

Project 2: Evolve a Density Classification CA

Code due Wednesday March 4th (any time before midnight)

Reports due at 12pm (noon) Friday March 6th in the CS front office

Follow the guidelines on the Assignments section of the webpage for readme files, using nice, turnin and the report template (from Project 1, with appropriate section number modifications).

Your task is to evolve a one-dimensional CA for density classification, following Mitchell et al 1994. To start, set the radius of your CA to 2 (CA neighborhood size = 5). Use as your CA initial configuration (IC) a bitstring of length 121 (Note that Mitchell uses an IC of length 149). Run your CA for (at most) 200 time steps (you will save computation time by stopping the CA when it has converged to an unchanging state). Repeat for a CA with radius 3 (CA neighborhood size = 7).

Define the genotype of your GA as a bitstring that specifies the CA rule set (ϕ) following specification (5.1) on page 8 of Mitchell et al 1994 (you will have a different length genotype for the radius 2 and radius 3 CA). If you are using CS Dept. Linux machines, limit your GA to a number of evaluations equivalent to population size 100, 50 generations and a fitness function that evaluates at most 50 ICs per rule set. If you have other computing resources, you can run longer, but provide baseline results at 250,000 evaluations.

Complete your report using the template from Project 1. Complete all sections of the report template. Instructions for Project 2 for Methods, Results and Discussion are below. Each section will be worth twice the number of points for Project 1 (i.e., Abstract is 10 points, Introduction is 10 points, Results are 120 points, etc.)

Specify in Methods (Section 4)

- 4.1 The parameters of your GA: your method for initializing the initial generation of genotypes (which specify the rule set ϕ), your selection method, fitness function and any form of elitism that you use. If you deviate from the methods in Mitchell 1994, specify how. Describe any techniques you use to speed up the evolutionary process or find better (or different) solutions.
- 4.2 Also describe how you determine the 50 CA ICs that you use to evaluate fitness, and any techniques you use to improve results.
- 4.3 Describe how you analyzed the data you present in Results.

Present Results in Section 2

Section 2.1: Analyze your results

Compare results from the radius 2 and radius 3 CAs. Produce figures similar to Figures 1,2, 3 and 11 in Mitchell et al 1994. Provide a meaningful caption for each figure, and explain what these figures mean in text.

Additionally analyze whether there is a relationship between the fitness of a rule and the length of the transient period before the CA settles into a fixed point. Are the dynamics of fit rules periodic, chaotic or something else prior to converging on a fixed point?

Section 2.2 Analyze mutational robustness

Explain what fraction of the rule space you could explore given the imposed GA limit for radius 2 and radius 3 CAs.

Analyze the mutational robustness of your genotype. Explain whether there are regions in your genotype landscape that are mutationally robust, meaning that similar genotypes have similar fitness values.

Choose 1 genotype with high fitness. Sample the fitness of strings in the mutational neighborhoods (as described by Wagner 2012) surrounding that genotype. You should sample from neighborhoods from 1 to at least 5 mutations away from your starting genotype. (Note that the mutational neighborhoods of the genotype are entirely different from the neighborhood of a cell in a CA). Provide a figure that shows how mean and maximum fitness changes as you sample neighborhoods that are a greater distance from your chosen genotype. Describe whether it is possible to traverse from your chosen genotype to a distant genotype (one a large hamming distance away) with similar fitness without substantially reducing fitness. (You may define “large”, “similar” and “substantially” to describe the most interesting case you find.)

Section 2.3 Improve the performance of your density classification CA.

The following are suggestions for how you might attempt to improve your CA (you may choose either the radius = 2 or radius = 3 CA). You may choose one of the suggestions below and investigate it thoroughly, two or more and compare their effect on performance or you may choose some other method. Provide 1 to 3 figures to demonstrate your results.

-Focus your evaluations on the best rule sets. Mitchell does this by evaluating a small subset of initial conditions on every CA, and a large set of initial conditions for the elite CAs. You may also test a very small number of ICs to eliminate obviously bad rule sets

-Reduce useless rule set evaluations by eliminating rule sets with lambda values that are unlikely to improve performance. Eliminate (or reduce) repetitive evaluation of the same rules in the GA by storing rules that have been evaluated. Use those computational cycles to analyze rule sets that are likely to improve fitness.

-Evaluate some feature of the CA dynamics for a small set of CA initial conditions to evaluate fitness, rather than evaluating many initial conditions.

-Dynamically change the fitness function or the procedure by which you choose CA initial conditions as you evolve.

- Use some method other than a GA to search for good CA density classification rules

- Search the literature for other attempts at density classification and repeat (or improve) their methods. List those papers in References and refer to them in Results.

Discussion (Section 3)

Summarize your results and discuss the most interesting finding in each of your results sections in the corresponding discussion section 3.1, 3.2 and 3.3. Each section should be approximately 2 paragraphs.