then the current variable cannot be true. This is represented by $s_{(i-1)k} -> \; p_i (CNF: \; c_{(i-1)k} \vee \; p_i)$. Finally, we add a constraint for all j such that $1 < j <= k: ((p_i \wedge c_{(i-1)(j-1)}) \vee c_{(i-1)j}) -> c_{ij} (CNF: (c_{ij} \vee \; p_i \vee \; c_{(i-1)(j-1)}) \wedge (c_{ij} \vee \; c_{(i-1)j})$. This states that if the current variable is true and the count has reached j-1 by the previous variable, or the count has reached j by the previous variable, then the count has reached j by the current variable. In total, we create k + (n-1)(2k+1) clauses. Given that we are dealing with very low cardinalities, this is effectively linear with regard to the number of variables.

User-Defined Constraints

Now, we handle our user-defined constraints: section constraints, class constraints, subject constraints, NUPath constraints, and time constraints. A section constraint requires that a particular section be included in the schedule and creates a single variable clause s, where s is the section id of the desired section. A class constraint requires that a particular class be taken, and is encoded as $s1 \vee s2 \vee \ldots \vee sn$, where all s_i are sections of the desired class. A subject constraint is encoded the same way as a class constraint, except that all the sections' classes are of the desired subject. A time constraint takes in a range of acceptable times for sections to be scheduled and creates clauses $s_1 \wedge \; s_2 \wedge \ldots \wedge \; s_n$, where all s_i are sections not in the specified range. When the user is ready, they ask the program to determine the satisfiability of their schedule constraints. If the schedule is satisfiable, the program will say so and give a possible assignment of classes; otherwise, it will say that the schedule is not satisfiable.

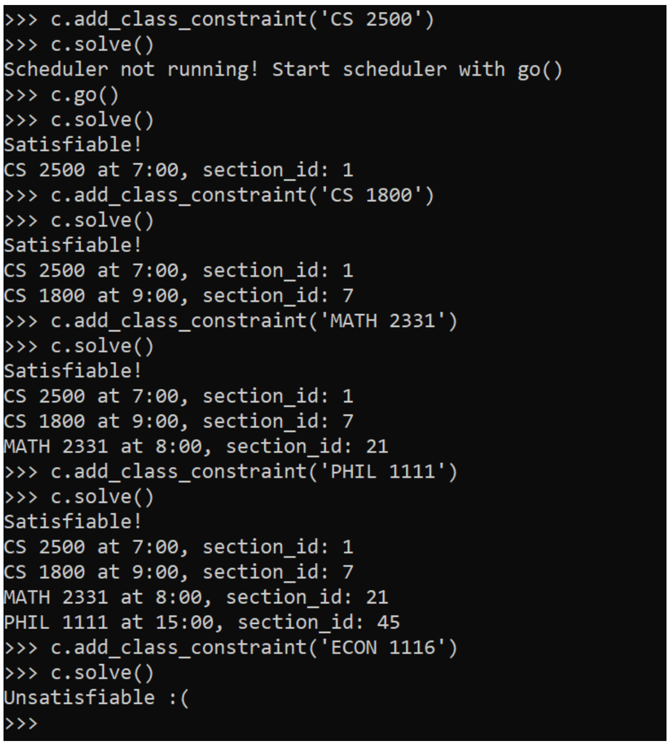
Results

Satisfying Goal

It was very interesting to analyze the output of our class scheduler along different constraints, and combinations of classes. Our goal was to find an available schedule when provided with a dataset of classes. We successfully completed our goal, as is demonstrated below. The sat solver finds a valid schedule based on the user's desired sections, subjects, classes, timeframe, and NUPath constraints.

