QUEST UNIVERSITY CANADA

A Summer of Implementing Gradual Improvements to Systems at Quest

Caleb Raible-Clark
Advised by Richard Hoshino
8/24/2013

An Overview of My Summer

Being an applied mathematics student, I am far more transfixed with what I did than what I attempted. With that limited scope, this brief report summarizes the three specific components of my academic fellowship that I consider to be most concrete. Appended, an interested reader can find richer discussion of most of these topics.

Algorithm and Survey Design for Roommate Matching and Condo Arrangement

Richard and I assisted the Student Affairs and Admissions teams by developing a program which used discrete optimization to suggest optimal roommate matching, based on data from Residence Life entry surveys from incoming first years. The goal of the project was to create a program which could aid in maximizing compatibility and equitability in the roommate matching process. Furthermore, we aimed to aid in overall housing organization (floor and village assignment) to enhance the whole Quest community. The largest change in survey design was the ability for incoming first year students to quantify how much they cared about various preferences (eg smoking, sleeping times), rather than simply stating those preferences.

In the end, we ended up using a "greedy" algorithm for roommate matching (see appendix E section 2.2 for a clearer explanation of this term). To assign students to floors, we simply added a function to sort for gender balance. We cannot be sure whether the system creates better matching than the previous system (I would not dare to attempt an objective measure of overall roommate happiness). We can, however, say with confidence that we were able to aid staff in completing a task in dramatically less time than previous years.

With ongoing work, continued meetings with Student Affairs and Admissions, and further research into the psychology of compatibility, there is further opportunity to develop a more robust survey (See Appendix E Section 3.1), a more optimal algorithm, and better compatibility with Microsoft SharePoint Survey functionality by next summer. We also, with members of the Student Affairs team, hope to publish our results and strategy in a location where they may be of use to other universities. (See Appendix D for the Announcement of Implementation; See appendix E for a summary of the roommate matching program)

27th Conference of the Association for the Advancement of Artificial Intelligence

Richard and I attended the 27th Conference of the Association for the Advancement of Artificial Intelligence (AAAI), in Bellevue, WA. At AAAI (the most prestigious conference on the topic), I had the opportunity to learn about a huge range of various interconnected ongoing Artificial Intelligence (including the scheduling work relevant to my fellowship), network with major leaders in the field, and bond with Richard in an off-campus setting. The experience was absolutely invaluable, and I would highly recommend that any Summer Fellow attend an Academic Conference, ideally with their advisor (See Appendix F for a blog post reflection on the conference).

Course Registration Algorithm Implementation

Encompassing the majority of my time and efforts this summer has been a project to change course registration processes at Quest. I set out at the beginning of the summer to design a program which could improve overall utility (i.e. students getting into their first choice courses each block) and equitability (i.e. more people getting closer to the mean count of number of first choice courses acquired) in Quest's course registration system. In retrospect, my intentions around the design of this program are dramatically disconnected from its actual design. That being said, the system will be an improvement over Quest's current system, and—barring any major unforeseen logistical issues—will be implemented for the 2013 October Registration Cycle. (See Appendix A for a report on course registration at Quest; See appendix B for a list of further recommendations beyond the scope of the work I completed; see appendix C for a discussion about strategies for administrative implementation of our recommendation)

List of Appendices

APPENDIX A—**Course Registration at Quest University**

1 Current Process: A Modified Random Serial Dictatorship

1.1 Description

Currently, Quest University employs a modified version of a method of known as a Random Serial Dictatorship (RSD)[1]. In the context of Quest, an unmodified RSD system would involve a random assignment of priority, with the highest priority student choosing their best available choices (for all four blocks relevant to the registration period), before the second highest priority student making any choices. The second highest priority student would then choose their best available choices (for all four blocks), before the third, and so on.

Quest's system is essentially this, but modified with three main priority mechanisms: the payment mechanism, where only students who have paid a deposit for the coming term may register; the year of study mechanism, where students classified as fifth year students register before fourth year students, fourth year students before third year, and so on; and the timing mechanism, where students who register sooner receive priority over students who register later.

We assume that these mechanisms still leave our ordering random—we do not believe that course-specific demand is correlated with the current fastest-finger method of priority ranking.

1.2 Issues

Firstly, the timing mechanism is problematic. There is no apparent justification to the implicit claim that students who register 5 seconds after the opening of registration are more deserving than students who register 7 seconds after the opening of registration. Furthermore, the timing mechanism is complicated by differences between internet and computer speeds, and thus favors those tech-savvy students and those who have been able to purchase higher-performing computers. We are thus favoring random factors beyond even lack of schedule conflict during registration opening and the savvy to register immediately.

Secondly, though the RSD mechanism incentivizes honesty and offers Pareto Optimal[2] results (assuming any two courses are non-equal in ranking for a given student), it offers less welfare and less equitability than alternative strategies. Basic intuition supports this. Equitability is obviously lacking when some students get to choose all courses before others choose any. Total welfare loss also follows intuitively: a high priority student probably cares less about losing their last choice course than a low priority student cares about losing their first choice course—*viz.* "the lucky gain less than the unlucky lose." This intuition has been confirmed in analysis of actual data from Harvard Business School[3].

