Using Boolean Satisfiability Solvers to Help Reduce Cognitive Load and Improve Decision Making when Creating Common Academic Schedules

Joshua C. Manzano
De La Salle University
Manila, Philippines
josh_manzano@dlsu.edu.ph

Adrienne Francesca O. Soliven
De La Salle University
Manila, Philippines
adrienne_soliven@dlsu.edu.ph

Antonio Miguel B. Llamas
De La Salle University
Manila, Philippines
antonio_llamas@dlsu.edu.ph

Shenn Margareth V. Tinsay
De La Salle University
Manila, Philippines
shenn_tinsay@dlsu.edu.ph

Briane Paul V. Samson
De La Salle University
Manila, Philippines
briane.samson@dlsu.edu.ph

Rafael A. Cabredo
De La Salle University
Manila, Philippines
rafael.cabredo@dlsu.edu.ph

ABSTRACT

Manual schedule creation often involves satisfying numerous unique and conflicting constraints, which becomes more cognitively demanding when creating a common academic schedule with other individuals. Poor decision making caused by cognitive overload can result in unsuitable schedules. This study proposes the use of Boolean satisfiability (SAT) solvers in an academic scheduling system to help students balance scheduling preferences and satisfy necessary constraints. Based on the availability of courses and the scheduling preferences of users, the system automatically resolves conflicts and presents possible schedules. In a controlled experiment with 42 undergraduate students, cognitive demand was reduced by eliminating menial decisions, which significantly optimized the creation of a common schedule among peers. We found that human errors and emotional stress were diminished, and schedules created using the system were more satisfactory to participants. Finally, we present recommendations and design implications for future academic scheduling systems.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; *Empirical studies in HCI*; Usability testing.

KEYWORDS

Constraint Satisfaction; Decision Making; Student Scheduling; Human-Centered Design; Boolean Satisfiability; Cognitive Load; SAT Solver

ACM Reference Format:
Joshua C. Manzano, Adrienne Francesca O. Soliven, Antonio Miguel B. Llamas, Shenn Margareth V. Tinsay, Briane Paul V. Samson, and Rafael A. Cabredo. 2021. Using Boolean Satisfiability Solvers to Help Reduce Cognitive Load and Improve Decision Making when Creating Common Academic Schedules. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3411764.3445681

CHI '21, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445681

1 INTRODUCTION

Many universities employ an approach to academic scheduling wherein students choose classes to enlist in for a given semester [9, 15]. Under this approach, students are effectively responsible for determining their own schedules. Scheduling in general tends to be a complex task, meaning that it primarily necessitates conscious, rather than intuitive, mental effort [3, 8]. Here, complexity arises due to the frequent need to satisfy diverse and conflicting constraints when making scheduling decisions [5]. A university may offer numerous classes for a single course that differ in terms of assigned faculty members, days, times, and classrooms [6]. When deciding which classes to include in their schedules, students must account for various constraints related to considerations that may be external (e.g., availability of slots for a specific class) or personal (e.g., preferences for specific classes based on certain attributes) in nature. Adding to this complexity, the scheduling process also often entails creating and deliberating over different possible schedule configurations in order to arrive at an optimal final schedule [16]. In cognitive load theory, the act of decision making requires more deliberative processing of information and consequently, more substantial cognitive resources [3, 11]. High cognitive load in scheduling, then, can be linked to the multitude of constraints that must be satisfied. Constraint satisfaction entails the need to make considered decisions regarding the most feasible time slots in which to assign different activities based on available information. Research further suggests that in more complex tasks, the quality of a final decision is more likely to be stifled by higher cognitive load [3, 17]. This may be attributed to the cognitive resources needed for decision making being consumed by auxiliary tasks [11].

We consider how a scheduling system might be designed to assist individuals in performing cognitive resource-heavy tasks involved in scheduling. Research that looked at scheduling in individual universities showed that manual approaches to scheduling, wherein students consult relevant course information then create schedules by writing or typing relevant details manually, remained popular

[34]. As manual schedule creation leaves students primarily reliant on their own mental faculties to accomplish relevant tasks such as evaluating and comparing possible schedule configurations, it retains the implicitly high cognitive load entailed in scheduling as a task. Alternative approaches to scheduling might therefore be developed under the guiding principle of lightening the cognitive load of the entire scheduling process. Our research explores using automation to augment decision making in scheduling, reducing the cognitive effort required to make decisions when creating schedules. We specifically focus on the incorporation of satisfiability (SAT) solver technology into a scheduling system design. Previous studies have used SAT solvers in designing systems in various fields to address constraint satisfaction [2, 20, 35]. A SAT solver can automate the process of creating schedules that are configured according to specified constraints, effectively reducing cognitive load by handling the cognitive resource-heavy task of deciding how to optimally map out activities in schedules. We design this system with the intent of supporting both individual and collaborative scheduling.

Our research activities were primarily centered around a local university that employs the academic scheduling approach outlined earlier. Undergraduate students from this university were surveyed and interviewed in order to identify common considerations of students in academic scheduling. We contemplated how such considerations could be reflected in a scheduling system design in order to alleviate specific aspects of manual scheduling perceived to be tedious. We then outlined a design for a semi-automated scheduling system that utilizes a SAT solver to create schedules based on specified constraints. We implemented this system in a web application for academic scheduling targeted for use by undergraduate students. Students were invited to participate in usability testing of the application, in which they were asked to perform several tasks with the assistance of the application and provide feedback on their experience. Data from testing was used to gauge the effectiveness and usability of the system among its intended end-user base. Throughout this paper, we outline insights obtained from our research activities and note how these findings informed our system design process. This work contributes a framing of SAT solver technology that highlights its potential application in academic scheduling, especially in domains wherein students are directly responsible for constructing their own schedules; and a detailed analysis, supported by empirical findings, of the positive attributes of semi-automated schedule creation in comparison to existing manual approaches. We close by discussing recommendations, design implications, and further opportunities for researchers exploring optimization of the scheduling process.

2 RELATED WORK

Our research explores how a satisfiability (SAT) solver might be incorporated into a scheduling system in order to facilitate decision making, thereby minimizing the cognitive load of scheduling. We describe the impact of constraint satisfaction on the decision making process with respect to cognitive load levels. This is followed by a discussion of the influence of cognitive load on human decision making. We also discuss insights from previous works on SAT solver application and scheduling interface design in various domains that informed our system design process.

2.1 Constraint Satisfaction in Scheduling

Efficient scheduling entails evaluating the feasibility and favorability of assigning activities to specific time slots and days. This evaluation is largely dependent on the constraints defined in a specific scheduling domain. As such, scheduling has been defined as a problem concerning the satisfaction of a series of constraints, which may be either hard or soft [1, 5]. Hard constraints must be satisfied, whereas soft constraints, under which constraints for preferences generally fall, are considered implicitly when creating a schedule but may be violated [5]. Bharadwaj et al. [5] attributes the complexity of scheduling to the presence of varied and often conflicting hard and soft constraints.

In a university setting, the formulation of student schedules for an academic term or semester, also referred to as timetabling [28], is persistently dictated by multiple interlocking constraints. In many universities, students have the responsibility of designing their own respective schedules [9, 15]. Employing this approach, a university first plans classes to be offered for a term – which generally entails multiple classes being planned for an individual course – that are then each assigned to particular days and times, as well as classrooms and faculty members [6]. Students are then tasked with selecting specific classes to enlist in given the fixed class timetable produced by the university; this is referred to as student scheduling [28]. Student scheduling necessitates satisfying various constraints in order to produce a schedule that might be deemed optimal by the scheduling student [28, 31]. In this domain, hard constraints reflect inviolable conditions such as not being able to enlist in classes without open enlistment slots, or whose inclusion in a schedule might result in a time conflict with another class currently in the schedule [12, 21]. Soft constraints reflect preferences for specific classes based on certain attributes (e.g., preferred days and times), and as such vary on a case-by-case basis. Cheng et al. [9] associates difficulties in performing student scheduling in a university with the large number of constraints typically encountered in a university scheduling domain. These were especially apparent in universities with dense student populations, given the increased difficulty of resolving hard constraints (owing to the limited number of enlistment slots for individual classes).