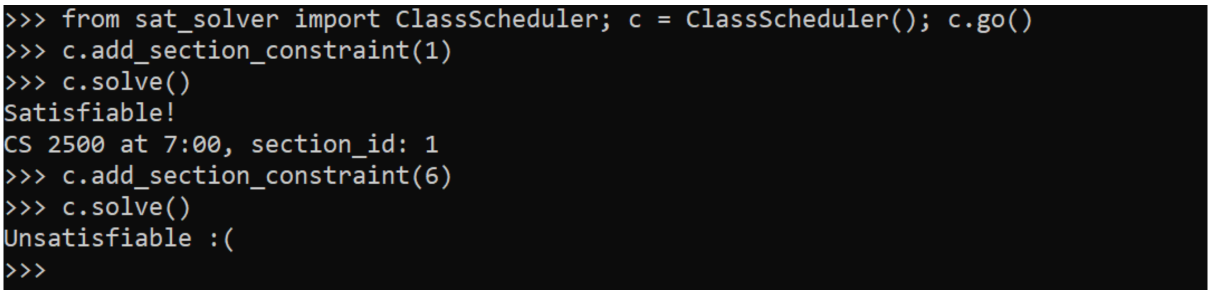
Adding Class Constraints

In Figure 1 we add a full schedule of classes, choosing sections with different starting times, and of different classes, and as we see the full schedule is satisfiable. Then we add 1 more class (ECON 1116), which should not work because we can only take 4 classes, and as we see the class scheduler returns that it is unsatisfiable.



```
>>> c.add_class_constraint('CS 2500')
>>> c.solve()
Scheduler not running! Start scheduler with go()
>>> c.go()
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
>>> c.add_class_constraint('CS 1800')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
CS 1800 at 9:00, section_id: 7
>>> c.add_class_constraint('MATH 2331')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
CS 1800 at 9:00, section_id: 7
MATH 2331 at 8:00, section_id: 21
>>> c.add_class_constraint('PHIL 1111')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
CS 1800 at 9:00, section_id: 7
MATH 2331 at 8:00, section_id: 21
PHIL 1111 at 15:00, section_id: 45
>>> c.add_class_constraint('ECON 1116')
>>> c.solve()
Unsatisfiable :(
>>>
```

Figure 1: Evaluating the satisfiability of adding class constraints

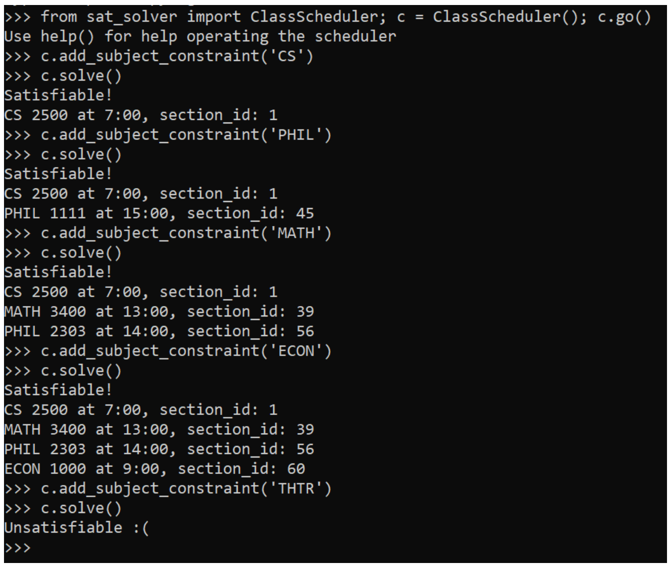


```
>>> from sat_solver import ClassScheduler; c = ClassScheduler(); c.go()
>>> c.add_section_constraint(1)
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
>>> c.add_section_constraint(6)
>>> c.solve()
Unsatisfiable :(
>>>
```

Figure 2: Evaluating a combination of sections to make the schedule unsatisfiable

Adding Section Constraints

Figure 2 adds the first section at 7:00, and then adds the sixth section which is also at 7:00. This is a contradiction, as we cannot take two classes at the same time, and the class scheduler appropriately informs the user that this combination of classes is unsatisfiable.

