Project 2

Creation of Spatial Structure by Activation/Inhibition Cellular Automaton

Due September 25, 2009

Background

Activation/Inhibition Cellular Automaton

In this project you will investigate and measure the creation of spatial structure by an activation/inhibition CA (AICA) such as we discussed in class.[1] Structure will be quantified in terms of *spatial correlation*, *joint entropy*, and *mutual information* (discussed below), which you will measure after the CA has converged to a stable state.

Recall that the state transition rule of an activation/inhibition network is given by a formula such as this:

$$s_i(t+1) = \text{sign}\left[h + J_1 \sum_{r_{ij} < R_1} s_j(t) + J_2 \sum_{R_1 \leq r_{ij} < R_2} s_j(t) \right]. \tag{1}$$

Since this is a 2D CA, the cell indices are two-dimensional vectors, $\mathbf{i} = (i_1, i_2)$, $\mathbf{j} = (j_1, j_2)$. As usual, we will also assume that the space is a torus, that is, the indices wrap around on the top and bottom and on the left and right edges. For the purposes of this project you may assume that $J_1 \geq 0$ and $J_2 \leq 0$ (which are the usual cases). Also, you should assume that the R_1 neighborhood includes the cell ($\mathbf{i}$) at its center.

The distance r_{ij} between cells can be measured in many different ways, for example by Euclidean distance (also known as the L_2 metric). For the purpose of this project it will be more convenient to measure distance by the L_1 metric, which is defined:

$$r_{ij} = |i_1 - j_1| + |i_2 - j_2|. \tag{2}$$

This gives neighborhoods that are diamond-shaped, rather than circular, but that doesn't really matter.

---

[1]Additional information and tips are available at Kristy's website: <http://www.cs.utk.edu/~kvanhorn/cs594_bio/project2/aica.html>.

Spatial Correlation

We are interested in the extent to which the states of cells at various distances are correlated to each other. Therefore, we define the *absolute spatial correlation at distance l* as

$$\rho_l = |\langle s_\mathbf{i} s_\mathbf{j} \rangle - \langle s_\mathbf{i} \rangle \langle s_\mathbf{j} \rangle| \text{ for all } \mathbf{i}, \mathbf{j} \text{ such that } r_{\mathbf{ij}} = l .\text{[2]}$$

The angle brackets mean "average value of." To understand this formula, suppose that the states $+1$ and -1 are equally frequent in the space. Then the individual state averages are 0: $\langle s_\mathbf{i} \rangle = \langle s_\mathbf{j} \rangle = 0$. Therefore $\langle s_\mathbf{i} s_\mathbf{j} \rangle$ is the average value of the product $s_\mathbf{i} s_\mathbf{j}$ for cells that are a distance l apart. If these cells tend to be $+1$ at the same time, or -1 at the same time, this average will be greater than zero (positive correlation). If they tend to have opposite signs, then it will be less than zero (negative correlation). If they tend to have the same sign as often as the opposite sign, then the average will be near zero (no correlation). By subtracting $\langle s_\mathbf{i} \rangle \langle s_\mathbf{j} \rangle$ we compensate for an overall bias toward positive or negative states (such as we get when $h \neq 0$). We take the absolute value, because we are not interested in whether the spatial correlation is positive or negative, only its magnitude. Note that $\rho_0 = 1 - \langle s_\mathbf{i} \rangle^2$.

Next let's consider more explicitly how to compute ρ_l. Suppose there are N^2 cells in the space, and let C_l ($1 \leq l < N/2$) be the circumference (in number of cells) of a neighborhood of radius l. For the L_1 metric, $C_l = 4l$. Thus there are C_l cells at a distance l from a given cell. Then we can see (make sure you really do see it!) that:

$$\rho_l = \left| \frac{2}{N^2 C_l} \sum_{\substack{\langle \mathbf{ij} \rangle \\ r_{\mathbf{ij}} = l}} s_\mathbf{i} s_\mathbf{j} - \left(\frac{1}{N^2} \sum_\mathbf{i} s_\mathbf{i} \right)^2 \right| .$$

The notation $\langle \mathbf{ij} \rangle$ under the summation means "all pairs of distinct cells $\mathbf{i}$ and $\mathbf{j}$" (taking each pair just once); therefore there are $N^2 C_l / 2$ of these pairs. (There is a diamond of C_l cells around each of the N^2 cells.) Thus we are averaging over all pairs at a distance of l. For purposes of computation, this can be made more explicit:

$$\rho_l = \left| \left(\frac{1}{4 l N^2} \sum_{i_1, i_2} \left(s_{i_1, i_2} \sum_{|j_1| + |j_2| = l} s_{i_1 + j_1, i_2 + j_2} \right) \right) - \left(\frac{1}{N^2} \sum_{i_1, i_2} s_{i_1, i_2} \right)^2 \right| .$$