---

[1] Abdulkadirouglu and Sönmez, "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems"; Budish and Cantillon, "The Multi-unit Assignment Problem."

[2] "Pareto Optimal" means that no agent in the system can gain anything without another agent losing something.

[3] Budish and Cantillon, "The Multi-unit Assignment Problem."

2 Alternatives Considered and Rejected

Before developing our final recommendation (see section 3.2), we considered many potential alternatives. This consideration involved research, to discover their theoretical strengths and vulnerabilities. We have also had many informal conversations, met with relevant administration (President, CAO, Registrar), and hosted two focus groups consisting of three faculty, three alumni, and four students, to answer relevant values questions. We used this valuable feedback, as well as the qualitative results from a 140-response survey we distributed last winter, to guide us in our discernment.

Furthermore, we considered and rejected changes to our prioritization paradigm. Currently, ranking priority is based on several factors: year at Quest, speed of internet connection, freedom (or lack thereof) from any scheduling conflicts during the actual moment of registration, and access (or lack thereof) to liquid cash assets in order to register on time. Following is a brief rationale of our rejected considerations:

- Prioritizing based on GPA seemed unfair and out of alignment with Quest's values.
- Preferring question-related courses to electives seemed to ignore the value of electives, and undervalue student self-knowledge and advisor oversight.
- A preference for students registering for classes required for graduate school application assumes that students attending grad school are more important than others.
- Lastly, year-at-Quest balance in classes (e.g. somehow incentivizing 2nd years in classes with mostly 1st years, or reserving slots for younger students) did not seem to be feasible within our proposed alternative (see section 3.2)

To assess the following structural (as opposed to above priority assignment) changes to our system, we considered the commonly sought attributes in allocation strategies: transparency, strategyproofness, Pareto efficiency, equitability, and overall welfare (without regard to distribution across students)[4].

2.1 Integer Programming Technique

Our first approach was based on "integer programming", a powerful technique in mathematical optimization. First, the students would assign "preference points" to all classes, varying from 0 points for a class one has no interest in, to 5 points for a class which would be ideal. The program would then search through all possible solutions, and output the master schedule with the highest total point value.

While we originally planned to propose this approach, we realized that this system could yield serious problems, provided students are dishonest. For instance, if a student would like a particular class which she thinks will be in high demand, she could rank that class a five, and all other classes a zero, even if she would be moderately interested in several other classes that block. This dishonesty would then place this student into her preferred class, allowing her to manipulate a mathematical system and

[4] Budish, "Matching 'versus' Mechanism Design"; Pathak and Sönmez, *School Admissions Reform in Chicago and England*.

prioritize her preference in a system which has the goal maximizing total preference points, over all students.

2.2 Colorado College Auction System

Colorado College uses an auction system to register for classes. In this system, each student is allowed to use 80 points for the year, across 8 blocks, bidding for only one course per block. Analysis of this system has shown that it offers unequal results, particularly divided by major [5]. In our environment, which is comparatively small (Colorado College has ~2,000 undergraduate students), bidding would likely be even less equitable. For example, math students who are used to courses with <10 students can leverage dependable assumptions, to never bet more than one point on a math course, giving them an unfair advantage over, for instance, neuroscience students, who regularly occupy 20-student courses. This solution is thus not strategyproof or equitable. Furthermore, the degree of transparency offered by the Colorado College system would likely face significant political opposition among faculty, given that Colorado College publishes history of enrollment, and that Quest tutors have in the past expressed opposition to publishing of data which might bias course preferences themselves.

2.3 Trading Market

After courses have been assigned by any mechanism, an algorithm would identify swaps of courses, where a set of individuals would each "down-grade" in one block for a more significant "up-grade" in another block. This is inspired by, but distinct from, recent work done on kidney swap cycle identification, where families can participate in swap cycles, giving up an incompatible kidney for a compatible one[6]. Such a market would be problematic because it would incentivize complex strategies (likely discovered and understood by a small subset of the students) involving preliminary requests for high-demand courses rather than courses which are actually ideal for the student. While this solution is Pareto Efficient, it is far from strategyproof, and highly unlikely yield a Nash Equilibrium[7]. As a brief aside, depending on the exact mechanisms used for the trading market, it is possible that this solution would yield identical results to the auction solution (2.2) or to our proposed solution (see 3.3), given students with perfect strategy.

2.4 Gale-Shapley Algorithm

In this Nobel Prize winning algorithm, students would rank their classes in each block according to their preference, while tutors in each block would rank students according to their preference (e.g. bumping up their own mentees who need their course for their Question). Then in each block, the students would "propose" to the tutors, the tutors would "marry" up to 20 students, and the simple algorithm would guarantee an optimal fit for both the students and the tutors.