It is common for students to use a more hands-on approach to creating academic schedules in preparation for enlistment [34]. A student may first consult available information regarding class offerings, evaluating relevant considerations and afterwards deciding a feasible arrangement of classes that best fits personal preferences [34]. The student may then produce a hand-drawn or digital visual representation of a created schedule for future reference [34]. This manual approach, however, possesses its limitations, particularly as concerns entailed mental demand. When determining how to map out specific classes in a schedule, each decision must be carefully evaluated according to numerous hard and soft constraints in order to ensure its feasibility. This mental demand is amplified when one considers the potential need to create numerous differing schedule configurations in order to evaluate a final configuration that might be deemed optimal [16]. Acknowledging the prospect of improving on manual approaches to schedule creation, Strazzarino and Henry [34] proposed automation as an effective means of addressing some of the aforementioned limitations. They subsequently developed and deployed a tool for a local university that extracted online class offering data, which was used to automatically create custom schedules for students.

Noting mental demand as a significant contributor to the difficulty of scheduling, we suggest that optimization of the scheduling process might focus on the principle of reducing cognitive load. This might be done by targeting aspects of scheduling that can be identified as particularly demanding of mental effort on the part of the doer, such as constraint satisfaction and by extension decision making.

2.2 Cognitive Load in Decision Making Processes

In the context of human information processing, cognitive load, or cognitive business, refers to the information held in human working memory at any given time [23]. The level of cognitive load entailed in a particular task is directly correlated to the amount of working memory resources required to perform the task [23]. Research has looked into the impact of cognitive load level on the outcomes of decision making processes in particular.

Allen et al. [3] performed an experiment that aimed to observe the effect of cognitive load manipulation on the ability of participants to accurately derive uncertainty information from presented graphs and make decisions about optimal behaviors based on this information. Cognitive load was increased for half of the participants in the form of an additional instruction to remember a specified eight-digit number for the duration of the experiment. Results indicated that while participants tended to perform consistently well on graph interpretation and mean and probability estimation questions, where information needed to ascertain answers was "overtly display[ed]" in the graphs "in an easy-to-read fashion", cognitive load impacted performance in behavioral choice questions, "presumably because of the need for more effortful attention and deliberative processes necessary to make [judgments]" [3]. These results supported the supposition that in a task involving decision making, the complexity of the task determines the extent to which decision making quality is hampered by high cognitive load [3]. Complexity here may be described as the extent to which a task requires conscious effort and deliberation on the part of the doer [3]. Chandra and Ghosh [8] posited that complex tasks require greater use of controlled or conscious knowledge to supplement automated or unconscious knowledge.

Desmarais [11] specifies a distinction between automatic processes, which "occur outside of awareness, are effortless, independent of other processes, and involuntary" and are thus unaffected by cognitive load, and controlled processes, which "require significant cognitive resources, and conscious effort and/or thought". The impact of high cognitive load on a controlled decision making process thus becomes apparent because here, decision making cannot be done efficiently in an instinctive and unconscious manner, instead requiring substantial cognitive resources of its own. The quality of a final decision may thus be stifled by cognitive overload [3, 17]. This may be attributed to the cognitive resources needed for decision making being consumed by external tasks [11]. Ferrari and Dovidio [13] linked cognitive overload to a decreased ability to process relevant information needed to make a decision, suggesting that the range of information that can be processed in a decision making task is limited by the availability of cognitive resources.

Considering the potential for cognitive overload to hinder decision making, optimization of a controlled decision making process may be approached under the principle of maximizing cognitive resources. Mentally-demanding subtasks present in a decision making domain might thus be specifically targeted in order to free up needed cognitive resources. This perspective informed our conceptualization and subsequent design of a student scheduling system targeted at automating the specific subtask of constraint satisfaction.

2.3 Boolean Satisfiability Solvers

A solver is a general-purpose algorithm that is applied to automatically search for solutions to a given problem. A solver approach may be utilized to avoid having to develop new algorithmic solutions from scratch for individual applications where intelligent search is needed, while providing high performance [7]. A satisfiability (SAT) solver is a solver that, given a list of clauses, either finds a model that can satisfy all of the clauses or reports unsatisfiability. SAT solvers are capable of solving hard-structured problems with over a million variables and several million constraints. They provide a minimalistic black box approach wherein no external tuning is needed to determine a solution to the Boolean satisfiability problem [7]. This problem, also referred to as the propositional satisfiability problem and the SAT problem, is a constraint satisfaction problem that entails determining whether there exists an assignment of variables that can satisfy a given set of constraints [4]. It is known to be NP-complete and is of tremendous importance in computer science [37].

SAT solver usage is a popular approach to addressing practical problems that may be encoded into SAT [2]. Despite their potential exponential run time, SAT solvers are increasingly being considered as general-purpose tools in multiple domains, including several involving scheduling [14]. Achá and Nieuwenhuis [2] and Strichman [35] implemented solutions to curriculum-based course timetabling and course scheduling using propositional SAT solvers. Demirović et al. [10] applied SAT solvers to employee scheduling, while Horbach et al. [20] applied them to tournament scheduling for sports leagues.

We expand on previous works by designing a semi-automated system for student scheduling that uses SAT solvers, which automate the menial tasks involved when creating a personal academic schedule. In tackling this domain, we consider how a SAT solver implementation might be undertaken outside the scope of previous implementations, so as to facilitate a scheduling process that is more directly tailored to considerations of individual users. In contrast to previous academic scheduling studies that have focused on creating universal schedules affecting a large group by balancing specifications of multiple individuals (e.g., course scheduling according to faculty preferences, as in Strichman [35]), or according to defined specifications as opposed to human preferences (e.g., curriculum-based course timetabling, as in Achá and Nieuwenhuis

[2]), this work aims to facilitate scheduling of a more personalized nature. The designed system is intended to be used by students on an individual basis in order to create schedules based on personal considerations, or when collaborative scheduling is concerned, considerations of immediate friend groups. This expanded focus on the scheduling experience with regard to the individual extends to the overarching aim of making scheduling a less cognitively-demanding process. As such, this work also contributes a more overtly human-centered system evaluation approach. Performance of the developed system is evaluated in terms of task completion rates and measured user satisfaction, cognitive load, and stress.

2.4 Scheduling Interfaces

Much work on the development of scheduling tools has focused on optimizing the scheduling process in terms of time and ease of completion through particular approaches. Such tools have been developed under the principle of improving on manual scheduling methods by providing more convenient means of accounting for intricate considerations such as user preferences and calendar context [22, 24], as well as facilitating collaborative scheduling [24, 36] and schedule visualization [36].

Several studies have specifically looked into integrating automation into interfaces for scheduling systems. Kim et al. [24] designed a scheduling system that attempts, through use of a deep neutral model, to learn user preferences and understand raw calendar contexts that include natural language, so as to plan optimal schedules for one or more users more intuitively. In a groupware calendar system developed by Tullio et al. [36], modules for intelligently identifying events common to calendars of multiple individuals and predicting event attendance, along with an interface for visualizing the outputs of these modules, were implemented in order to facilitate more efficient collaborative event planning. Huh et al. [22] developed an automated interactive system intended for use by case managers in clinics in order to schedule appointments, accounting for preferences of patients and health care providers alike. Of particular note was the emphasis of the study on evaluating the degree to which user control might be maintained in a system interface so as to optimize user experience. Three interfaces with varying degrees of user control in manipulating schedules were designed and subsequently tested with users [22]. Findings ultimately pointed to the favorability of maintaining a balance between automated and manual functionalities; the proponents stressed the importance of providing a user of a system with detailed information on how the system arrives at automated decisions and factors in specifications manually entered by the user [22]. On user—AI collaboration, Oh et al. [26] similarly advises that user experience might be improved throughout interactive processes by providing users with adequate control over decision making, as well as detailed instructions.

