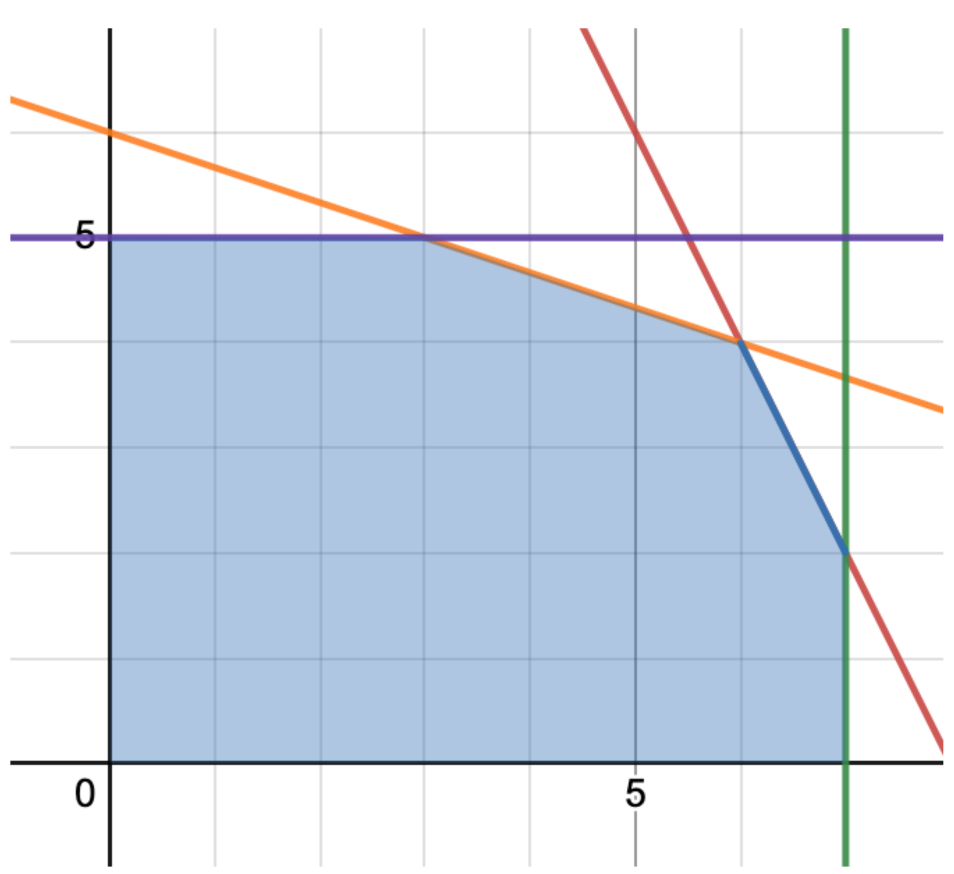
Integer Linear Programming for Stable Matching

Mary Blanchard, Batmend Batsaikhan, Cecilia Erlichman, Andreas Miller, Petrichor Park, and Hugh Shanno

What is Integer Linear Programming?

- Integer linear programming (ILP) is a subset of **linear programming** (LP). It is a method to formulate a problem as variables and linear constraints.
- **Linear constraints** are linear equations or linear inequalities that relate the variables to each other in the form $ax + by + \ldots = cz$.
- **Solutions** are found when variable values satisfy all constraints.
- **Optimal solutions** are found when a solution maximizes or minimizes the **objective function**.
- ILP further constrains variables to only take on **integer values**.



In this example, we have two variables (x and y), four linear constraints, and two sign constraints. This creates a region of viable solutions. If we wanted to maximize or minimize an objective function relating x and y, the maximum or minimum would occur on one of the boundaries of the viable region.

What is Stable Matching?

- Stable matching problems involve using rank-ordered **preferences** to create pairs or groups.
- A **blocking pair** is made up of two entities who are not paired together that would strictly prefer to be paired with each other over their current matches.
- A **stable matching** is a set of pairings with no blocking pairs.
- In the **CS match**, students rank courses by their preference, and courses create "preferences" of the students by class year and previous courses taken.

Acknowledgements

We would like to thank Layla Oesper for being our advisor, and Dave Musicant for his help and insight into the Carleton CS Match.

ILP Formulation of the CS Match

Variables

- Student-course pairings (1 = matched, 0 = not matched)
- Student "weight" or "happiness" (auxiliary)

Constraints

- Students can only be matched to one course.
- Courses can not have more students than seats.
- There can be no student-course blocking pairs, enforced by ensuring that if a student x is not matched to a course, that course is matched only to students it ranks higher than student x.