This algorithm has many practical uses, including the national matching system of residents to hospitals in the United States[8]. Our focus group strongly rejected the idea of tutors ranking students, and felt that each tutor's preference list should be a randomized list of 4th years, followed by a randomized list of 3rd years, followed by a randomized list of 2nd years, followed by a randomized list of 1st years. They felt

---

[5] Benson, Johnson, and Lybecker, *Bidding for Classes*.

[6] Roth, Sönmez, and Ünver, "Kidney Exchange"; Gentry, Montgomery, and Segev, "Kidney Paired Donation."

[7] Kalinowski et al., "Strategic Behavior When Allocating Indivisible Goods Sequentially."

[8] Robinson, "Are Medical Students Meeting Their (best Possible) Match."

that the prospect of faculty creating special “preference lists” would create a logistical and ethical nightmare, after considering optics and the added time that would be required by faculty to meet with students and their advisors, to discern which students should be allowed into a course.

3 The Four-Round Snake-Order Draft

A simple mathematical proof shows that--assuming priority would be consistent across all courses in a given block, and balanced across all students in a given year—the above Gale-Shapley Algorithm reduces to a one-way multi-round “draft”, analogous to a fantasy sports pool, where students would pick their courses, one by one, in some prescribed order. A cyclically-shifted snake-order draft across four rounds yields an equitable distribution of dibs across four rounds, within a given grade. For instance, if we had 8 fourth years, the dibs order for the fourth year turn would be as follows:

First round: 1, 2, 3, 4, 5, 6, 7, 8

Second round: 8, 7, 6, 5, 4, 3, 2, 1

Third round: 5, 6, 7, 8, 1, 2, 3, 4

Fourth round: 4, 3, 2, 1, 8, 7, 6, 5

With these rounds thus determined, there are two ways to set up the four-round system: one which is block-specific, or one which is block-nonspecific.

3.1 Block-Specific

This option would require students to list their top choices for each block. For block A, the student offers all of the courses which they would be willing to take that block. A student is allowed to prioritize anywhere from 0 to *n* courses, where *n*=the total number of classes being offered that block. If in block A the student cannot make it into any of the courses they have chosen to prioritize, the student will not be registered in any block A course. The student then completes the same process for blocks B-D.

At this point, we use our four round orders as explained above. The “first round” of the draft is for Block A (January); the “second round” is for Block B (February); and so on. The algorithm then assigns the student with the highest ranking in January their highest-ranked course which is not yet full. Next, it assigns the highest-ranked students their top choice which is not yet full for block B, C, and D, in that order. Following this, it moves to the second ranked person in block A, then blocks B, C, and D, in that order. The process repeats until each student has been either assigned to their top still-available choice for each block or to no course where none of their choices are available.

By structuring the program this way, all students in the same year have the same power in the system, regardless of strategy. There is no way for a student to improve their course registration by reporting anything other than what they believe to be an honest ordering of their preferences. However, this does not allow students to pick where they have priority. For instance, a student who really needs course A1 may randomly be placed in the lowest rank of their year in block A, and end up receiving their first choices for blocks B, C, and D, instead.

3.2 Block-Nonspecific

This option works similarly, except that each student provides a ranking of their relative choices of classes, without sorting by block. If each block has three courses, Student 1 would have ranked the courses in the four blocks in the Block-Specific method in this form:

{[A1, A2, A3], [B1, B2, B3], [C1, C2, C3], [D1, D2, D3]},

In the Block-Nonspecific method, Student 1 would offer a choice order like this (preserving choice order for each block):

[B1, B2, A1, A2, A3, C1, C2, B3, D1, D2, D3, C3].

Courses thus ranked, we use the same snake order before, but create one master list by stacking the four rounds. All of the fourth years appear four times; next come all of the third years, each appearing four times; then the second years in the same fashion; then the first years in the same.

At this point, the student with 1st choice chooses their top choice still-available course, followed by the 2nd, the 3rd, etc. The process continues until all assignments are complete (to a course or to none where no options are still available).

By structuring the program this way, students can prioritize certain blocks over each other, to ensure that they “spend their dibs” on courses of greatest value to them. However, this system is somewhat vulnerable to gaming, for the simple reason that if student 1 knows that B1 usually only has <10 students in it, student 1 might rank B1, B2, B3 at the very end of the list, even if B1 is the course that is of most value to their academic path.

3.3 Rationale For Recommending Block-Nonspecific approach

This approach is Pareto efficient, but not strategyproof. The strategy is relatively simple—each student ranks the usually-full top choices above the usually-less-than-full choices—and in its simplicity, the system offers a type of equality[9]. That being said, strategic (i.e. dishonest) play can significantly decrease overall utility. A study of Harvard Business School found that strategic play (though still significantly better than the RSD) in this system yields about 25% less net perfect student-course assignments than honest play.[10]. However, strategic play should have less of an impact at Quest, since our block system diminishes negative consequences due to greater instance of schedule conflict.