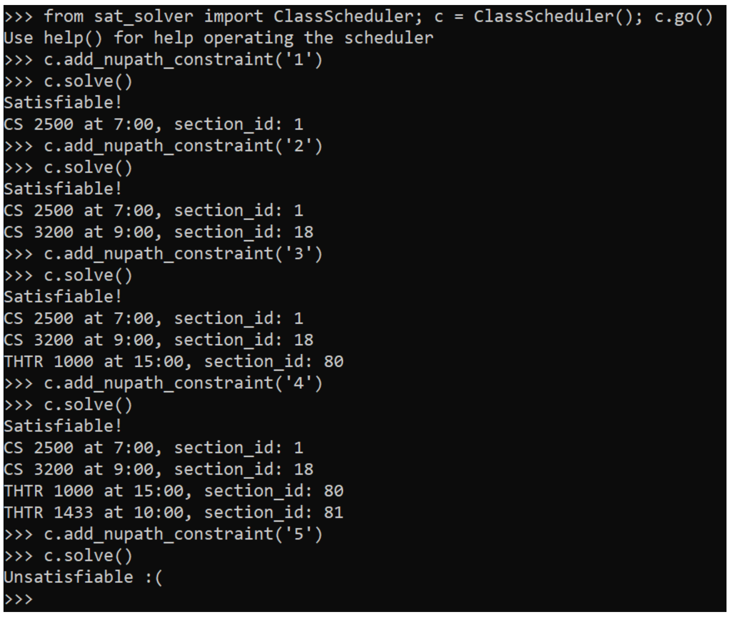


```
>>> from sat_solver import ClassScheduler; c = ClassScheduler(); c.go()
Use help() for help operating the scheduler
>>> c.add_subject_constraint('CS')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
>>> c.add_subject_constraint('PHIL')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
PHIL 1111 at 15:00, section_id: 45
>>> c.add_subject_constraint('MATH')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
MATH 3400 at 13:00, section_id: 39
PHIL 2303 at 14:00, section_id: 56
>>> c.add_subject_constraint('ECON')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
MATH 3400 at 13:00, section_id: 39
PHIL 2303 at 14:00, section_id: 56
ECON 1000 at 9:00, section_id: 60
>>> c.add_subject_constraint('THTR')
>>> c.solve()
Unsatisfiable :(
>>>
```

Figure 3: Evaluating combinations of subjects to make the schedule unsatisfiable

Adding Subject Constraints

Figure 3 evaluates adding in constraints for each subject. The Scheduler finds a minimum satisfiable section to fit from that subject. As we see trying to add in a 5th subject would not work as you can't take 5 classes.

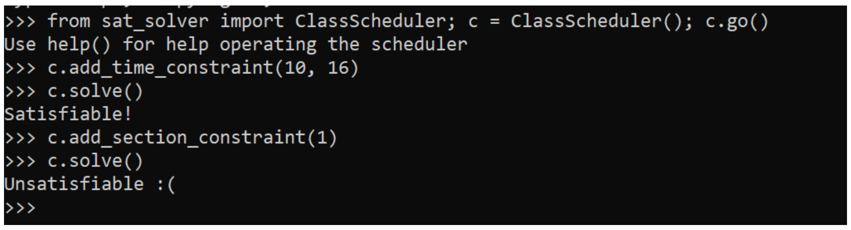


```
>>> from sat_solver import ClassScheduler; c = ClassScheduler(); c.go()
Use help() for help operating the scheduler
>>> c.add_nupath_constraint('1')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
>>> c.add_nupath_constraint('2')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
CS 3200 at 9:00, section_id: 18
>>> c.add_nupath_constraint('3')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
CS 3200 at 9:00, section_id: 18
THTR 1000 at 15:00, section_id: 80
>>> c.add_nupath_constraint('4')
>>> c.solve()
Satisfiable!
CS 2500 at 7:00, section_id: 1
CS 3200 at 9:00, section_id: 18
THTR 1000 at 15:00, section_id: 80
THTR 1433 at 10:00, section_id: 81
>>> c.add_nupath_constraint('5')
>>> c.solve()
Unsatisfiable :(
>>>
```

Figure 4: Evaluating nupath constraints added to the schedule

Adding NUPath Constraints

Figure 4 adds various NUPath constraints. Here adding a 5th NUPath constraint causes an error, as a student can't take 5 classes.