The above-mentioned findings on the significance of according users control over decision making suggest benefits to implementing a semi-automated interface in designing a scheduling system, as has been done in this work. This significance could be posited to be more evident in domains such as student scheduling, where users, in this case students, are directly concerned with creating their own schedules. This can be contrasted against scheduling systems whose intended users are administrators or concerned third parties tasked with creating schedules on behalf of others.

3 NEEDFINDING ACTIVITIES FOR UNDERSTANDING THE SCHEDULE CREATION PROCESS

We began by disseminating a survey to undergraduate students to establish common procedures, priorities, and frustrations experienced during the schedule creation process. Afterwards, we conducted semi-structured interviews with select participants in order to derive additional insights and formalize initial findings. Contextual inquiry was later conducted to conceptualize potential improvements to existing approaches to schedule creation.

3.1 Methodology

We administered an online needfinding survey for undergraduate students from a local university. The survey was entirely voluntary and was distributed to students from different disciplines. Responses were collected from 54 students (30 female, 24 male). The base of respondents comprised 14 first-year students, 8 second-year students, 16 third-year students, and 16 fourth-year students. This data was consolidated into affinity diagrams and subsequently analyzed.

Out of the 54 needfinding survey respondents, 16 were chosen to participate in a semi-structured, in-depth interview. All chosen participants agreed to be interviewed, and they all completed their respective interviews. The participants comprised regular (N=10), delayed (N=4), and shiftee students (N=2). They were chosen among the needfinding survey participants because they responded to questions posed in the survey with answers that we deemed unique or unexpected. Interviewees were asked to expound on their perspectives regarding certain points initially raised in the survey. This facilitated a more in-depth understanding of the experiences of students with schedule creation.

Contextual inquiry was conducted with a new set of 5 undergraduate students from the same university as participants, all of whom had not been previously surveyed or interviewed. We observed participants as they demonstrated their typical approaches to schedule creation. The effectiveness of these approaches was evaluated based on observation and inquiry of participants. Participants were asked about their sentiments regarding their utilized scheduling approaches, including their general satisfaction with these approaches and possible areas in which these approaches could be made more efficient.

3.2 Findings

Based on analysis of needfinding data, six primary categories of student scheduling preferences—time, day, human relations, proximity, course priority, and workload—were identified, along with their corresponding subcategories, as shown in Table 1. The preferences reflected in the eventual design of the proposed scheduling system encompassed these categories. Common approaches to and experiences with student scheduling were also identified.

3.2.1 Procedures and Utilized Tools. All participants in needfinding activities noted that they first consult their respective academic flowcharts to determine courses that they need to take in a given

Table 1: Preference taxonomy schema.

Depth 1	**Depth 2**	**Depth 3**
Time preference	Break preferences	Long breaks
		Short breaks
	Distribution preference	Compressed
		Spread out
	Start time preference	Early start time
		Late start time
	End time preference	Early end time
Day preference	Limit number of days	
	Free day	
Human relations preference	Faculty preference	Professors to take
		Professors to avoid
	Friend preference	
Proximity preference	Nearby rooms	
	Far away rooms	
Course priority preference	"Must have" courses	
	“Nice to have" courses	
Workload preference	Number of daily courses	
	Balancing the spread of heavy and light courses	

semester, then view online details on available class offerings for these courses. From here, participants noted different tools that they respectively use to facilitate subsequent schedule creation. Survey respondents indicated that they used online schedule creation websites[1] (69%), spreadsheet software[2] (13%), and note-taking applications (2%). 31% of respondents indicated that they created schedules without the assistance of any tools.

Participants perceived the process of schedule creation to be potentially stressful due to several "worst-case" considerations. They considered long-term impact, noting that failure to enlist in certain courses could lead to delayed graduation. The impact of schedules on quality of life was also noted, with participants emphasizing the stress of adhering to an unsatisfactory routine for the entire duration of a semester. Considering these scenarios, participants noted that they create backup schedules if circumstances do not permit desired schedules to be followed, as enlistment tends to be time-sensitive.

3.2.2 Perceived Redundancy and Inefficiency. Participants identified several menial tasks entailed in manual schedule creation, including sorting through all offered classes; evaluating which classes to select; mentally resolving potential time conflicts between classes; and manually inputting classes into a visual schedule. These tasks were generally described as tedious given their repetitive and time-consuming nature, which is compounded when creating multiple possible schedules.

[1]https://classup.plokia.com/, https://freecollegeschedulemaker.com/, https://schedninja.com/
[2]Microsoft Excel, Google Spreadsheets

3.2.3 Perceived Information Overload. Students must factor in various considerations in determining a particular schedule configuration that they might deem most satisfactory. Participants noted that when comparing offered classes for a particular course, they consider time slots and professors of individual classes. They also evaluate the spread of class workload within a particular day or week. It was noted that participants often found themselves shifting back and forth between pages displaying information for specific class offerings in order to facilitate the aforementioned comparison. Instantaneous processing of large amounts of information may be overwhelming, and viewing multiple schedules at once may be confusing.

Some participants made use of color labels to determine a hierarchy of possible scheduling choices. They marked individual class offerings with specific colors to denote their perceived favorability (i.e., using different color labels for "good" classes, "acceptable" classes, and classes that should be avoided). However, this strategy becomes less effective when dealing with numerous classes that are all perceived to be similarly favorable to add to a schedule. Color differentiation is not particularly meaningful in helping a student decide between two classes of similar priority.

3.2.4 Proneness to Errors. Common problems, such as pain points and workarounds, were observed, and suggestions were solicited as participants accomplished schedule creation tasks over the course of contextual inquiry. We asked the participants to naturally perform their respective schedule creation methods using their preferred tools, if they had any. Many participants created schedules without realizing the presence of conflicts between classes. This problem was more apparent among participants who rendered their schedules as pure text as opposed to those who rendered them visually, indicating that visual cues helped to make conflicts more readily identifiable. Misspellings of course codes were also observed to have been committed and overlooked by participants.

4 SYSTEM DESIGN

The scheduling system covered features for schedule creation and management. Features were determined according to potential usefulness for students, which was established based on needfinding data. This data encompassed insights regarding practices followed by students in the schedule creation process as well as suggestions from students on how this process could be improved. We designed eight modules that a user may interact with: view course, schedule management, view peer information, compare schedule, copy schedule, coordinate schedule, profile management, and export. Reduction of cognitive load is primarily facilitated by the schedule management module, which covers features for preference specification and schedule generation. Collaborative scheduling is facilitated by the view peer information, compare schedule, copy schedule, and coordinate schedule modules.

4.1 Preference Specification

The system enables the specification of various preferences that reflect the aforementioned primary preference categories. Preferences are specified by a user. A user may input information into as many preference fields as one deems necessary. As such, preference specification is not necessary if a user does not have preferences in any of the covered preference areas. Inputted preferences assist the system during the schedule generation process in narrowing down generated schedule variations. The feature is primarily intended to assist each user in deciding a schedule configuration that best suits personal preferences. The interfaces for preference specification and related functionalities are shown in Figure 1.

4.2 Schedule Generation

The system generates ten schedule variations based on course and section selections made by a user. During schedule generation, class conflicts are automatically resolved by the system based on defined hard and soft constraints. Hard constraints include unavailability of classes, overlapping timeslots, and taking only a single class for each course [28]. Preferences specified by a user are reflected as soft constraints. A user may set selected courses to "high" or "low" priority, enabling the system to recognize courses to prioritize for inclusion in generated schedules. A user may further indicate specific classes that one wishes to include in or exclude from generated schedules.

A user can manually edit a generated schedule. This allows a user to refine a generated schedule that might not be deemed fully satisfactory. Resulting conflicts are automatically resolved by the system, with the user being immediately notified about the presence of a conflict and given the option to either proceed with the edit and automatically resolve the conflict, or to cancel the edit.