4 Implementation

Barring any major unforeseen logistical issues our registration system will be implemented for the 2013 October Registration Cycle. We are currently collaborating with the IT team and Registrar Tim Schoahs to transfer our program’s output into the university’s registration system. We are also developing a web

[9] Notably, students can still have different levels of advantage according to the amount of unpopular courses they desire, like in the Colorado College auction system.
[10] Budish and Cantillon, “The Multi-unit Assignment Problem.”

interface which can give us course preference ordering information from students in a format which is useful to us.

4.1 Caveats

To create offer a system which leaves students as empowered as possible in choosing their courses, they may also provide caveats.

We will include "A only if B" commands to deal with courses where the student is attempting to register for two courses, where one course is the prerequisite for the other. In the case of a student receiving course A, but not the course required to satisfy their "A only if B" constraint, the system will backtrack to the point of the assignment process where the student was assigned course A, and instead assign the student their next available choice. This preserves the power of information currently available to students during registration.

We will also include a mechanism where a student can create "exclusive" commands. A student could say "I want only one of the following courses this term: [Fate and Virtue (January), Reason and Freedom (January), Identity and Perspective (March)]", or say "I only want *two* of the following courses this term" and list the same courses. This would automatically apply to multiple courses that satisfy the same requirement (e.g. no more than one foundation math, no more than two foundation social sciences). In the case of the exclusive command, we will simply halt assignment once the quota for that given set is filled.

4.2 Timing of Advisor Approval

Quest's Faculty have unanimously agreed to shift advisor approval of course registration to "pre-registration." In other words, faculty advisors must view student course-choice rankings, and approve those rankings, before they are submitted and thus before the registration takes place.

5 Conclusion

Overall, the shift from a Random Serial Dictatorship to an equitably ordered multi-round draft will indeed improve the system. It should provide a more equitable distribution of high-demand courses, and a greater number of total students getting into their first choices.

That being said, the course planning at Quest is far from perfect. In the future, the registration structure itself could be altered to include a proxy draft, where the program strategizes for each student which course is the best to choose, based on its enrollment in that round, rather than depending on an essentially set (besides its caveats) list (see Appendix C).

Furthermore, other features of the system could likely produce far more utility than the specific structure of the registration algorithm (see Appendix B). Fortunately, this project seems to indicate that Quest's community and power structures are ready and willing to consider and implement changes which can be justified based on the tenets of quantitative reasoning, on short time frames. In other words, if you are a Quest student, you can make a difference.

Bibliography

Abdulkadirouglu, Atila, and Tayfun Sönmez. "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems." *Econometrica* 66, no. 3 (1998): 689–701.

Benson, Cassandra Michele, Daniel K. N. Johnson, and Kristina M. Lybecker. *Bidding for Classes: Course Allocation Under the Colorado College Auction System*. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, April 24, 2013. http://papers.ssrn.com/abstract=2257921.

Budish, Eric. "Matching 'versus' Mechanism Design." *SIGecom Excehanges* 11, no. 2 (2012): 4–15.

Budish, Eric, and Estelle Cantillon. "The Multi-unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard" (2010). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1539313.

Gentry, Sommer E., Robert A. Montgomery, and Dorry L. Segev. "Kidney Paired Donation: Fundamentals, Limitations, and Expansions." *American Journal of Kidney Diseases* 57, no. 1 (2011): 144–151.

Kalinowski, Thomas, Nina Narodytska, Toby Walsh, and Lirong Xia. "Strategic Behavior When Allocating Indivisible Goods Sequentially." In *Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*, 2013. http://www.cse.unsw.edu/~tw/knwxaaai13.pdf.

Pathak, Parag A., and Tayfun Sönmez. *School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation*. National Bureau of Economic Research, 2011. http://www.nber.org/papers/w16783.

Robinson, Sara. "Are Medical Students Meeting Their (best Possible) Match." *SIAM News* 36, no. 3 (2003): 8–9.

Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver. "Kidney Exchange." *The Quarterly Journal of Economics* 119, no. 2 (May 1, 2004): 457–488. doi:10.1162/0033553041382157.

APPENDIX B—**Further Steps to Improve Course Registration**

From the survey, focus groups, and reflection, we also gathered some ideas of how we might improve issues not directly related to or solved by either of the above options. They can be found below, in no particular order:

- Changes to priority which we have not yet dismissed
    - Adjust the course count for year at Quest by one or two blocks to better represent time at Quest (notably, we received suggestions to move the barrier both directions—perhaps we could base it on semesters rather than years)
    - Extend waitlist acceptance time period to more than 48 hours
    - Offer a wait-list bump-up to students who are booted from classes cancelled after registration, so that their wait-list priority is as high as it would have been had they joined a waitlist
- Planning and Information sharing
    - Publish time of day with schedule before registration
    - Publish estimated book/equipment price with course description before registration
    - Require each faculty to write a course description for all courses offered for registration (rather than using the same description for all courses of the same title, but by different tutors)
    - Either finalize the schedule sooner or place prerequisite courses in the same term (and before the class that requires the prerequisite), so that people can take needed concentration classes without having a dependable schedule ahead of time
    - Clarify process for skipping prerequisites, and allow more options to take a concentration course in lieu of a foundation, where students have background or demonstrated aptitude in the subject
    - Create a single, fair, and clear way for students to receive tentative schedules
- Plan the schedule with better information
    - Across Areas:
        - Many complaints about concurrent courses in different areas appealing to the same interest groups (eg the overlap of the LoK theme stream with Logic & Metalogic; Godel and Turing ought not be mutually exclusive)
        - One specific request that stats be offered in the winter, since life science students often need to go outdoors in the fall and spring
    - Many requests that student desire to take courses be collected in some way, to influence the order courses will be offered in (similarly, some students requested that desire affect decisions about which courses to be offered)