Objective Function

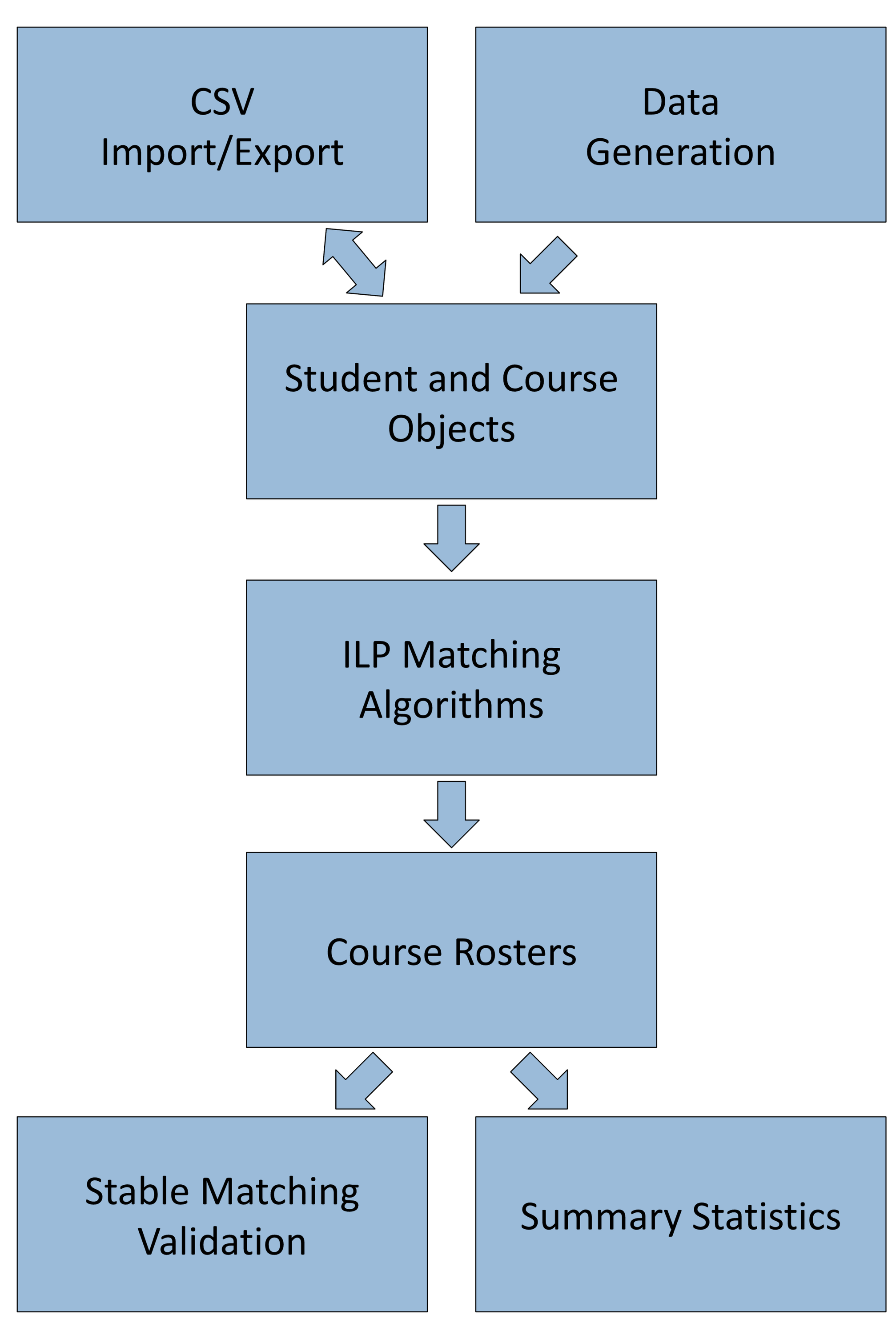
- Gale-Shapley emulation maximizes number of students matched and minimizes preference index.
- Hospital-Resident model maximizes number of students matched to courses.
- Weighted match maximizes both number of students matched and the weights of the matched pairs.

