CS 100: Simulation and Randomness

Chris Kauffman

Week 12-1
Logistics

HW6 Up Soon

- Last of the Semester
- Security Upgrades
- AI and Automation questions

Reading:

- Pattern on the Stone Ch 8 Machines that Learn
- Article: A Plan for Spam

End Game

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 12</td>
<td>Tue 4/18</td>
<td>Simulation</td>
</tr>
<tr>
<td></td>
<td>Thu 4/20</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>Week 13</td>
<td>Tue 4/25</td>
<td>Machine Learning</td>
</tr>
<tr>
<td></td>
<td>Thu 4/27</td>
<td>Automation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mini-Exam 4</td>
</tr>
<tr>
<td>Week 14</td>
<td>Tue 5/2</td>
<td>Intellectual Property Review</td>
</tr>
<tr>
<td></td>
<td>Thu 5/4</td>
<td>HW 6 Due</td>
</tr>
<tr>
<td>Week 15</td>
<td>Thu 5/11</td>
<td>Final Exam</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10:30am-1:15pm</td>
</tr>
</tbody>
</table>
1. True or False: The number of transistors in modern computers is increasing. (If so, by how much)

2. True or False: Modern computers continue to increase in speed dramatically.

3. True or False: With 5 computers, one can always get a program to finish in 1/5th the time

4. If False: why? What is the slow down?

5. What is a histogram?
Simulations

- An important application of computation
- One of its earliest uses: artillery firing tables
- Lots of modern incarnations, many of which use parallel computation
- We will explore a few today
Schilling’s Segregation Model

- White is empty space
- City people are little red/blue squares, different "classes"
- People want to have neighbors w/ same class
- Tolerance threshold for like to not-like neighbors
- Too few like neighbors, move to a random empty location

<table>
<thead>
<tr>
<th></th>
<th>Like</th>
<th>Unlike</th>
<th>Total</th>
<th>Ratio Like</th>
<th>Percentage</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Left</strong></td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>1/6</td>
<td>16.6%</td>
<td>30%</td>
</tr>
<tr>
<td><strong>Result Move</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Like</th>
<th>Unlike</th>
<th>Total</th>
<th>Ratio Like</th>
<th>Percentage</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Right</strong></td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>3/7</td>
<td>42.9%</td>
<td>30%</td>
</tr>
<tr>
<td><strong>Result Stay</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For Fun Disease Model

- Play with http://mattbierbaum.github.io/zombies-usa/
- Read about what is in the model here: http://arxiv.org/abs/1503.01104
- Consider what’s good and bad about the model

Figure out

- What is the parameter $\alpha$ described as the "kill to bite ratio"
- What can one learn by playing with the model?
Other Models

All models are wrong, but some are useful.
– George Box, Statistician

What other computer simulations affect your life?

➤ List at least 2
➤ What do you think is involved with the codes there?
➤ Could you run that code on your laptop?
Examples

Weather Prediction

Auto Safety During Crashes

Any volunteers to replace the simulated dummy?

Siri: Will it rain this afternoon?
More Examples

Wireless Network Simulation

Source: Higashino Lab

Protein Structure Prediction

Source: Fold.it

Puzzle game that simulations biology, protein folding, disease and drug development research.

Can you hear me now?
Randomness

Chance Events

- Segregation: location to move to is random
- Chance events are good for games too

Google Me This

- Are computers random at all?
- How does a computer produce random numbers or random bits?
- Are they really random numbers?
Pseudo-Random

A classic random number generator from *The C Programming Language* by Kernighan and Ritchie

/* Tracks state of random number generator */
unsigned long int next = 1;

/* rand: return pseudo-random integer on 0..32767 */
int rand() {
    next = next * 1103515245 + 12345;
    return (unsigned int)(next/65536) % 32768;
}

/* srand: set seed for rand() */
void srand(unsigned int seed) {
    next = seed;
}

Is there anything random about it?
Example Program

Random Draws

/* Draw 10 cards from a deck of 52 cards with replacement */
int main(){
    srand(12345);
    int i;
    for(i=0; i<20; i++){
        int myrand = rand() % 52;
        printf("%2d ",myrand);
    }
    printf("\n");
}

Always produces the sequence

44  4  17  14  27  29  33  6  38  41
12  43  34  27  31  24  25  28  33  8

Code Explained

▶ rand() produces a positive number
▶ Range 0 to 32767
▶ Limit to 0 to 51 by dividing by 52 and taking remainder
Random Numbers

- Pseudo-random sequences are deterministic: always produce the same sequence if you start in the same spot.
- If you want a different sequence, start somewhere different.
- Most common trick: use the time of day to seed the random sequence.
- Since time changes all the time (smirk) will get you different looking random sequences.
Rate Anything!

- You’re Running a Business Online
- Free web service
- RateAnything.com (currently available for purchase)
- Users can submit a person, place, or thing and start rating and feedback
- To make money: sell ads
- How do you choose what ads to sell?

Source: RateMyProfessor.com
Data Associations

Rating A Person

CS Professor  Ads for programming books
Politician    Political ads
Dentist       Ads for Toothpaste, electric toothbrush

Rating a Place

Vacation Hotel Ads for local tour, restaurant
Restaurant   Ads for other restaurants, coupon books

Rating a Thing

Frying Pan   Ads for tongs, spatula, new oven
New Car      Ads for insurance, car wash
Getting Paid

- You get paid by advertisers every time people click on your ads
- Want people to click as much as possible
- Try to make ads relevant to
  - Object being rated
  - User interests
- Any idea how do do this with a computer?
Machine Learning Can Help

- Algorithms that can learn patterns
  - This object relates to this ad
  - This user is interested in this stuff
- Requires **information** in a machine friendly form
  - List of words in rated object description
  - List of objects user has viewed
  - List of ads user has clicked on, which they have ignored
- Most machine learning requires **training**
  - Explicitly label "this object is like this object", "this ad should be served for this object", "the user clicked on this ad and ignored this ad"
  - Can learn the patterns so that a new ad can be associated to new rated objects
- **Chapter 8 of The Pattern on the Stone** discusses some machine learning, more next time