APPENDIX C—**Non Mathematical Questions of Implementation**

One major unresolved question is how Quest administration will present and implement this program. One (likely unpopular among faculty) approach would be to offer a complete record of historical course registration, for students to use to predict desirable courses.

Another option would be to offer a preliminary survey asking students to fill it honestly, with non-identifiable results published to offer equal information to all students. This would give less advantage to knowledgeable students, but more advantage to those with “natural advantage” (eg the math students rather than neuroscience students, as discussed in Appendix A Section 2.2). It would also likely reduce net welfare loss due to sub-strategic play, where a student assumes a course is less popular than it is, and thus gets a course they care less about than the course they would have received from honest reporting in the Block Non Specific Snake Order Draft.

A third option would be to ask students to answer honestly, appealing to the reasoning that if all students report honestly, utility will be maximized.

A fourth (and most likely, at least in the short term) option is the *laissez-faire* information and comment approach, where administration makes no recommendations about honesty, and offers no information not currently offered.

Finally, a fifth option would be the development of a strategic computer proxy, empowered to make strategic choices based on how much space remains in each course on a student’s preference list. While likely the ideal implementable option, it is likely not feasible for this registration round, and is also much more complex (and thus less transparent) than the above options.

APPENDIX D—**Announcement of Roommate Matching Program Implementation**

Posted on the Facebook page for the incoming class of 2013:

Ever wondered how we assign your housing at Quest? Check this out!

Like many things at Quest, we do it a bit differently. Our goals this year was to expand on our already hands-on, human process to get the best possible matches, and be as inter-disciplinary as possible in our thinking.

These photos are of Richard Hoshino, Math Tutor, and Kim Mortreuil, Student Service Coordinator from today. They are using some custom algorithms written by Richard based on Nobel prize winning pattern matching theory to help us find good matches for your room-mate and floor preferences. Remember that long form you filled out? Here is where the data ended up!

Building on a student class project, and the initiative of Caleb (a Quest Student) and Richard, we redesigned the housing application this year to use this new matching theory approach. We wanted to not just know what your preference was, but how important it was to you. Instead of just randomly assigning students together like other universities, we take a very involved approach to creating each floor and each condo-group and this matching technique puts more order to randomness, while being time efficient.

Many factors are used to create a balanced mixture of students on each floor (year of study, international mix, preferences, male/female balance and our unique personal knowledge of people). We balance all of these to create, what we hope is the best community, while also trying to match your preferences as much as possible.

Another cool thing... Unlike any other university, Admissions Counsellors at Quest made a first pass as creating room-mate pairs and condo groups based on how they know you, and the data from your applications. Another way to demonstrate that you are not just a number at Quest!

Finally we in Student Affairs work hard to use these complex preference matching criteria to try to find the best possible matches. Once the formula pops out good matches we use a whole bunch of good old human consideration to help finalize your placements. Richard and Caleb have been awesome and we have had a lot of fun working together collaboratively. It's a best combination of human intelligence, technology and people coming together to collaborate.

Pretty cool huh?!?

Here are some articles if you want to learn more...

http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2012/popular-economicsciences2012.pdf http://www.nytimes.com/2012/10/16/business/economy/alvin-roth-and-lloyd-shapley-win-nobel-in-economic-science.html?_r=0

John Scheunhage

Campus Counsellor

3200 University Blvd. Squamish BC V8B 0N8

604 898 8084

www.questu.ca <http://www.questu.ca/>

QUEST UNIVERSITY CANADA

Because education is not about learning the right answers; it's about learning how to ask insightful questions.

APPENDIX E—**Roommate Matching Program**

1 Motivation

The Quest University Residence life team, in an attempt to match incoming first years into roommate pairs by some criteria besides a similarity in the early letters of their last name, has worked hard over the years to place students in rooms which will allow each student to be comfortable, and to foster a campus community in line with Quest's ethos.

Because Quest administration recognizes the importance of students' home environments during their first year, they have traditionally devoted tremendous staff time to the task, and been bound by the computational limits of pen-and-paper approaches.

We aimed to significantly reduce staff time required for the roommate matching program by providing an integer program and modifying the survey. By doing so we aimed to widen our range of data and better determine the relevance of specific data points, to increase both time efficiency and overall assignment utility.

1.1 Original System

The original system worked in 2 main steps. Firstly, the Admissions Team would work to match incoming students with one another. Generally admission counselor A would pair a given student with another student counseled by admission counselor A.

