William Scott
Phong Huang
Alexander Wei

10 May 2021
MCMC 260

Topics in Simulated Annealing

1. Annealing on the Flip Chain

Presented by William Scott

In this presentation, we discuss districting plans through the toy example of grid graphs, explaining how the districting problem emphasizes sampling but also has a flavor of optimization. We explain the tendency of the flip chain to produce misshapen plans, and use simulated annealing to enforce a preference for shorter boundaries. We then delve into three phases of the annealing process: hot, cooling, cold. While the cooling phase is generally quite satisfactory, the hot and cold phases quickly become infeasible on grids of appreciable size; we provide intuition and loose computational bounds explaining why. We share a "successful" run on a 40x40 grid, and discuss reasons for pessimism on larger grids. Finally, we mention some extensions and modifications which have the potential for greater scalability.

The accompanying code implements simulated annealing on flip chains with strict connectivity and district size requirements, and contains some canned experiments showing that regime change is feasible for grid sizes up to around 40x40. I also experiment with a generalized flip chain (multiple reassignments and/or loosened connectivity requirements), which was feasible at small scales but failed to fix the inherent problems with the flip chain. Code can be found at <github.com/alexander-wei/MCMC-Annealing-2021>

2. Simulated Annealing: Aircraft trajectory planning

Presented by Phong Hoang

In this presentation, we will talk about the aircraft trajectory planning problem. Given the set of different trajectories where each trajectory is represented by its three-dimensional route and departure time, we want to reassign those routes as well as departure time so that the number of interactions is minimum. Solving this problem requires two steps which are modeling the problem and applying the simulated annealing method.

We will turn this problem into a combinatorial optimization problem. In order to do so, we must come up with an objective function as well as some constraints. First, we assume that there are only M waypoints that can be used to change the trajectory. The first constraint will be placed on the departure time. Since it is unacceptable to delay a flight for too long, the departure time shift is bounded above. Then, that time interval will be discretized as we are solving a combinatorial optimization problem. We also need to put a constraint on the route length extension as an increase in route length will lead to an increase in fuel consumption. Furthermore, we will have to place a constraint on waypoint locations to avoid sharp turns. Similar to what we have done with departure time, we will also discretize the possible locations of each waypoint. Our objective function will be the total number of interactions; an interaction happens if two trajectories are close at a given time. Because there will be a tremendous number of trajectories, the presentation will also talk about a quick way to detect interactions. The problem asks us to minimize the objective function, given some constraints.

In order to solve this problem, we will apply the simulated annealing method. First, we will talk about how to generate a neighborhood solution. We will choose a random trajectory i, if the number of interactions for this trajectory is smaller than a user-defined threshold, we will choose a different flight. Otherwise, we will modify i to get a neighborhood solution. Then, we will randomly choose a waypoint. After that, we will change that waypoint with a given probability; and we will change the departure time with a different given probability. For a given temperature, we will generate a constant Lk number of neighborhood solutions (Lk = 3500 in our case). The initial temperature will be the average variations of 100 different neighborhood solutions. The cooling schedule will follow geometrical law with the constant a = 0.99. Our algorithm will terminate if the final temperature Tf = c.T0, where $0 < c < 1$ is a user-defined constant.

3. Solving a Substitution Cipher: Locally Informed Proposals in Simulated Annealing

Presented by Alexander Wei

In the presentation, we define the substitution cipher as a combinatorial optimization problem and propose Simulated Annealing (SA) as a suitable solution technique. The discrete domain of possible cipherkeys is simple enough for us to gauge performance, and our natural grasp of language allows us to intuitively propose optimization techniques. With reference to literature on enhanced targeting of desired distributions via Metropolis Hastings (MH) MCMC, we have a reasonably simple domain of 26! permutations to present a few specialized techniques. The space of cipherkeys is much smaller than for a flipchain on a moderately sized grid, but we nonetheless confront issues like bottlenecks and the need for a targeted solution.

To provide the cipher scores for SA to improve on, we propose a Keyrank heuristic on the basis of chi-squared frequency comparison for choosing which letters to be reassigned with each MH step. The task of solving a cipher is posed in terms of matching letter-transition frequencies; amazingly, this frees us from having to check a word dictionary when reassigning cipherkey letters, instead making locally-informed key proposals that converge quickly. Following the suggestions of Grathwohl, et al. in “Oops I Took A Gradient” and Zanella in “Informed proposals for local MCMC in discrete spaces,” we choose to bias the MH walker toward reassignment of letters that have poor Keyrank, employing a two-stage SA process. Code is on github: <github.com/alexander-wei/MCMC-Annealing-2021>.

Works Cited

1. Annealing on the Flip Chain

Bertsimas, Dimitris et al. “Simulated Annealing.” 1993. *Statistical Science*. <https://www.mit.edu/~dbertsim/papers/Optimization/Simulated%20annealing.pdf>
The authors review early annealing methods on finite state spaces, and relate convergence it to bottlenecks in the objective function. They show asymptotic bounds on runtime but caution their practical use, and also indicate some applications where simulated annealing is competitive with heuristic and other approaches. These include graph partitioning and graph coloring.