Efficiency of an ILP Solver Algorithm

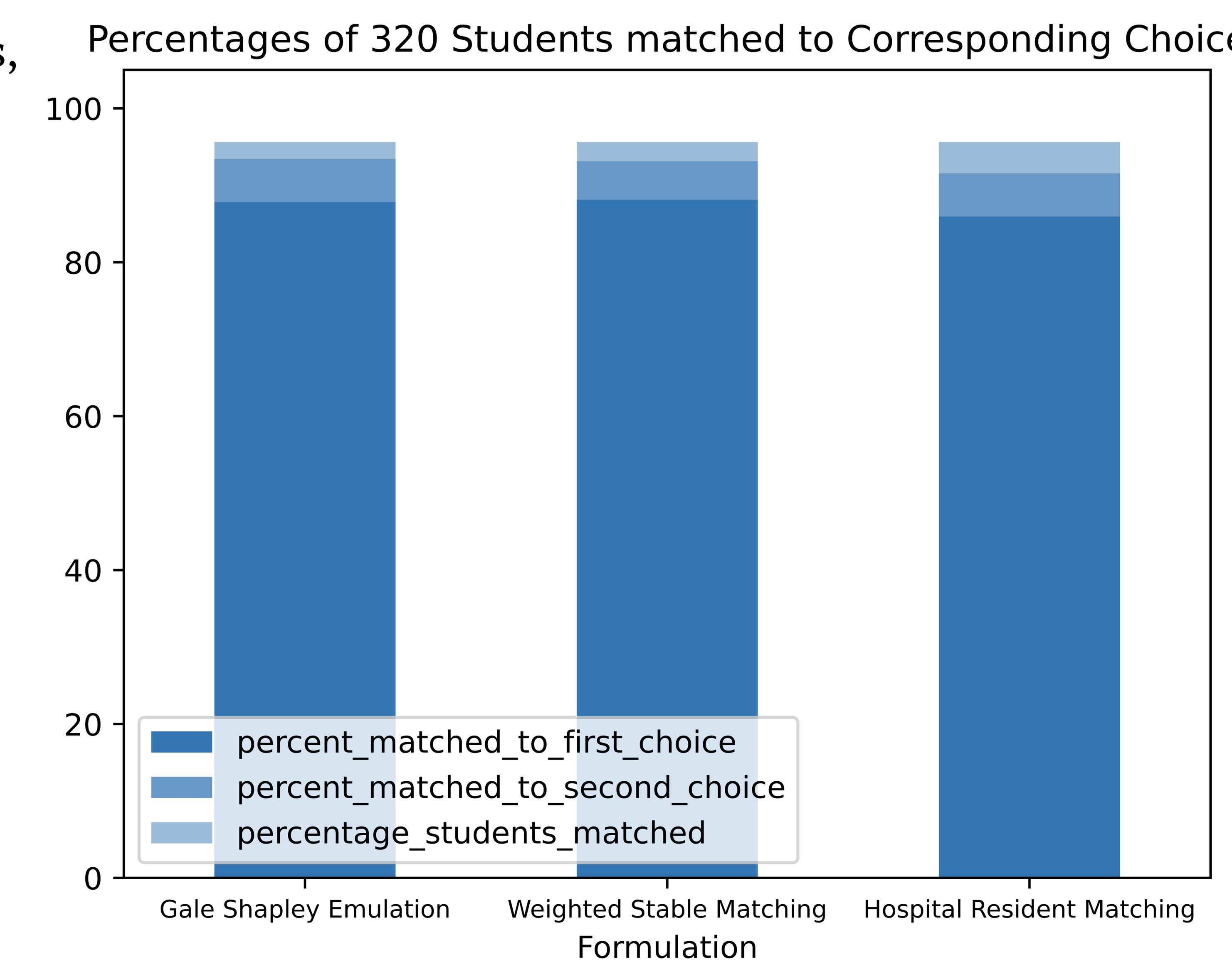
Simplex Algorithm

This algorithm substitutes inequalities for equations by introducing "slack" variables. It then searches for solutions by **traversing facets of the polytope** of the feasible region. Moving from one facet to the next is called a pivot and is conceptually similar to rearranging the system of equations to solve for one variable. The efficiency of the simplex algorithm varies based on the problem. For average problems, it is **very efficient**, typically running in linear time, but for worst-case problems, it can take exponential time. One such problem is the **Klee-Minty cube**, which was used to show worst-case running time for this algorithm. A Klee-Minty cube is an n-dimensional polytope with perturbed edges. Many standard simplex algorithm implementations will have a worst-case execution that includes visiting every facet, or 2^n pivots. Using randomization, this has been improved to $\mathbf{2^{O(\sqrt{n}\log m)}}$ for n dimensions and m constraints.

Our Code



We generated datasets of students, courses, and preferences to simulate the Carleton CS Match using three ILP formulations. In the example matching shown here, all three formulations resulted in 95.6% of students being matched to a course, with around 88% matching to their first-ranked course. Although this is better than the 74% of students in real CS matches who get their first course, the high variability between our generated data sets leaves too much room for error to recommend implementing an ILP solution over the current CS match algorithm.



References

Delorme, M., García, S., Gondzio, J., Kalcsics, J,. Manlove, D., & Pettersson, W. (2019). Mathematical models for stable matching problems with ties and incomplete lists, *European Journal of Operational Research*, *277*(2), 426-441. https://doi.org/10.1016/j.ejor.2019.03.017.

Gusfield, D. (2019). *Integer Linear Programming in Computational and Systems Biology: An Entry-Level Text and Course*. Cambridge: Cambridge University Press. doi:10.1017/9781108377737

Klee, V., Minty, G. (1969). How good is the simplex algorithm?, *Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969*, 159–175. Academic Press, New York-London, 1972