This being done, the Student Affairs Team would review the roommate matches, and look for issues based on the survey. Next, the Student Affairs Team would assign condo matches (4 people total—2 roommate pairs). The relevant questions in the survey to this process were:

- Country & State/Province of Residence
- Gender
- Personality—Please choose the characteristics that describe your personality
    - Introverted
    - Extroverted
    - Artistic
    - Active
- Lifestyle Preferences 1—Please choose the characteristics that describe your lifestyle. Choose ALL that apply
    - Early Riser
    - Night Owl
    - Orderly
    - Messy
- Lifestyle Preferences 2-- Please choose the characteristics that describe your lifestyle. Choose ALL that apply
    - Smoker

- Occasional Smoker
  - Non-Smoker
  - Non-Drinker
- About yourself (Ideal living scenario)—Please write a short description about yourself and what you are looking for in residence. Say a bit about what you are looking for in room-mates, your basic lifestyle preferences sand values, and the kind of living experience you would like to create. Examples include hobbies, interests, language abilities, social preferences, etc.
  - (blank text box)
- Fluent Languages—Please enter the languages you can speak fluently other than English
- Other comments/Medical Considerations

1.2 Critiques of Original System

The first issue, particularly for the incoming class of 2013 (~230 students incoming, in a school to be ~530 total), was a tremendous population boom, perhaps because of this, the admissions team was not able to provide match suggestions within the required time frame. To improve timeliness, we needed to help provide a more organized set of potential matches, instead of a list of names in alphabetical order.

Secondly, each admissions counselor matching students with other students with the same counselor is problematic. Partially, admissions counselors are geography specific, which meant that people from similar geographic regions ended up as roommates (not ideal in achieving Quest's goal of an integrated international community). Furthermore, breaking up the whole incoming class into micro-populations decreases the potential for overall utility. Two students who might match perfectly, but who are served by different counselors, are less likely to end up with each other than would be desirable.

Thirdly, the questions themselves lend themselves to a "one-size-fits-all" assumption about how much various issues matter. For instance, if student N was a non-smoker, the data would seem to imply that student N should not be put with student M, who is an occasional smoker. But if student N doesn't mind the smell of smoke sometimes, and student M is trying to quit (and would thus like a roommate who isn't constantly inviting M to smoke), our system offers a misassumption.

2 The Revised Approach

In order to address these issues, we implemented a matching program—based on a method originally conceived of by 4 Quest Students, Khoya Craig, Aida Ndiaye, Graham Fischer, and Becky Schile—with a survey that allowed students to rate the importance of each issue they provided data on. We then used a greedy algorithm to match students into high-compatibility pairs. The overall process was basically unchanged—student names to admissions, and then to the Student Affairs Team for matching into condos and floors—except that the list given to the Admissions Team was supplemented with a list of recommended matches and compatibility scores.

2.1 Revised Survey

We refined the relevant portion of the survey into a specific format, with four preference question pairs (8 questions total), each pair asking about a lifestyle preference (*u*), and the degree to which the student cares about that given preference (*v*). We also included one question which we heavily weighted automatically. The relevant portion of the survey is as follows:

- Social Interaction (O)
  - Scale from 1 (introverted) to 5 (extroverted)
  - Scale from 1 (I don't care about my roommate being like me in this regard) to 5 (I care very much that my roommate is similar to me in this regard)
- Sleep Preferences (L)
  - Scale from 1 (morning person) to 5 (late night person)
  - Scale from 1 (I don't care about my roommate being like me in this regard) to 5 (I care very much that my roommate is similar to me in this regard)
- Room Environment (E)
  - Scale from 1 (very messy) to 5 (very tidy)
  - Scale from 1 (I don't care about my roommate being like me in this regard) to 5 (I care very much that my roommate is similar to me in this regard)
- Smoking (M)
  - Scale from 1 (Never) to 5 (more than once per day)
  - Scale from 1 (I don't care about my roommate being like me in this regard) to 5 (I care very much that my roommate is similar to me in this regard)
- Substance-Free: "How important is it that you are matched with condo-mates who live a substance-free (alcohol, drugs, smoking) lifestyle?" (F)
  - Not at All, Neutral, or Very Much

2.2 Compatibility Scores

To calculate uncompatibility score $S^{1,2}_O$ (difference between student 1 and student 2 in the plane of social interaction compatibility), we multiply the personality difference (O^1_u-O^2_u) by the sum of how much each one prioritizes the issue (O^1_v+O^2_v). An identical process is used for the $S^{1,2}_O$, $S^{1,2}_L$, $S^{1,2}_E$, and $S^{1,2}_M$.

To calculate $S^{1,2}_F$ we assume a high priority score (automatically 15, compared to maximum of 5 for each other compatibility plane), and thus assign it 30*(O^1_u-O^2_u).

To calculate overall uncompatibility $S^{1,2}$, we sum $S^{1,2}_F+S^{1,2}_O+S^{1,2}_L+S^{1,2}_E+S^{1,2}_M$.