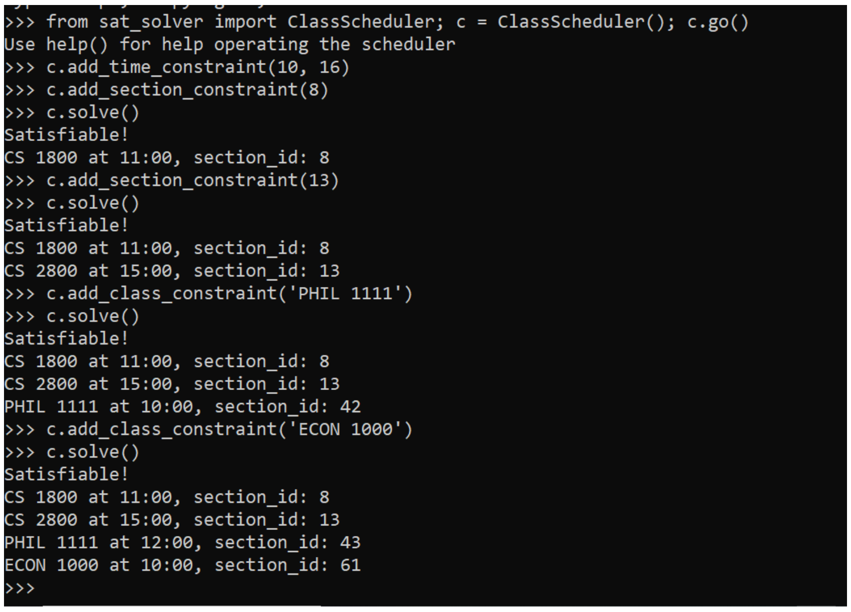


```
>>> from sat_solver import ClassScheduler; c = ClassScheduler(); c.go()
Use help() for help operating the scheduler
>>> c.add_time_constraint(10, 16)
>>> c.solve()
Satisfiable!
>>> c.add_section_constraint(1)
>>> c.solve()
Unsatisfiable :(
>>>
```

Figure 5: Evaluating adding section not within the time constraints

Adding Time Constraints

In Figure 5 we evalutate time constraints, as this student wants to only take classes between 10:00 and 16:00. We try selecting section number 1 for them, which is a 7:00 class, and doesn't fit in their time constraint. Therefore this schedule is unsatisfiable.

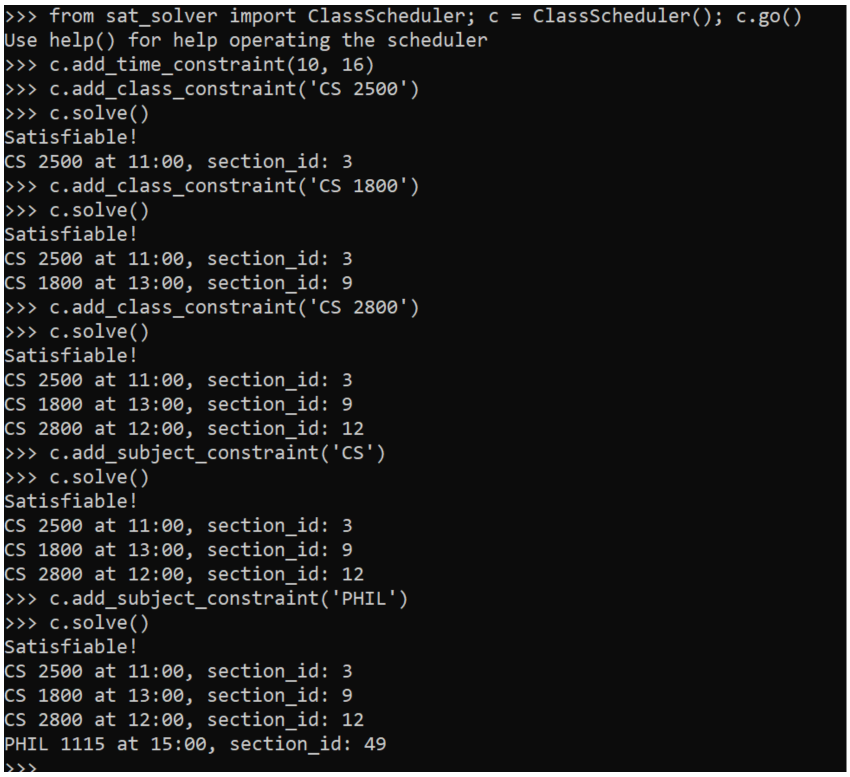


```
>>> from sat_solver import ClassScheduler; c = ClassScheduler(); c.go()
Use help() for help operating the scheduler
>>> c.add_time_constraint(10, 16)
>>> c.add_section_constraint(8)
>>> c.solve()
Satisfiable!
CS 1800 at 11:00, section_id: 8
>>> c.add_section_constraint(13)
>>> c.solve()
Satisfiable!
CS 1800 at 11:00, section_id: 8
CS 2800 at 15:00, section_id: 13
>>> c.add_class_constraint('PHIL 1111')
>>> c.solve()
Satisfiable!
CS 1800 at 11:00, section_id: 8
CS 2800 at 15:00, section_id: 13
PHIL 1111 at 10:00, section_id: 42
>>> c.add_class_constraint('ECON 1000')
>>> c.solve()
Satisfiable!
CS 1800 at 11:00, section_id: 8
CS 2800 at 15:00, section_id: 13
PHIL 1111 at 12:00, section_id: 43
ECON 1000 at 10:00, section_id: 61
>>>
```