Automatic conflict resolution and schedule generation are handled by an open-source SAT solver[3]. To interface with the SAT solver, three modules were designed and implemented: an encoding module, a SAT solver module, and a decoding module. The encoding module receives JavaScript Object Notation (JSON) inputs from the schedule management module and transforms preference data into propositional satisfiability constraints. These constraints are then processed by the SAT solver module. The SAT solver module receives conjunctive normal form (CNF)-formatted input from the encoding module and processes this input using a SAT solver. After processing, the corresponding information is passed to the decoding module, which converts the SAT solver outputs into a schedule formatted as a JSON file. It passes outputted JSON files back to the schedule management module for display.

4.3 Collaborative Scheduling

Functionalities that facilitate collaborative scheduling were implemented in the system. The system allows users to add other users as friends. The schedule comparison feature enables side-by-side comparison of a schedule saved by a user and a schedule saved by a friend of the user. For each schedule, a visual representation of the schedule is displayed alongside an informative table containing specific details about each class in the schedule. Common classes between two schedules are highlighted for easier identification.

A user may copy schedules of friends for personal reference. The implementation of this feature was spurred by the insight that some students choose to adhere to schedules of peers whom they would like to attend classes with. The feature is thus intended to alleviate the tediousness of manually reproducing an existing schedule that a user wishes to adhere to.

[3]https://github.com/Z3Prover/z3

Finally, the system facilitates schedule generation that accounts for preferences specified by the user and preferences specified by selected friends of the user, as well as courses that both parties, based on their current saved schedules, plan on taking. Three schedule variations are produced as a result of this process.

5 EVALUATION

To better understand the impact of the proposed scheduling system on schedule creation performed by students, we compared use of the system against use of existing manual scheduling methods in a controlled experiment.

5.1 Participants

During testing, participants were divided equally into 2 groups: the control group and the experimental group. Each group comprised 7 cliques (groups of friends) that performed the same sets of tasks, amounting to 14 cliques in total. The cliques were composed of actual peers that had experienced at least three trimesters of coordinating schedules with each other. Each clique consisted of 2–4 members; cliques were balanced according to demographics and their indicated amounts of experience with schedule creation. This was accounted for considering that participants were similar in several aspects, including year level, course degree, and student status[4]. For testing, a new set of participants was recruited, consisting of 42 students (18 female, 24 male), all of whom had not taken part in the prior needfinding and contextual inquiry activities. The participants are comprised of 12 second-year students, 18 third-year students, 4 fourth-year students, and 8 fifth-year students recruited from one college.

5.2 Procedure

Each test was administered remotely, considering that students typically discuss and perform schedule creation in this manner. Audio and screen activity were recorded for the duration of each test. Online chat groups were created for each clique in order to facilitate communication and observe interactions between group members. Each test lasted around 60–70 minutes.

Data on hypothetical class offerings was created for use in testing procedures. This data was based on historical course offering data made publicly available by the university that served as the testing domain. To reduce potential bias, this data was not shown to participants prior to commencement of testing. Participants were given varying lists of required courses to enlist in. The intent was to observe how they would handle this constraint when creating schedules.

The test setup comprised three main procedures. In the first procedure, participants were tasked with individually creating their schedules, only taking their own personal preferences into consideration and not communicating with any of their group members. In the second procedure, participants were tasked with coordinating with their group members to create their respective schedules; they were prompted to use their usual methods of schedule creation and communication. In the third procedure, groups were posed with a scenario in which some or all of their members are unable to

[4]Pertains to the current standing of the student as pertains to graduation, i.e., being "delayed" for graduation or otherwise "regular"

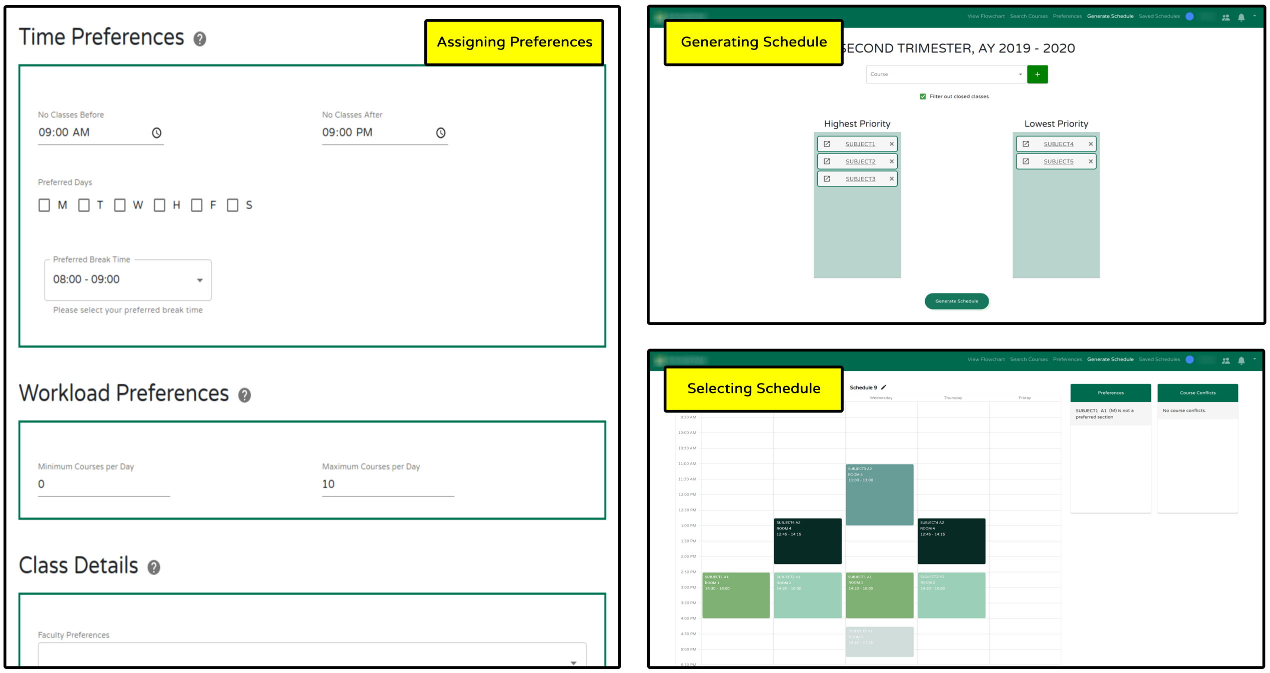

Figure 1: Interfaces for specifying preferences, adding courses to be reflected in generated schedules, and selecting preferred schedules among generated schedule variations.

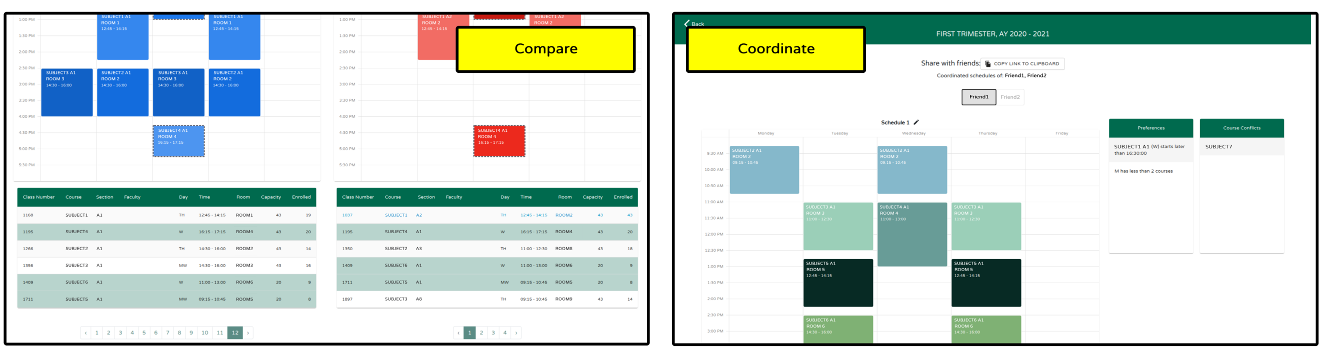

Figure 2: Interfaces for comparing schedules with a friend and collaborative schedule generation with one or more friends.

enlist in desired classes owing to sudden unavailability of slots, requiring them to adjust their schedules accordingly. Times taken by participants to complete each procedure were recorded.