Fifield, Benjamin et al. “Automated Redistricting Simulation using Markov Chain Monte Carlo.” December 2019. *Journal of Computational and Graphical Statistics*. <https://imai.fas.harvard.edu/research/files/redist.pdf>
The authors aim to uniformly sample from population-balanced, contiguous districting plans. Instead of single units, connected components of units are swapped between districts. The authors experiment with parallel tempering (multiple chains are run at different temperatures, temperatures are swapped between the chains). In practice, the proposal scheme is quite good at avoiding rejections due to connectivity problems. Finally, some other, more ad-hoc simulated redistricting algorithms are presented for comparison. Overall, the approach is quite principled but care is still required.

Johnson, David S. et al. “Optimization by Simulated Annealing: an Experimental Evaluation.” Feb. 1989. <http://faculty.washington.edu/aragon/pubs/annealing-pt1.pdf>
The authors treat annealing for optimal graph partitioning in detail. Limitations of the annealing approach are discussed, as well as the types of problems annealing can solve. The proposed flip chain allows but penalizes variation in component sizes, and is tested on a 500-vertex graph. The test graph is much less regular than the grids we investigate, but has similar density overall and is thus a good indication of what to expect. Key observations include (1) long cooling times are required, (2) allowing, but penalizing, deviations from acceptable solutions improves connectivity and performance, and (3) the particular choice of temperature decay is not particularly important.

2. Simulated Annealing: Aircraft Trajectory Planning

Daniel Delahaye, Supatcha Chaimatanan, Marcel Mongeau. Simulated annealing: From basics to applications. Gendreau, Michel; Potvin, Jean-Yves. Handbook of Metaheuristics, 272, Springer, pp.1- 35.ISBN 978-3-319-91085-7, 2019, International Series in Operations Research & Management Science (ISOR), 978-3-319-91086-4. ff10.1007/978-3-319-91086-4_1ff.

Supatcha Chaimatanan, Daniel Delahaye, Marcel Mongeau. Strategic deconfliction of aircraft trajectories. ISIATM 2013, 2nd International Conference on Interdisciplinary Science for Innovative Air Traffic Management, Jul 2013, Toulouse, France. p xxxx. Ffhal-00868450f.

Arianit Islami, Supatcha Chaimatanan, Daniel Delahaye. Large Scale 4D Trajectory Planning. Air Traffic Management and Systems – II , 2016, Lecture Notes in Electrical Engineering.

Supatcha Chaimatanan, Daniel Delahaye, Marcel Mongeau. A methodology for strategic planning of aircraft trajectories using simulated annealing. ISIATM 2012, 1st

International Conference on Interdisciplinary Science for Air traffic Management, Jun 2012, Daytona Beach, United States.

3. Solving a Substitution Cipher: Locally Informed Proposals in Simulated Annealing

Deford, Daryl, Duchin, Moon, et al. “Recombination: A family of Markov chains for redistricting.” Mar. 27, 2020. MGGG Redistricting Lab. <mggg.org/uploads/ReCom.pdf>.

The authors review the contemporary use of flip-chains in evaluating state districting as a problem in gerrymandering. A novel recombination approach is proposed: using spanning trees of a “dual graph,” they demonstrate that the innate clustering of populace becomes an advantage under the new approach that appropriately targets natural districting structure. We take cue from Deford, Duchin, et al. to lend a similar focus on the substitution cipher task in creating informed cipherkey proposals.

Garg, Poonam. “Evolutionary Computation Algorithms for Cryptanalysis: A Study.” *IJCSIS International Journal of Computer Science and Information Security*, Vol. 7, No. 1, 2010. Arxiv. <arxiv.org/pdf/1006.5745>.

Garg reviews several “evolutionary” techniques for combinatorial cryptography problems. They involve improvement of a proposed solution in search of a global optimum and are well-suited to MCMC. With respect to the methods covered here, the substitution cipher affords a unique perspective of attack; we infer that simulated annealing is an optimal algorithm since we are able to track decipher progress on-the-go.

Grathwohl, Will, et al. “Oops I Took A Gradient: Scalable Sampling for Discrete Distributions.” 2021. Arxiv. <arxiv.org/pdf/2102.04509>.

The authors propose a gradient-based proposal scheme for Metropolis-Hastings (MH) algorithms. Their highlighted “Gibbs with Gradients” proposal is compared with several other locally-informed proposal methods that can be used in biasing a MH walker toward a region of interest. Particular emphasis is placed on the preservation and interpretation of innate topological structure to the walker.

Ireland, Michael, et al. “Monte-Carlo Imaging for Optical Interferometry.” 2007. Arxiv. <arxiv.org/pdf/2007.00716>.

The authors present a simulated annealing approach to image reconstruction by optical scans. These are the data-sets returned by an X-ray scan at the dentist's office or the full-body scanners at the airport. Just like with solving a substitution cipher, the task is to minimize the chi-squared scores of the proposed solution. Some peculiarities of the annealing approach are either covered or posed as questions for further review. For example: how would we fully automate this process? What types of image scans should be favored by "regularization," and how do we design a regularization procedure that is consistent with a typical scan?

Zanella, Giacomo. "Informed proposals for local MCMC in discrete spaces." Nov. 21, 2017. Arxiv. <arxiv.org/pdf/1711.07424.pdf>.

Zanella presents slow convergence as a characteristic problem of the "blind" MH walk. The paper is tasked with tilting a walker toward the region of interest; Zanella reviews some locally-informed techniques and devises a universal framework of "pointwise informed" proposals that is adaptable to both the continuous and discrete settings of MH.