Figure 6: Evaluating adding time and class constraints

Adding Time and Class Constraints

In Figure 6 we evaluate a full schedule of a student with strict time constraints. The desired schedule times are from 10:00 to 16:00. We add the sections 8 and 13. Then we want to take two more classes, so we choose those class constraints. The scheduler chooses sections so as to not interfere with the prior sections 8 and 13.

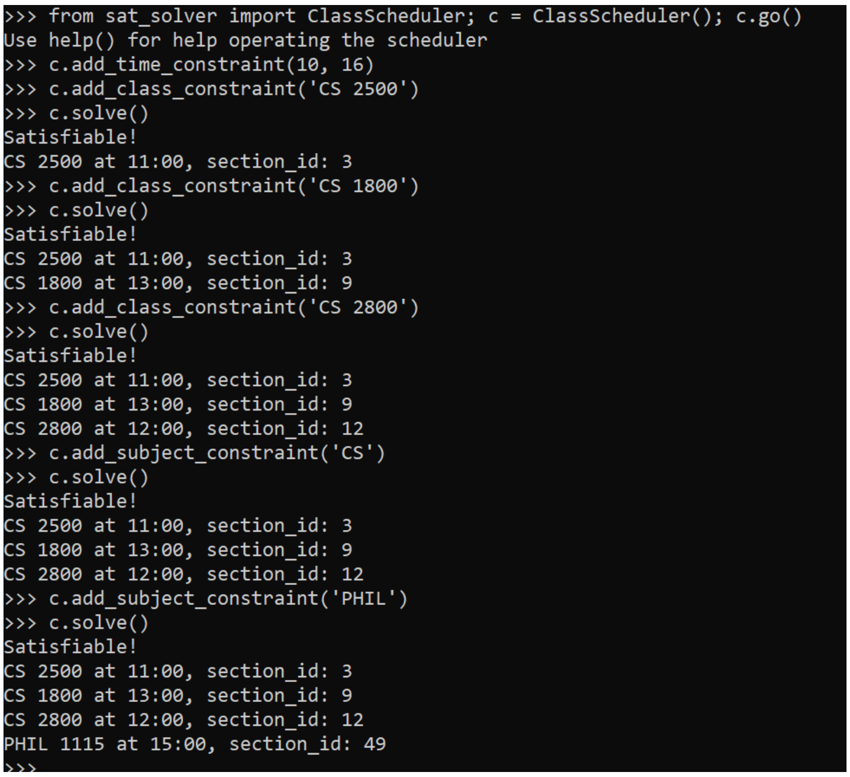


```
>>> from sat_solver import ClassScheduler; c = ClassScheduler(); c.go()
Use help() for help operating the scheduler
>>> c.add_time_constraint(10, 16)
>>> c.add_class_constraint('CS 2500')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
>>> c.add_class_constraint('CS 1800')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
CS 1800 at 13:00, section_id: 9
>>> c.add_class_constraint('CS 2800')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
CS 1800 at 13:00, section_id: 9
CS 2800 at 12:00, section_id: 12
>>> c.add_subject_constraint('CS')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
CS 1800 at 13:00, section_id: 9
CS 2800 at 12:00, section_id: 12
>>> c.add_subject_constraint('PHIL')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
CS 1800 at 13:00, section_id: 9
CS 2800 at 12:00, section_id: 12
PHIL 1115 at 15:00, section_id: 49
>>>
```