The two groups were tasked with accomplishing the aforementioned three main procedures. As such, all cliques across both groups performed the same set of tasks. However, each group was asked to complete these tasks using a different tool. Cliques in the control group were asked to freely demonstrate their own respective schedule creation methods using their preferred tools, if they had any, while cliques in the experimental group were asked to use the proposed scheduling system. Before accomplishing the tasks, the experimental group was only given a general description of the system, in order to observe its intuitiveness for first-time users.

At the conclusion of testing, each participant was presented with a survey and a questionnaire to accomplish. Responses were used to assess cognitive load and stress experienced by participants while using the system. Follow-up interviews were conducted with participants at least a week after the testing session in order to evaluate quality of and satisfaction with decisions made during testing.

5.2.1 User Experience Questionnaire. The User Experience Questionnaire (UEQ) was used to measure the classical usability and the user experience of the system. It can be used to measure against competing products, measure the improvement of user experience in an iterated product, determine sufficiency in user experience, and determine areas requiring improvement. The UEQ complements evaluations that have subjective quality assessment because it provides supplementary data [25]. It encompasses scales for six categories: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty.

The UEQ also encompasses a benchmark evaluation that facilitates interpretation and evaluation of the user experience of a product. Benchmark categories are primarily applicable to business and web applications [29].

5.2.2 NASA Task Load Index. The NASA Task Load Index (NASA-TLX) assessment tool was used to gauge the amount of mental workload perceived by participants in the process of creating schedules. The NASA-TLX is administered as a questionnaire that evaluates perceived workload according to six categories: mental demand, physical demand, temporal demand, performance, effort, and frustration [18]. Each category is rated on a 10-item scale, with 1 representing "low" and 10 representing "high". The frustration metric may be used to measure stress levels in a specified domain [30].

5.2.3 Positive and Negative Affect Schedule. The Positive and Negative Affect Schedule (PANAS) questionnaire was used to gauge affects experienced by participants when creating schedules. As the PANAS evaluates the elicitation of psychological stress reactions in individuals, it has been utilized in studies concerning decision making under stress in order to establish stress induction in a specified domain [32, 33]. We specifically administered the Positive and Negative Affect Schedule — Short Form (PANAS-SF), a concise 10-item variation of the PANAS [38]. Participants were asked to rate five adjectives measuring positive affectivity and five measuring negative affectivity on a five-point scale, with 1 representing "none" and 5 representing "very much".

5.2.4 Subjective Mental Effort Questionnaire. The Subjective Mental Effort Questionnaire (SMEQ) was used to gauge the amount of mental effort perceived by participants as being necessary to create schedules. The SMEQ consists of a single 150-point scale with nine labels positioned at certain points. Each label describes a perceived level of effort to perform a task, ranging from "not at all hard to do" to "tremendously hard to do" [27].

5.2.5 Survey Questions. Four-point survey questions were also administered in order to evaluate the system in terms of perceived efficiency and satisfaction. Participants were asked to rate how much they agreed with a series of statements on a four-point scale, with 1 representing "strongly disagree" and 4 representing "strongly agree". These statements inquired regarding perceptions of performing key tasks entailed in schedule creation. Presented statements were formulated with the intent of these serving as indicators for four self-defined metrics: stress, cognitive load, ability for efficient decision making, and ability for efficient group collaboration.

5.3 Follow-Up Interviews

Participants were interviewed in order to evaluate efficiency of and satisfaction with the system, as well as quality of decisions made while using the system. We found that a week-long interim period between testing and interview dates was sufficient to ensure that participants did not recall created schedules and were no longer fatigued by the experiment, which may have potentially skewed answers. Questions were formulated using a study on meeting scheduling by Higa et al. [19] as a reference point.

Each participant was asked to answer a series of four-point survey questions. The participant was also asked to perform three comparisons corresponding to the aforementioned three main testing procedures. In each comparison, two schedules – one of which corresponded to the actual final schedule created by the participant during a particular testing procedure, though the participant was not made aware of this – were presented, and the participant was asked to indicate the more preferable option. This was done to gauge the quality of schedules created using the system. Participants were then presented with the schedules they created during testing and asked to rate each schedule based on perceived satisfaction on a scale of 1 to 4, with 1 representing "not satisfied" and 4 representing "most satisfied".

6 RESULTS

6.1 Impact on Cognitive Load

One of the primary goals of this paper is to evaluate the impact of a SAT solver-based scheduling system on the cognitive load of the schedule creation process. Figure 3 shows the difference in cognitive load experienced by participants during schedule creation. Total NASA-TLX scores indicated that participants in the experimental group (M = 17.04, σ = 5.48) perceived less cognitive workload compared to participants in the control group (M = 26.52, σ = 8.58). A Student's paired two-tailed t-test between the experimental and control groups indicates a significant decrease in the level of perceived cognitive workload for the experimental group, $t(40) = 4.26$, $p < 0.001$.

Subjective mental effort scores indicated that schedule creation necessitated less mental effort to be undertaken for participants in the experimental group (M = 17.38, σ = 8.32) compared to participants in the control group (M = 37.14, σ = 33.00). A Student's paired two-tailed t-test between the experimental and control groups showed a significant decrease in the level of necessitated mental effort for the experimental group, $t(40) = 2.66$, $p < 0.05$.

Various tasks entailed in schedule creation were found to be less effortful to undertake by the participants in the experimental group. The most significant difference in perceived levels of effort between the two groups was observed in the task of keeping track of the availability of slots for individual classes. The control group considered this task "very difficult", while the experimental group considered it "very easy". Adjusting created schedules and resolving conflicts between classes in a schedule were considered "difficult" by the control group and "easy" by the experimental group. Within the experimental group, tasks necessitating working with multiple schedules were likewise found to be less effortful to undertake. Creating different possible schedule configurations and keeping track of multiple schedules were both considered "difficult" by the control group, whereas the experimental group found these tasks to be "very easy" and "easy", respectively.

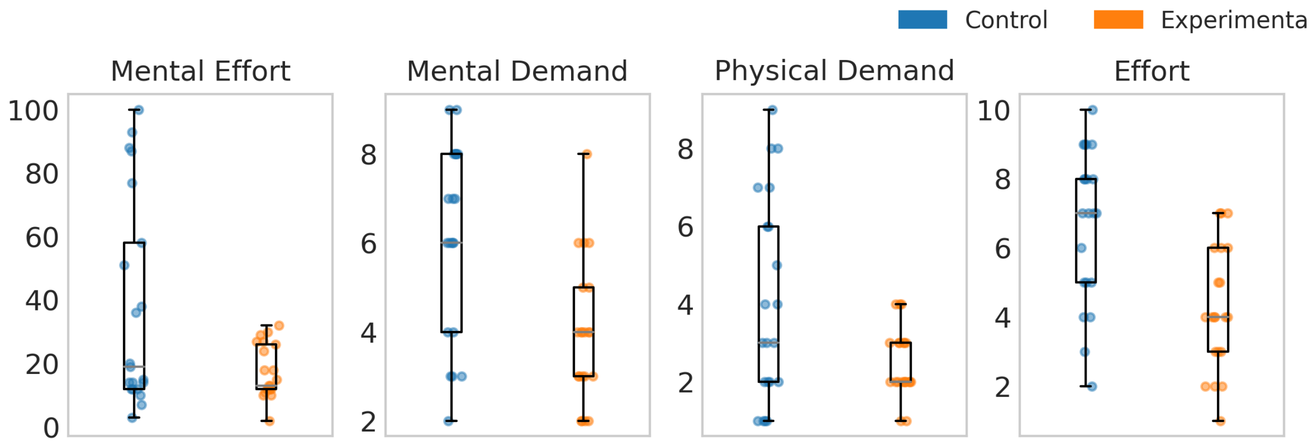

Figure 3: Cognitive load experienced by participants during schedule creation.

