Research Advisor: Ariel Procaccia

Short Abstract

Wesley Pegden and Ariel Procaccia recently proposed a non-partisan redistricting protocol, whereby parties can split a state into districts by taking turns proposing a districting and freezing a single district. Throughout the course of this senior thesis project, we will examine both the properties of this and other redistricting protocols and their effects when used in practice. In order to support the investigation of these issues, we construct an interactive online simulation to allow users to play the protocols against artificial agents within a simplified game UX that corresponds to real world electoral data. Using this platform, we hope to (1) investigate how efficiently compute an optimal or near-optimal strategy, (2) help conceptualize the protocol by mapping it from simplified geographical representations to the real world, (3) discover potential second order effects of the protocol when used in practice, and (4) raise awareness of the protocol among relevant audiences.

Description of the Problem:

Gerrymandering is a process where a state's legislature divide a state's congressional districts in a manner maximizing the advantage of one political party. Although extensive investigation has been conducted into quantifying gerrymandering and examining its effects, most existing solutions rely on non-partisan agents, such as judges or independent commissions, to moderate the redistricting process. Unfortunately, the use of independent agents creates additional problems: its difficult to find agents both experienced and capable at performing the task while remaining truly impartial. In their working paper [1], Pegden and Procaccia propose an I-Cut-You-Freeze protocol, whereby two parties acting selfishly create a non-partisan electoral division. In the protocol, parties take turns dividing the state map into districts, and then choosing from the districts proposed by the other party on which district to "freeze." In addition to guaranteeing districts are divided roughly in proportion to the party membership of the state, this protocol promises that one party cannot "pack" a minority group into less than $$\sqrt{n}\vert{2}$$ districts if the other party opposes it.

While the protocol has some nice theoretical properties, there are many unanswered questions regarding its effects in practice. A significant amount of work needs to be undertaken in order to understand how the protocol affects cracking or packing under non-idealized conditions, and to understand whether the protocol produces socially beneficial outcomes under more complex scenarios, such as when in the presence of multiple or non-disjoint minority groups. Investigations into how to efficiently play the protocol optimally have not been conducted, making it difficult to understand what kind of second order effects are produced from merely near-optimal play. Producing an algorithm to demonstrate that a party can play optimally or near-optimally would help thoroughly explore these effects and would help convince users that such a protocol would be viable in practice. Other open questions related to the work include examining what effect the protocol produces when parties want to ensure that they have a stronger hold on the districts than a simple majority, or exploring what happens when under different geographical constraints. By gathering data about how the protocol works in practice, we hope to answer some of these unknowns, identify new questions and better educate users on the usefulness of such an approach.

- **Planned research contribution:**

Throughout the next two or three semesters, I plan to solve the following sub-problems:

1. Locate a good dataset from multiple states or combine election data with 2010 census data. If a consistent dataset does not exist for many states, focus on individually producing datasets for some of the more heavily gerrymandered states.
2. Find a bijective mapping between this data and some sort of grid of selectable squares, where each selectable square represents a constant sized subpopulation of the state. Whatever mapping is chosen should be amenable to a friendly user interface. This will require either modifying an existing topology-preserving cartogram method or coming up with our own solution.
3. Construct an artificial agent capable of efficiently playing the game with some reasonable ability. Initial approaches to this problem could be simple, but later approaches should incorporate more advanced methods.
4. Stitch the above tasks together into a professional looking website.
5. Analyze the data produced from users play against the artificial agent (and possibly each other).

Research timeline:

In order to better break up the research, I'll split the work into half-semester chunks. In general, I believe that these chunks represent independent tasks, while simultaneously frontloading the risk of the project so that, if necessary, alternative approaches may be devised.

- January- March 15, 2018:
  - (Task 1) Attempt to find a generalizable way to extract election data from different state datasets. Finding potential sources of consistently formatted data will be a challenge, but potential sources include David Bradlee's gardlow.com or non-profits. If no easy generalized approach exists, for 3+ states, one of which being Pennsylvania, gather and format a usable dataset such that population voting preferences is known to a reasonably granular level.
  - (Task 2) Experiment with a few different methods in order to identify a way to efficiently represent the districts. Some initial approaches could include (1) adopting a cartogram method and overlying a grid on top or (2) finding some objective function that provides a rough heuristic to preserve geography for the mapping and then using ILP to minimize this objective function
- March-May, 2018:
  - (Task 4) Use the method identified from task 2 to construct an aesthetic user interface that allows users to select districts within the constraints set forward by the protocol.
  - (Task 3 cont.): Begin investigating ways that this can be played as a game. See if there are any obvious properties of the protocol that may be exploited to construct a competitive artificial agent. If no such easy approach exists, attempt to define an evaluation function that gives a reasonably good scoring of how each side is performing, to be later used in a minimax approach.
- September-October 15, 2019:
  - (Task 3 cont.) Use the heuristic or evaluation function identified in the previous semester to develop an AI. Continue to try to find a way to build an AI with more formal guarantees about its performance based off the properties of the protocol. Write a report detailing preliminary investigation into task 3.
  - (Task 4 cont.) Incorporate the AI into the user interface of the website. Work with designers to launch the website. Though questions about improving the AI might remain, a sufficient amount of work should exist by this no open questions relating to the launch of the site remain.
- October 15-December 15, 2018:
  - (Task 4 cont.) Complete any remaining tasks necessary towards launching the website in 11/18, allowing users to experiment with the protocol and to see how the first-run AI plays against them.
  - (Task 3 cont.) If an AI can't play the game optimally or within some reasonable guaranteed bound, see if the AI could be improved by using techniques such as NEAT, CFR, or any machine learning methods. Explore if the user data generated is sufficient to bootstrap an ML approach.