Notice that the second summation is over positive and negative j_1, j_2 in the

---

[2] Note: A true correlation coefficient is normalized to [–1, 1] by dividing by standard deviations of the variables: $\rho_l = \left| \left(\langle s_\mathbf{i} s_\mathbf{j} \rangle - \langle s_\mathbf{i} \rangle \langle s_\mathbf{j} \rangle \right) / \sigma_{s_\mathbf{i}} \sigma_{s_\mathbf{j}} \right|$ for all $\mathbf{i}$, $\mathbf{j}$ such that $r_{\mathbf{ij}} = l$. You can compute it this way if you want, but make sure to tell us.

range $-l \le j_1, j_2 \le l$. The coefficient on the first summation is:

$$\frac{1}{2} \frac{2}{N^2} \frac{1}{C_l} = \frac{1}{4 l N^2} .$$

The $\frac{1}{2}$ factor is to avoid double counting the pairs $\langle \mathbf{ij} \rangle$. Make sure that you understand the formula for ρ_l . Remember that cell indices wrap around both vertically and horizontally.

Characteristic Correlation Length

(This section is relevant to r**equired measurements for CS 594 students**. CS 420 students can skip it, unless they want to do these measurements for extra credit.)

Often spatial correlation decreases approximately exponentially with distance, $\rho_l \propto e^{-l/\lambda}$, where λ , the *characteristic correlation length*, measures how quickly spatial correlation decreases with distance. By assuming that spatial correlation is exponentially decreasing, we can estimate λ . Let α be an arbitrary constant of proportionality $\rho_l = \alpha e^{-l/\lambda}$. Then,

$$\rho_0 = \alpha e^0 = \alpha ,$$

$$\rho_\lambda = \alpha e^{-\lambda/\lambda} = \alpha e^{-1} = \rho_0 \Big/ e .$$

Therefore we can estimate $\lambda \approx l$ such that $\rho_l = \rho_0 / e$. That is, the characteristic correlation length is that length at which the spatial correlation has decreased to $1/e$ of its maximum value ρ_0 . (*Characteristic correlation time* can be defined similarly.)

Mutual Information Between Distant Cells

As discussed in class, one way to measure the correlation between cell states is by *average mutual information*. The average mutual information between cells at a distance l is related to the joint entropy between cells at this distance. Therefore, first define the average entropy $H(S)$ of the cellular space S :

$$H(S) = - \sum_{s \in \{-1,+1\}} \Pr\{s\} \lg \Pr\{s\} .$$

Remember that we take $0 \lg 0 = 0$. For state counting convenience, let $\beta(s) = (1+s)/2$ so that $\beta(+1) = 1$ and $\beta(-1) = 0$ (β converts a *bipolar* number $\in \{-1,+1\}$ into a *binary* number $\in \{0,1\}$). Then the probabilities of the states can be computed:

$$\Pr\{+1\} = \frac{1}{N^2} \sum_i \beta(s_i),$$

$$\Pr\{-1\} = 1 - \Pr\{+1\}.$$

The probability that two cells at distance l have state $+1$ is:

$$P_l(+1,+1)=\frac{2}{N^2 C_l}\sum_{\substack{\langle \mathbf{ij}\rangle \\ r_{\mathbf{ij}}=l}}\beta(s_{\mathbf{i}})\beta(s_{\mathbf{j}})\ .$$

Similarly,

$$P_l(-1,-1)=\frac{2}{N^2 C_l}\sum_{\substack{\langle \mathbf{ij}\rangle \\ r_{\mathbf{ij}}=l}}\beta(-s_{\mathbf{i}})\beta(-s_{\mathbf{j}})\ .$$

These summations can be computed in the same way as ρ_l . Finally,

$$P_l(+1,-1)=P_l(-1,+1)=1-P_l(+1,+1)-P_l(-1,-1)\ .$$

(Notice that there are only three distinct possibilities for cells at distance l : both +1, both –1, or opposite signs.) The joint entropy between cells at a distance l , H_l , is then defined in terms of the probabilities in the usual way:

$$H_l=-\left(P_l(+1,+1)\lg P_l(+1,+1)+P_l(-1,-1)\lg P_l(-1,-1)+P_l(+1,-1)\lg P_l(+1,-1)\right)\ .$$

The *average mutual information* between two sources A and B is defined $I(A,B)=H(A)+H(B)-H(A,B)$. Therefore, the average mutual information between cells at distance l is defined:

$$I_l=2H(S)-H_l\ .$$

4

Experiments

In this project you will be investigating how spatial structure (as measured by spatial correlation, mutual information, etc.) is affected by the parameters (J_1, J_2, R_1, R_2, h) of an AICA. (Graduate students should also compute characteristic correlation length λ; undergraduates can do this for extra credit.) Make sure to check Kristy's website for information and suggestions on doing your project:

<http://www.cs.utk.edu/~kvanhorn/cs594_bio/project2/aica.html>.

The goal is to get the measures of structure for a variety of different combinations of parameters for each of the following three experiments. The following table shows the minimum set of parameter combinations that you should use. You can try additional combinations to earn some extra credit.