Both groups were of the prevailing sentiment that schedule creation did not take long to accomplish. This sentiment was noted to be more pronounced among participants in the experimental group.

6.2 Impact on Perceived Stress

As measured through PANAS and NASA-TLX scores, perceived stress was generally lower among participants in the experimental group compared to participants in the control group. Figure 4 shows the difference in perceived stress experienced by participants during schedule creation. The total PANAS negative affect scores indicated that participants in the experimental group (M = 6.85, σ = 1.79) experienced less negative affects associated with stress compared to participants in the control group (M = 10.28, σ = 3.22), $t(40) = 4.25$, $p < 0.001$. The scores in the frustration subscale of the NASA-TLX indicated that participants in the experimental group (M = 2.90, σ = 1.86) experienced less irritation, stress, and annoyance compared to participants in the control group (M = 4.90, σ = 2.68), $t(40) = 2.80$, $p < 0.05$.

A Student's paired two-tailed t-test between the experimental and control groups for both assessment tools resulted in $p < 0.001$, which indicates a significant decrease in perceived stress for the experimental group.

Relative to the control group, participants in the experimental group were noted to have found schedule creation less stressful and more efficient to undertake.

6.3 Impact on Ability for Decision Making

The amount of time taken for participants to create an academic schedule in general was significantly smaller for the experimental group (M = 40.35 min, σ = 8.04 min) compared to the control group (M = 51.77 min, σ = 15.11 min). A Student's paired two-tailed t-test between the experimental and control groups resulted in $t(40) = 3.06$, $p < 0.001$. This indicates that participants in the experimental group were able to make choices in significantly less time than participants in the control group.

It was noted that the task of deliberating over choices of similar perceived favorability, which was found to be "difficult" by the control group, was found to be "easy" by the experimental group. Furthermore, it was noted that among several decision-reliant schedule creation tasks that the control group found "easy" to perform manually, greater ease of task completion was nonetheless observed within the experimental group. These tasks included assessing the favorability of possible class selections, organizing all information necessary to create a schedule, and ultimately selecting classes, weighing schedule possibilities, and creating a final schedule.

Responses from participants in the control group to the inquiry of whether decision making (undertaken while performing the scheduling method utilized in question) could still be improved indicated a general sentiment of "agree", whereas responses to this inquiry from participants in the experimental group using the proposed system indicated a general sentiment of "strongly disagree". Participants in the experimental group were noted to have predominantly expressed that they found it easier to focus on making decisions due to the reduced need to expend effort on menial tasks such as creating multiple possible schedules and considering possible conflicts within created schedules.

Various human errors were noted to have been committed by participants in the control group. Several groups created schedules with conflicts between classes, which they were unable to identify until late in the process. Other errors included classes being missing from generated schedule visualizations and failure of participants to initially notice certain class offerings that they later indicated they would have preferred to include in their schedules.

6.4 Impact on Ability for Group Collaboration

Both the control and experimental groups were noted to have taken roughly the same amount of time to accomplish the first group task. For the second group task, the experimental group was noted to have taken a significantly shorter time to finish, as shown in Figure 5. This could be seen as an indication that the task of adjusting schedules in a group may be undertaken more efficiently using the proposed system. It was observed that within the experimental

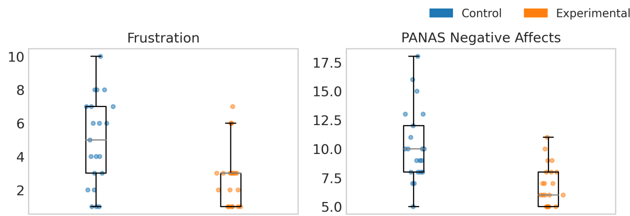

Figure 4: Perceived stress experienced by participants during schedule creation.

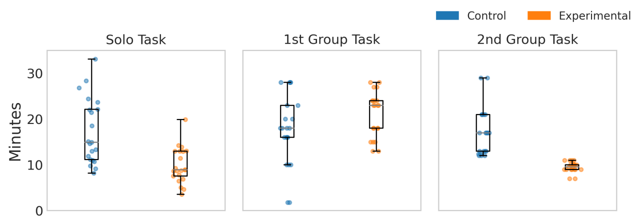

Figure 5: Times taken for participants to create a schedule individually and in a group.

Table 2: UEQ scores of the proposed system, with corresponding benchmark evaluations.

Aspect	Score	Benchmark evaluation
Attractiveness	2.11	Excellent
Perspicuity	1.43	Above average
Efficiency	1.86	Good
Dependability	1.60	Good
Stimulation	1.92	Excellent

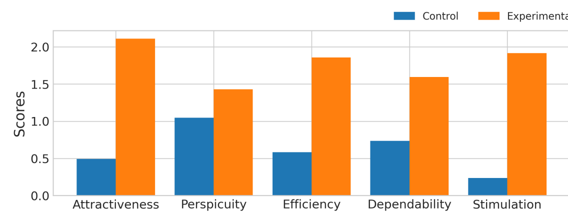

Figure 6: Scores of the proposed system in relevant classical usability metrics (efficiency, perspicuity, and dependability) and user experience metrics (attractiveness and stimulation).

group, creation of possible schedules with multiple individuals was found to have been easier to accomplish. It was also noted that participants who used voice calls to coordinate schedule creation were able to finalize a schedule faster.

Within the control group, human errors in performing group tasks were noted. Creation of schedules containing course conflicts was still a common occurrence. There were instances wherein participants would apply suggestions of peers to add certain classes to their schedules, resulting in course conflicts. Some participants intended to simply copy a schedule already created by another peer, only to fail to properly replicate this schedule.

6.5 User Experience and Usability

We also look at user experience and usability of the proposed system. UEQ results suggested that the participants found the system to be acceptable in most relevant aspects. A benchmark evaluation of the system is shown in Table 2. A Student's paired two-tailed t-test between the experimental and control groups on the aspects of attractiveness ($t(40) = 4.10$), efficiency ($t(40) = 2.71$), dependability ($t(40) = 2.45$), and stimulation ($t(40) = 4.08$), resulted in $p < 0.05$ (Figure 6). This indicates that participants in the experimental group perceived better experience with the system in these aspects than participants in the control group. In perspicuity, the system registered a benchmark evaluation of "above average"; as such, the system can be seen as lacking in this aspect based on benchmark data[5]. No significant difference was observed between the two groups in terms of perspicuity.

Regarding satisfaction with created schedules, responses from participants in the control group indicated a general sentiment of "satisfied", while responses from participants in the experimental group indicated a general sentiment of "very satisfied". A Mann–Whitney U test between the experimental and control groups resulted in $p < 0.05$, which indicates that participants in the experimental group were more satisfied with the schedules they created than participants in the control group.

[5]Since the benchmark contains data from established products, a new product should ideally reach an evaluation of at least "good" on all scales [29].

7 DISCUSSION

7.1 Reduction of Workload and Stress

Automation of menial tasks entailed in scheduling—such as creating multiple possible schedules and resolving possible conflicts within created schedules—could be deemed impactful on the reduction of cognitive load and stress when creating personalized schedules. It was noted that participants utilizing the semi-automated system perceived performing various tasks entailed in schedule creation to have been less cognitively effortful and stressful compared to participants who used existing scheduling approaches. This was particularly noted among tasks that involved dealing with multiple schedule variations, which were identified as being high in entailed effort and stress by participants who used existing approaches.

The system was also noted to alleviate physical demand entailed in schedule creation, as shown in Figure 3. Follow-up interviews indicated that the act of switching between browser tabs and applications was perceived as physically demanding to participants who utilized existing approaches. By contrast, participants who utilized the system found the schedule creation process, and the aforementioned tasks in particular, to have necessitated relatively minimal physical demand. This could suggest that users who experience high cognitive load might in turn perceive physical tasks to be more demanding.