- January – March 15, 2019:
  - (Task 3 cont.) Collect and analyze the data produced from users play against the artificial agent (and possibly each other). Evaluate the proposed redistricting using some of the Gerrymandering tests proposed in prior work in this area, such as the tests proposed in [5], [6] and [7]. Write a preliminary report on the results of user play.
  - Continue to implement the AI improvements from the previous semester. Write a report on investigations into machine learning approaches into creating an AI.
- March 15 – May, 2019:
  - **Compile written reports into a senior thesis project.**

Background reading to be carried out:

Throughout the course of the research project, I will be gathering information from a number of sources depending on both the needs of the project and direction from my research advisor. However, the following are sources that I believe would be strong start for background reading:

- Research on cartograms, including [4], [2] and [3]
- Background readings on other protocols, including [10] and [11]
- Readings on methods used to quantify Gerymandering, including [5], [6] and [7]
- Cake cutting methods, including selected sections from [8] and [9]
- Background reading on different approaches to solving perfect information games of this type. This could include CFR from [12], and other readings as directed by my research advisor
- Background readings on statistical approaches to game AI's
- Further background readings on ILP, mostly from lecture slides and no specific research paper

Expected results of the research:

- Identify and implement way to create a mapping useful during interactive play that allows users to select districts and see the results of their selections on a real world state map
- Devise an efficient algorithm to run an AI to play the I-Cut-You-Freeze redistricting protocol for the geometric case. Included in these results would be an evaluation of which approach to creating this AI was the most effective.
- A website to both illustrate the protocol and gather data a good dataset from human play. This could guide future research into understanding human player behavior when playing games simulating these political protocols. Such a dataset could aid future research into psychology, political science and HCI by giving insight into how players behave under certain game scenarios or conditions.
- Better understanding of the second order effects these protocols can have when used in practice
- Increased awareness by both the general public and academia of redistricting protocols

Signature of your research advisor, signifying endorsement of the project and willingness to supervise and evaluate it:

_____

A bibliography of related work (all references belong here):

[1] Wesley Pegden and Ariel Procaccia. A Partisan Districting Protocol With Provably Nonpartisan Outcomes. Working paper.
[2] Sun S. A fast, free-form rubber-sheet algorithm for contiguous area cartograms. International Journal of Geographical Information Science. 2013 Mar 1;27(3):567-93.
[3] Sagar BD. Cartograms via mathematical morphology. Information Visualization. 2014 Jan;13(1):42-58.
[4] Keim DA, North SC, Panse C. Cartodraw: A fast algorithm for generating contiguous cartograms. IEEE Transactions on Visualization and Computer Graphics. 2004 Jan;10(1):95-110.
[5] Samuel S-H Wang. Three tests for practical evaluation of partisan gerrymandering. Stanford Law Review, 68(6):1263, 2016.
[6] Michael D McDonald and Robin E Best. Unfair partisan gerrymanders in politics and law: A diagnostic applied to six cases. Election Law Journal, 14(4):312–330, 2015.
[7] Schuck, Peter H. "The Thickest Thicket: Partisan Gerrymandering and Judicial Regulation of Politics." *Columbia Law Review* 87, no. 7 (1987): 1325-384. doi:10.2307/1122527.
[8] Steven J Brams and Alan D Taylor. Fair Division: From Cake-Cutting to Dispute Resolution. Cambridge University Press, 1996.
[9] Ariel D Procaccia. Cake cutting: Not just child's play. Communications of the ACM, 56(7):78–87, 2013.
[10] Zeph Landau, Oneil Reid, and Ilona Yershov. A fair division solution to the problem of redistricting. Social Choice and Welfare, 32(3):479–492, 2009.
[11] Zeph Landau and Francis Edward Su. Fair division and redistricting. AMS Special Sessions on The Mathematics of Decisions, Elections, and Games, pages 17–36, 2010.
[12] Zinkevich M, Johanson M, Bowling M, Piccione C. Regret minimization in games with incomplete information. InAdvances in neural information processing systems 2008 (pp. 1729-1736).