Figure 7: Evaluating adding time and subject constraints

Adding Time and Subject Constraints

In Figure 7 we set the same time constraints of 10:00 to 16:00, and add 3 class constraints. We then add Philosophy as a subject to find a section of philosophy to fit within those time constraints and it successfully chooses a generic philosophy class this student could fit into their schedule.



```
>>> from sat_solver import ClassScheduler; c = ClassScheduler(); c.go()
Use help() for help operating the scheduler
>>> c.add_time_constraint(10, 16)
>>> c.add_class_constraint('CS 2500')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
>>> c.add_class_constraint('CS 1800')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
CS 1800 at 13:00, section_id: 9
>>> c.add_class_constraint('CS 2800')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
CS 1800 at 13:00, section_id: 9
CS 2800 at 12:00, section_id: 12
>>> c.add_subject_constraint('CS')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
CS 1800 at 13:00, section_id: 9
CS 2800 at 12:00, section_id: 12
>>> c.add_subject_constraint('PHIL')
>>> c.solve()
Satisfiable!
CS 2500 at 11:00, section_id: 3
CS 1800 at 13:00, section_id: 9
CS 2800 at 12:00, section_id: 12
PHIL 1115 at 15:00, section_id: 49
>>>
```

Figure 8: Evaluating not enough time constraint

Adding Limited Time Constraint

In Figure 8 we show how a schedule with only 3 hours of time constraints is unsatisfiable when choosing 4 classes, as we could not choose 4 sections without a time overlap.

Generalization of Project

The class scheduler can be generalized really well. We could feed in a real schedule, such as Northeastern's Spring 2021 schedule, to find a satisfiable schedule of sections we could take, based on which classes we need, which time constraints we have, or even which subject we might need another class in.

Lingering Thoughts

There are some other factors we would have liked to consider for the Class Scheduler. To be more helpful for upper class students, we could add a way to exclude certain classes from being selected when adding in a subject constraint. We wouldn't want the Scheduler to tell a student to take a CS class they've already taken, so it would be useful for the sat solver to consider already taken classes. These are things that would have required expanding our cardinality constraints, which would have been tough given the time frame of our project. Another Aspect of this project we wanted to expand upon was to consider the optimal schedule, not just a minimum satisfiable schedule. There are multiple satisfiable schedules based on class constraints, subject constraints, etc. so it would be amazing to expand the scope of our project to consider other factors or even allow for the user to select an optimal schedule out of the satisfiable schedules. We could have considered closeness of classes, and some favorability rankings to determine an optimal schedule out of a series of satisfiable schedules.

Summary

As we can see from the provided examples, the class scheduler provides a solid framework for determining a valid schedule for a student considering some constraints. The scheduler has proven to fulfill many uses for students. Some of these include needing to find a last section to fill in their schedule to fit within a certain time constraint, a section to fit within a certain subject constraint, or needing to take classes but not knowing which times to take them. The problem of determining the optimal schedule out of a series of satisfiable schedules is an interesting problem that was made easier by the completion of this satisfiability problem. We now understand that it is feasible to use a SAT solver to approach the scheduling problem. Future work could center around adding constraints that are focused on optimization of the schedule.

References

1

Ignatiev, A., Morgado, A., amp; Marques-Silva, J. (2018). SAT technology in Python¶. Retrieved November 02, 2020, from https://pysathq.github.io/index.html

2

Reasoning, A. (2020, September 03). Lecture 08-2 Encoding cardinality constraints. Retrieved November 02, 2020, from https://www.youtube.com/watch?v=PKbAICDB9tM