Your CA space should be $N = 30$ in each dimension. Use asynchronous updating (i.e., update the cells one at a time rather than all at once), and update the states randomly, not in a systematic order, or you will get artifacts.

For each set of parameters you will make several runs, with different random initial states, so that you can average your results. Do a few trial runs to see how much the measures vary from run to run. If they don't vary much, then you don't need to do many runs for each set of parameters (maybe one run will be enough). If they do vary, then you will have to make several runs (less than 5) and average them together, and you won't have to investigate so many parameter values.

For each run you should allow the CA to stabilize, and then compute $H(S)$ estimated λ (this is optional for CS 420), and also plot H_l, I_l, and ρ_l against the l values. For each set of parameter values, compute the average of all of these and print them out (generate plots for H_l, I_l, and ρ_l). You don't need to hand in a graph for every set of parameters, just those that are relevant to your hypotheses.

Following are the descriptions of specific experiments that you should try. In each case, discuss and explain your results.

Descriptions of Experiments:

1. (**Optional for CS 420**; will earn extra credit.) Set $J_2 = 0$ to disable the inhibition system and set $J_1 = 1$. Quantify the spatial structure for at least the specified range of R_1 and h values.

2. (**Optional for CS 420**; will earn extra credit.) Set $J_1 = 0$ to disable the activation system and set $J_2 = -0.1$. Quantify spatial structure as in Experiment (1) for at least the specified parameter values.

3. Set $J_1 = 1$ and $J_2 = -0.1$. Investigate, as in the preceding experiments, the spatial structure for at least the specified values of R_1, R_2, and h.

Based on all the preceding experiments, draw conclusions about the dependence of spatial correlation, joint entropy, and mutual information on the parameters of the AICA. See if you can draw some quantitative conclusions.

Experiment 1 (required for 594)			Experiment 2 (required for 594)			Experiment 3 (required for 420 & 594)		
R_1	R_2	h	R_1	R_2	h	R_1	R_2	h
1	15	-1	1	2	0	1	2	0
3	15	-1	1	4	-2	1	5	-4
6	15	-2	1	4	-1	1	5	-2
			1	4	0	1	5	0
			1	6	-5	1	9	-6
			1	6	-3	1	9	-3
			1	6	0	1	9	0
			1	9	0	1	14	0
			1	13	0	3	5	-1
			4	5	0	3	5	0
			4	7	-5	3	9	-6
			4	7	-3	3	9	-3
			4	7	0	3	9	0
			4	12	0	3	14	0
			9	12	-6	7	9	0
			9	12	-3	7	14	0
			9	12	0	12	14	0

Convergence of AICA

Problem 1 in this part of the project is **required for CS 594** students and **extra-credit for CS 420 students**; Problem 2 is **extra-credit for everyone**. Consider again the state update rule (Eq. 1, p. 1). Recall that $r_{\mathbf{ij}}$ represents the distance between cells $\mathbf{i}$ and $\mathbf{j}$, so the first summation is over all cells within a distance of R_1 to cell $\mathbf{i}$, and the second summation is over all cells with a distance from R_1 to R_2 (Eq. 2, p. 1). For simplicity in this part of the project, assume that the R_1 neighborhood does *not* include the center cell $\mathbf{i}$.

The state of a CA can be updated either *synchronously* or *asynchronously*. Recall that with synchronous updating all the states are updated simultaneously. With asynchronous updating the cells are updated one at a time (usually in some random order).

This part of the project explores the convergence of this AICA; that is, does it inevitably reach a stable state?

Problems

Problem 1 (required for 594, extra-credit for 420)

Prove, by exhibiting a counter-example, that if synchronous updating is used, then the AICA may not reach a stable state.

Hint: Construct a very simple AICA, obeying the above state update equation (Eq. 1), that cycles between two different states.

Problem 2 (extra-credit for both CS 420 and CS 594)

Prove that if the states are updated asynchronously, then the AICA must reach a stable state.

Hint: Define the following function (called an *energy* or *Lyapunov* function) of the total state of an AICA:[3]

$$E[s(t)] = -\frac{1}{2}\sum_{\mathbf{i}} s_{\mathbf{i}}(t)\left[h + J_1 \sum_{r_{\mathbf{ij}} < R_1} s_{\mathbf{j}}(t) + J_2 \sum_{R_1 \le r_{\mathbf{ij}} < R_2} s_{\mathbf{j}}(t)\right].$$

Show that updating any single cell, according to the state update rule, cannot increase this function (that is, $\Delta E \le 0$). What else do you need to show in order to guarantee convergence to a stable state?

Additional Extra Credit: Assume that the R_1 neighborhood *does* include the center cell, and explore any additional assumptions that might be needed to guarantee convergence.

---

[3] For this energy function, look in Bar-Yam on p. 630 (section 7.2.2) and p. 170 (sec. 1.6.6).