For a simple example, Imagine two students, X and Y.

X and Y both are introverts (1), both are night people (5), both are regular time people (3), both are Neutral on Substance-Free, but they differ on smoking: X smokes never (1), and does not care whether a roommate smokes (1); Y smokes socially (3), and cares very much that a roommate would join (5). This makes their total uncompatibility=0+0+0+0+((3-1)*(5+1))=12.

2.3 Greedy Algorithms

We choose the first student, and assign them the nearest match (lowest score). We then do the same to the next student, and the next, until all students are paired up.

Following this, the list of recommended pairs is given to Admissions, and then to student affairs. Once all students have been assigned a roommate, we again use a greedy algorithm for condo assignments (two pair groupings).

We select a roommate pair A, with students 1 and 2 and compare it to a series of other pairs (B, C, D…). To check uncompatibility between pair A and pair B (students 3 and 4), we sum $S^{1,3}$+$S^{1,4}$+ $S^{2,3}$+ $S^{2,4}$. We then use the same method to check between pair A and pair B, pair A and pair C,…

When pair A's lowest uncompatibility pair is found, we assign them that pair, remove both pairs from the pool of unassigned matches, and continue to the next pair. The process continues until all pairs are matched.

3 Conclusions

According to verbal feedback, we provided a significantly more time efficient solution, based on a more robust interpretation of student-provided data. That being said, we are unsure about the net impact on system utility (total happiness with roommates). Firstly, the incoming first years have not yet lived with their roommates. Secondly, such happiness would be extremely complicated to measure. We may be able to examine retroactive data about how many people switched roommates during the year last year (although we are disadvantaged by sample size), and about how many people stayed with room or condo mates into subsequent years (although we are wary to assume that continuity represents desirable outcomes according to the Quest ethos).

Given the above measurement issues (as well as issues relating to the complex psychology of self-reported behavior and preference), we cannot conclude that we necessarily increased (nor that we decreased) system net utility compared to the previous system. We do think that, based on our intensive usage and sorting of data, our system performs much better in a short time frame (several days) than the previous one could.

Furthermore, since lifestyle preferences matter very much to some students, we can reasonably conclude that our system still provides considerable utility advantages over a randomized system, used by many other universities, without any significant difference in run time.

3.1 Further Changes

In the process of revising the survey, we developed a slate of modifications we could make to our system.

For instance, we might improve the Substance Free priority score, given that the question is on a 3 point scale. The maximum difference between two people using our method is 60, whereas the maximum difference between two people in any other plane is 50. Perhaps we should give more weight to the Substance Free attribute.

Another change would be to change the gender box from a m/f choice to an open box (to be in right relationship with the gender queer community) and to ask students which gender they would prefer in a roommate (there is no fundamental reason why all students of one gender would prefer a roommate of the same gender).

We also would like to consider changing some 2-question sets into 3-question sets—for example, an introvert might highly value the opportunity of living with an extrovert; our system only allowed them to say they don't care who they live with, or that they value an introvert.

Furthermore, we had a slew of potential other questions to add to the survey (in no particular order):

- Athlete to non-athlete and priority
- Video game player to non-video game player and priority
- Replace tidiness question with three questions
    - Bathroom cleanliness and priority (matching roommates)
    - Bedroom cleanliness and priority (matching roommates)
    - Common room cleanliness and priority (matching roommates and condomates)
- Length of time having lived away from home/been out of school priority (1, 2, 3, 4, 5+ years) and priority
- "Open Door" social space to "Sanctuary" refuge space and priority
- Sexual activeness and priority
- Music volume and priority
- Temperature and priority
- Study partner to party partner and priority
- Dietary restrictions (Vegan, Vegetarian, Pescatarian, etc) and priority
- Religiousness and priority

There may be problems with adding too many dimensions of priority to the system, in that some students may be dissatisfied if many of their concerns go unmet. That being said, if we trust students to accurately self-report preferences, we can reasonably assume that an expansion of the survey like this would improve overall happiness. That being said, each of these questions requires a value-based conversation about its accordance with Quest's goals. For instance, in our first example, allowing Athletes to request other Athletes may reinforce the extant divisions in the Quest community. Each question (including the ones already decided upon) also warrants further study into the psychology of compatibility, to discover where student preference self-reporting may or may not be dependable (eg do extroverts who request extroverts *really* enjoy living with extroverts, or do they just think they do)/

Lastly, a major improvement that could be made to the system (especially if we add a number of attributes) would be a more comprehensive version of the greedy algorithm. One way to do this would be to run the greedy algorithm n times, where n= the number of students, and then the condo-matching algorithm n/2 times. Each time it would start on a different student. The time that yielded the maximum system-wide optimum could be chosen as the final solution for roommate pairs. The condo matching algorithm would then maximize condo assignments based on those roommate assignments. Alternatively, we could run the whole process n*(n/2) times, to yield the best overall condo optimums. Furthermore, heuristics could be combined with a non-greedy method to prefer system-optimum-allowing, pair-suboptimal assignments within a certain uncompatibility score maximum.