It was noted that visualization through interface design elements such as calendar views facilitated the perceived simplification of tasks. Participants found that they comprehended schedules better when these were presented in a more visual-oriented weekly view format. Feedback indicated that users might desire a degree of liberty in designing visual components, such as when designing visual representations of schedules. Additionally, colors were found to be intuitive, allowing users to more immediately discern individual classes in schedules. Nonetheless, color choice should be thoughtfully considered in order to assure accessibility and readability of visual elements containing text.

7.2 Optimization of Decision Making

Our findings indicated that the integration of a SAT solver in a semi-automated system could be deemed effective in optimizing the creation of a common schedule among peers. This echoed the results of previous studies conducted by Achá and Nieuwenhuis [2] and Strichman [35] on SAT solver usage in a scheduling context. Use of the system was observed to pose benefits in decision making. Automation of menial tasks was noted to have helped reduce the time necessary to create and finalize schedules. Greater perceived ease of completion of decision making tasks was likewise noted among participants who utilized the system. Participants generally perceived automatic conflict resolution and subsequent generation of schedules to have been beneficial to easing decision making.

The significant difference between the times taken by the experimental group to complete the first and second group tasks could be attributed to a learning curve period, as pointed out by participants in follow-up interviews. Participants attributed the relatively longer time it took for them to complete the first group task to lack of familiarity with the system. They indicated that they used most of their time familiarizing themselves with features related to collaboration, and noted that by the second group task they were able to familiarize themselves with the system enough to be able to utilize it efficiently. Apart from this, participants noted that being able to automatically resolve conflicts made it easier for them to decide between schedules. It is interesting to note that during collaboration, some participants opted not to fully automate the creation of their schedules, instead relying on using specific features such as conflict resolution and editing of preexisting schedules.

Participants in the control group also cited other factors that they believed hampered efficiency of schedule creation and could thus be condensed. They found it more difficult to make informed decisions on which schedules to select as final schedules when these were presented in a primarily text-based manner, without corresponding visualization. The ability to visualize a schedule was thus deemed to be crucial to allowing individuals to more immediately determine how satisfactory a schedule might be. Although having more choices of classes to add was welcomed as facilitating more flexibility in schedule creation, it was noted that students may experience decision paralysis in situations necessitating the evaluation of larger numbers of choices, given the significant amount of information processing and deliberation that would be entailed.

The extensive features encompassed in the design of the system eliminate the need to perform tasks perceived as menial and redundant. Nonetheless, a seeming trade-off to this ease of use was observed, as participants expressed that they felt that they needed to allot time to more fully comprehend the system and maximize its capabilities. It was observed that after some initial time spent exploring the system, most participants were able to accomplish specified tasks with relative ease. This may be correlated to the relatively low UEQ perspicuity scores that were recorded.

7.3 Towards Automation of Schedule Creation

Although automation was perceived as generally beneficial by participants, they nonetheless expressed a preference for maintaining control over final decisions. This suggests that while automation may pose benefits in terms of facilitating faster and less strenuous decision making, fully automating the process of creating a common schedule with peers may not be deemed entirely optimal due to human factors, with users still desiring a degree of autonomy in decision making.

While a majority of the participants expressed overall satisfaction with their created schedules, results indicated that participants who utilized the SAT solver-based system were more satisfied with their schedules compared to those who did not. This perhaps suggests an increase in the quality of final decisions that may be attributed to reduction of cognitive load, as noted by Allen et al. [3] and Halstrom [17]. It was also noted that participants were more satisfied when they maintained a degree of control over the process, and indeed often utilized automated features to varying degrees to achieve this control. This indicates that when designing an automated scheduling system, it is important to consider freedom accorded to the user throughout the scheduling process. A balance between automated and manual decisions should be considered to maximize satisfaction with the process.

User experience evaluation of the SAT solver-based system indicated that users expressed hesitance with what they perceived to be the ambiguity of automated decisions performed by the system. This might imply that when automating schedules, users may feel more assuaged if they were presented with informative feedback regarding the reasoning behind generated outcomes, especially if these outcomes are sensitive to user inputs.

8 LIMITATIONS AND FUTURE WORK

In this work, three data gathering activities were undertaken. Students from different colleges (within the university that served as the testing domain) were selected as participants for needfinding and contextual inquiry in order to collect varied perspectives from students. In contrast, for the final controlled experiment, students from the same college were selected as participants to facilitate easier comparison and testing administration, with groups of participants being limited to 2–4 members. Testing with larger groups could be undertaken to verify if similar findings regarding efficiency of collaboration are upheld. Furthermore, improvements in collaboration were only measured through analysis of survey data and time on task measurement. Future work might thus explore the use of other metrics such as quantifying the amount of and difference in communication between peers.

Collaborative features in this study were mostly asynchronous in implementation. It may be noted that asynchronous collaboration lacks the responsive feedback typical of synchronous collaboration. As responsive feedback could potentially further optimize efficiency of collaboration, integration of additional synchronous collaborative features, to be used alongside a SAT solver, could be explored.

The proposed system was primarily dependent on an implemented SAT solver to facilitate features involving automated schedule generation. It allowed users to automatically resolve conflicts and generate potential schedules. However, the speed of automation was noted to decrease when users attempted to input greater amounts of preferences and constraints. Different automation algorithms such as genetic algorithms, tabu search, and other constraint satisfaction algorithms [28] can be considered in future works to benchmark the speed and efficiency of a SAT solver.

Future work tackling the implementation of automation in a scheduling system might entail undertaking testing with more complex use cases, such as when more conflicting preferences or more choices for a user to decide from are present. This can serve to gauge the extent to which the reduction of stress and cognitive load suggested in the results of this study holds true under more potentially demanding conditions. Future work could also look into applying SAT solvers or automation in general to facilitate improved decision making in use cases in other non-academic domains that may necessitate the creation of mutual schedules.

9 CONCLUSION

In this paper, we presented a design for a schedule creation system that assists users in balancing consideration of scheduling preferences with satisfaction of necessary constraints. We found that in an academic scheduling domain, the use of a Boolean satisfiability (SAT) solver was effective in reducing the cognitive load and stress of the schedule creation process, which led to beneficial effects such as reduction in time taken to accomplish schedule creation tasks, along with an increase in perceived satisfaction with decisions made. From our findings, we emphasize consideration of the use of informative feedback on automated decisions in order to reduce the likelihood of perceived ambiguity and confusion with such decisions, as well to aid in weighing possible decisions. Furthermore, we recommend preserving a balance between automated and manual decisions. This would allow users to maintain a sense of autonomy in decision making, while simultaneously reducing the effort needed to efficiently create a schedule.

ACKNOWLEDGMENTS

We would like to thank the Center for Complexity and Emerging Technologies for the extensive support that they have given us. We would also like to thank all testers and respondents, who generously provided their time to participate in our experiments for this study; as well as Jordan Deja and Unisse Chua, for offering much-appreciated feedback on this paper.

REFERENCES

[1] Slim Abdennadher and Hans Schlenker. 1999. Nurse scheduling using constraint logic programming. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence and Eleventh Conference on Innovative Applications of Artificial Intelligence*. AAAI Press / The MIT Press, Menlo Park, CA, USA, 838–843. http://www.aaai.org/Library/IAAI/1999/iaai99-118.php

[2] Roberto Asín Achá and Robert Nieuwenhuis. 2014. Curriculum-based course timetabling with SAT and MaxSAT. *Annals of Operations Research* 218, 1 (2014), 71–91.

[3] Pamela M. Allen, John A. Edwards, Frank J. Snyder, Kevin A. Makinson, and David M. Hamby. 2014. The effect of cognitive load on decision making with graphically displayed uncertainty information. *Risk Analysis* 34, 8 (2014), 1495–1505.