APPENDIX F—**Reflections From an Artificial Intelligence Conference**

http://credotwopointoh.blogspot.ca/2013/08/reflections-from-artificial.html

Last month, I went to the 27th conference of the Association for the Advancement of Artificial Intelligence. I tried to publish a paper to present at the conference, and failed, but went anyways since it was in Bellevue (across the water from Seattle, where I'm from but not where I live). I'm told by my academic advisor (who accompanied me) that this is the world's foremost academic conference on the topic. I believe him.

Within a few hours of showing up, I realized that he hadn't been lying when he said that I'd be one of the very few undergrad students there. In fact, (to the best of my knowledge), I was the only one.

My time at the conference could be broken up by the following approximations:

30% having NO IDEA what the hell the person at the front of the room was trying to communicate--or even like the general context of where one would apply their work.

30% understanding why the speaker's work was important but not getting their conclusion and/or reasoning.

30% actually kinda getting it!

10% schmoozing

One thing I noticed about AI research is how much has been developed, but not implemented. So much is just waiting to be commercialized or otherwise made available. But more on that in later posts...

Overall, I basically realized how wide and deep the field is. I heard about highly contrived game theoretic examples about how rules impact wait times in lines--in two restaurants across the street from each other serving identical food at identical prices, with people who wouldn't leave a line once in it. I also heard about concrete disaster management data analysis using massive amounts of social media data.

I heard about abstract mathematics, and, just a few minutes later, about engineering issues for robotics. I heard about how robots analyze humans--navigating a crowd, for instance. And I heard about how humans can deal with robots, like the study on autistic kids relating better to robots than to other children.

The problem, though, with having such a massive field, and so many concepts, is that people can't keep abreast of everything going on. Usually, this is fine. Obviously if I'm an engineering person, highly abstract mathematics don't have a huge amount of bearing on me--sure, it would be nice if I knew about it all, but not necessary. But in a field where everyone's research falls on some n-dimensional gradient, a lot of people are working on problems a lot more similar than they think.

Sometimes, someone is working on machine learning for visual processing. Someone else might be working on machine learning for metaphor identification. At this conference, it's unlikely that these two people would ever be in the same room--the robotics sessions happen in one room, while the semantics sessions happen in another. I had this very experience.

I attended an integer mathematics session on scheduling problems. This is the session that my attempted submission would have fallen into. Right afterwards, coming out of an adjacent room, was a man headed towards the poster area (where the posters for that time-slot were). He looked like an interesting guy, so I strolled towards the poster he ended up parking himself beside--clearly an author of this particular poster. I began to skim it, and noticed to my surprise, that it was eerily similar to the project I submitted a paper on.

In fact, his method was the very method I had recently decided to be the approach I recommended after doing research after my initial submission. We were working on literally the same problem, but I would have been in the "scheduling problems" track, whereas he was in "auctions and allocations" or something like that; I was doing integer programming and he was doing game theory. But we both were working on the same problem. In fact, from the title of his article (and the language that it used) I never would have guessed that it was talking about the very same issues and approaches that I was.

From conversations and my own observations of this sort of redundancy, I've gathered that this was not an isolated incident. Now, I'm sure that this sort of thing happens all of the time in other fields, too. But there's a special irony to it happening in the area of AI research.

First, let me offer an example. A project I'll be starting this October is coding a program to build a course schedule based on student desires. The idea is that the faculty would give me a list of all the courses to be taught, and even perhaps how many of each would be taught, and then that I would send a survey to students, asking them which courses they would like. This type of thing is uniquely helpful at a block school like quest (1 month at a time), because we see a lot of conflicts, like two courses, each about one of the two most influential mathematicians, one of whose work was built directly on the other's, both taught at the same time. Any pure mathematics student would desperately want to take both courses. Or internet politics at the same time as constitutional law--at a school with less than 200 upper-year students, any politics student would have likely wanted to take both courses.

That project is an AI project. Not an innovative one, in a theoretical sense--just a simple application of old techniques.

Now let's look at the fact that the AI conference isn't set up for me to attend two different presentations on the exact same topic. It could benefit from the same very simple type of integer programming to create a schedule to avoid these types of issues, with a basic survey sent out to registrants (and, of course, whatever additional restraints, like some speakers giving more than one talk).

Given that these are AI **experts** running a conference, I feel like we could use more AI to achieve our ends.

In fact, it's even more generalizable than that. One could imagine an AI-powered system designed to identify similarities between solutions or between problem-solution pairs by combing through AI literature for various open issues in conclusions of papers, and . It might even be able to analyze and answer questions posed through a wolfram-alpha style user interface. Luckily there's already some cool stuff being done on this. AITopics.org is a small team of people working on aggregating AI information using AI techniques--but it's only step one, and they're a very small team.

I suspect that if the AI community rallied around using AI to improve communication in their field (likely the work could be applied to other fields with relative ease), a large-scale effort in the short term could drastically improve the pace of progress for the whole field.

Just my two cents.