[4] Fadi Aloul, Bashar Al-Rawi, Anas Al-Farra, and Basel Al-Roh. 2006. Solving employee timetabling problems using Boolean satisfiability. In *2006 Innovations in Information Technology*. Institute of Electrical and Electronics Engineers, New York, NY, USA, 1–5. https://doi.org/10.1109/INNOVATIONS.2006.301886

[5] Anandhi Bharadwaj, Ajay S. Vinze, and Arun Sen. 1994. Blackboard architecture for reactive scheduling. *Expert Systems with Applications* 7, 1 (1994), 55–65.

[6] Stefan D. Bloomfield and Michael M. McSharry. 1979. Preferential course scheduling. *Interfaces* 9, 4 (1979), 24–31.

[7] Lucas Bordeaux, Youssef Hamadi, and Lintao Zhang. 2006. Propositional satisfiability and constraint programming: A comparative survey. *ACM Computing Surveys (CSUR)* 38, 4 (2006), 12.

[8] Naveen Chandra and Jayanta Kumar Ghosh. 2017. A cognitive method for building detection from high-resolution satellite images. *Current Science* 112, 5 (2017), 1038–1044. http://www.jstor.org/stable/24912499

[9] Eddie Cheng, Serge Kruk, and Marc Lipman. 2003. Flow formulations for the student scheduling problem. In *Practice and Theory of Automated Timetabling IV*, Edmund Burke and Patrick De Causmaecker (Eds.). Springer, Berlin, Germany, 299–309.

[10] Emir Demirović, Nysret Musliu, and Felix Winter. 2019. Modeling and solving staff scheduling with partial weighted MaxSAT. *Annals of Operations Research* 275, 1 (2019), 79–99.

[11] Sarah Louise Desmarais. 2005. *The effects of cognitive busyness and question wording: Responses to an intimate partner abuse vignette*. Ph.D. Dissertation. Simon Fraser University.

[12] M. A. Saleh Elmohamed, Paul Coddington, and Geoffrey Fox. 1998. A comparison of annealing techniques for academic course scheduling. In *Practice and Theory of Automated Timetabling II*, Edmund Burke and Michael Carter (Eds.). Springer, Berlin, Germany, 92–112.

[13] Joseph R. Ferrari and John F. Dovidio. 2001. Behavioral decision making strategies by indecisives. *Personality and Individual Differences* 30 (2001), 1113–1123.

[14] Carla P. Gomes, Henry Kautz, Ashish Sabharwal, and Bart Selman. 2008. Satisfiability solvers. *Foundations of Artificial Intelligence* 3 (2008), 89–134.

[15] Jennifer Gonzalez. 2012. Less choice, more structure for students: In a Tennessee system, it works. *The Chronicle of Higher Education* 58, 42 (2012), A12–A13.

[16] Donald G. Hackmann and Jerry W. Valentine. 1998. Designing an effective middle level schedule. *Middle School Journal* 29, 5 (1998), 3–13.

[17] Hannah Halstrom. 2016. *Made in America: The affects cognitive load consumer ethnocentrism and country of origin have on consumer purchasing decisions*. Honors Thesis. Union College.

[18] Sandra G. Hart. 2006. NASA-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909 arXiv:https://doi.org/10.1177/154193120605000909

[19] Kunihiko Higa, Bongsik Shin, and V. Sivakumar. 1996. Meeting scheduling: An experimental investigation. In *1996 IEEE International Conference on Systems, Man and Cybernetics: Information Intelligence and Systems*, Vol. 3. Institute of Electrical and Electronics Engineers, New York, NY, USA, 2023–2028. https://doi.org/10.1109/ICSMC.1996.565442

[20] Andrei Horbach, Thomas Bartsch, and Dirk Briskorn. 2012. Using a SAT-solver to schedule sports leagues. *Journal of Scheduling* 15, 1 (2012), 117–125.

[21] Chin-Ming Hsu and Hui-Mei Chao. 2009. A student-oriented class-course timetabling model with the capabilities of making good use of student time, saving college budgets and sharing departmental resources effectively. In *Proceedings of the 2009 WRI Global Congress on Intelligent Systems (GCIS '09, Vol. 2)*. IEEE Computer Society, Washington, DC, USA, 379–384. https://doi.org/10.1109/GCIS.2009.138

[22] Jina Huh, Martha Pollack, Hadi Katebi, Karem Sakallah, and Ned Kirsch. 2010. Incorporating user control in automated interactive scheduling systems. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems* (Aarhus, Denmark) *(DIS '10)*. Association for Computing Machinery, New York, NY, USA, 306–309. https://doi.org/10.1145/1858171.1858226

[23] Slava Kalyuga and Anne-Marie Singh. 2016. Rethinking the boundaries of cognitive load theory in complex learning. *Educational Psychology Review* 28, 4 (2016), 831–852.

[24] Donghyeon Kim, Jinhyuk Lee, Donghee Choi, Jaehoon Choi, and Jaewoo Kang. 2018. Learning user preferences and understanding calendar contexts for event scheduling. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (Torino, Italy) *(CIKM '18)*. Association for Computing Machinery, New York, NY, USA, 337–346. https://doi.org/10.1145/3269206.3271712

[25] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *HCI and Usability for Education and Work*, Andreas Holzinger (Ed.). Springer, Berlin, Germany, 63–76.

[26] Changhoon Oh, Jungwoo Song, Jinhan Choi, Seonghyeon Kim, Sungwoo Lee, and Bongwon Suh. 2018. I lead, you help but only with enough details: Understanding user experience of co-creation with artificial intelligence. In *Proceedings of the*

2018 CHI Conference on Human Factors in Computing Systems (Montreal, QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174223

[27] Jeff Sauro and Joseph S. Dumas. 2009. Comparison of three one-question, post-task usability questionnaires. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1599–1608. https://doi.org/10.1145/1518701.1518946

[28] Andrea Schaerf. 1999. A survey of automated timetabling. *Artificial Intelligence Review* 13, 2 (1999), 87–127.

[29] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Construction of a benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 4 (2017), 40–44.

[30] Alper Seker. 2014. Using outputs of NASA-TLX for building a mental workload expert system. *Gazi University Journal of Science* 27, 4 (2014), 1131–1142.

[31] Krzysztof Socha, Michael Sampels, and Max Manfrin. 2003. Ant algorithms for the university course timetabling problem with regard to the state-of-the-art. In *Applications of Evolutionary Computing*, Stefano Cagnoni, Colin G. Johnson, Juan J. Romero Cardalda, Elena Marchiori, David W. Corne, Jean-Arcady Meyer, Jens Gottlieb, Martin Middendorf, Agnès Guillot, Günther R. Raidl, and Emma Hart (Eds.). Springer, Berlin, Germany, 334–345.

[32] Katrin Starcke and Matthias Brand. 2012. Decision making under stress: A selective review. *Neuroscience & Biobehavioral Reviews* 36, 4 (2012), 1228–1248.

[33] Katrin Starcke, Oliver T. Wolf, Hans J. Markowitsch, and Matthias Brand. 2008. Anticipatory stress influences decision making under explicit risk conditions. *Behavioral Neuroscience* 122, 6 (2008), 1352.

[34] Robert Strazzarino and Tyson R. Henry. 2005. Volunteerism: A new idea for filling university information technology needs. In *Proceedings of the 33rd Annual ACM SIGUCCS Conference on User Services* (Monterey, CA, USA) *(SIGUCCS '05)*. Association for Computing Machinery, New York, NY, USA, 372–375. https://doi.org/10.1145/1099435.1099519

[35] Ofer Strichman. 2017. Near-optimal course scheduling at the Technion. *Interfaces* 47, 6 (2017), 537–554.

[36] Joe Tullio, Jeremy Goecks, Elizabeth D. Mynatt, and David H. Nguyen. 2002. Augmenting shared personal calendars. In *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology* (Paris, France) *(UIST '02)*. Association for Computing Machinery, New York, NY, USA, 11–20. https://doi.org/10.1145/571985.571988

[37] Yakir Vizel, Georg Weissenbacher, and Sharad Malik. 2015. Boolean satisfiability solvers and their applications in model checking. *Proc. IEEE* 103, 11 (2015), 2021–2035.

[38] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54, 6 (1988), 1063.