P o C S What's The Story?

Principles of Complex Systems, CSYS/MATH 300
University of Vermont, Fall 2018
Assignment 7 • code name: Beginner Pottery

Dispersed: Thursday, October 12, 2018.
Due: 11:59 pm, Friday, October 19, 2018.
Some useful reminders:
Deliverator: Prof. Peter Dodds (peter.dodds@uvm.edu)
Assistant Deliverator: David Dewhurst (david.dewhurst@uvm.edu)
Office: Farrell Hall, second floor, Trinity Campus
Office hours: 10:15 am to 11:30 am, Tuesday and Thursday, and 2:00 pm to 4:00 pm, Wednesday
Course website: http://www.uvm.edu/pdodds/teaching/courses/2018-08UVM-300
Bonus course notes: http://www.uvm.edu/pdodds/teaching/courses/2018-08UVM-300/docs/dewhurst-pocs-notes.pdf

---

All parts are worth 3 points unless marked otherwise. Please show all your workingses clearly and list the names of others with whom you collaborated.

Please obey the basic life rule: Never use Excel. Or any Microsoft product except maybe Xbox (which sadly will likely not help you here.)

Graduate students are requested to use LaTeX (or related TeX variant).

email submission: 1. Please send to david.dewhurst@uvm.edu.
2. PDF only! Please name your file as follows (where the number is to be padded by a 0 if less than 10 and names are all lowercase): `CSYS300assignment%02d$firstname-$lastname.pdf` as in `CSYS300assignment06michael-palin.pdf`

Please submit your project's current draft in pdf format via email. Please use this file name format (all lowercase after CSYS):
`CSYS300project-$firstname-$lastname-YYYY-MM-DD.pdf` as in
`CSYS300project-lisa-simpson-1989-12-17.pdf` where the date is the date of submission (and not, say, your birthdate).

---

1. (3 + 3 + 3)

   Highly Optimized Tolerance:

   This question is based on Carlson and Doyle's 1999 paper "Highly optimized tolerance: A mechanism for power laws in design systems" [1]. In class, we made our way through a discrete version of a toy HOT model of forest fires. This paper revolves around the equivalent continuous model's derivation. You do not have to perform the derivation but rather carry out some manipulations of probability distributions using their main formula.

Our interest is in Table I on p. 1415:

$p(x)$	$p_{\text{cum}}(x)$	$P_{\text{cum}}(A)$
$x^{-(q+1)}$	x^{-q}	$A^{-\gamma(1-1/q)}$
e^{-x}	e^{-x}	$A^{-\gamma}$
e^{-x^2}	$x^{-1}e^{-x^2}$	$A^{-\gamma}[\log(A)]^{-1/2}$

and Equation 8 on the same page:

$$P_{\geq}(A) = \int_{p^{-1}(A^{-\gamma})}^{\infty} p(\mathbf{x})\mathrm{d}\mathbf{x} = p_{\geq}\left(p^{-1}\left(A^{-\gamma}\right)\right),$$

where $\gamma = \alpha + 1/\beta$ and we'll write $P_{\geq}$ for P_{cum}.

Please note that $P_{\geq}(A)$ for $x^{-(q+1)}$ is not correct. Find the right one!

Here, $A(\mathbf{x})$ is the area connected to the point $\mathbf{x}$ (think connected patch of trees for forest fires). The cost of a 'failure' (e.g., lightning) beginning at $\mathbf{x}$ scales as $A(\mathbf{x})^{\alpha}$ which in turn occurs with probability $p(\mathbf{x})$. The function p^{-1} is the inverse function of p.

Resources associated with point $\mathbf{x}$ are denoted as $R(\mathbf{x})$ and area is assumed to scale with resource as $A(\mathbf{x}) \sim R^{-\beta}(\mathbf{x})$.

Finally, $p_{\geq}$ is the complementary cumulative distribution function for p.

As per the table, determine $p_{\geq}(x)$ and $P_{\geq}(A)$ for the following (3 pts each):

(a) $p(x) = cx^{-(q+1)}$,

(b) $p(x) = ce^{-x}$, and

(c) $p(x) = ce^{-x^2}$.

Note that these forms are for the tails of p only, and you should incorporate a constant of proportionality c, which is not shown in the paper.

2. The discrete version of HOT theory:

From lectures, we had the following.

Cost: Expected size of 'fire' in a d-dimensional lattice:

$$C_{\text{fire}} \propto \sum_{i=1}^{N_{\text{sites}}} p_i a_i$$

where a_i = area of ith site's region, and p_i = avg. prob. of fire at site i over a given time period.

The constraint for building and maintaining $(d-1)$-dimensional firewalls in d-dimensions is

$$C_{\text{firewalls}} \propto \sum_{i=1}^{N_{\text{sites}}} a_i^{(d-1)/d} a_i^{-1},$$

where we are assuming isometry.

Using Lagrange Multipliers, safety goggles, rubber gloves, a pair of tongs, and a maniacal laugh, determine that:

$$p_i \propto a_i^{-\gamma} = a_i^{-(1+1/d)}.$$

3. (3 + 3 + 3 + 3)

   A courageous coding festival:

   Code up the discrete HOT model in 2-d. Let's see if we find any of these super-duper power laws everyone keeps talking about. We'll follow the same approach as the $N = L \times L$ 2-d forest discussed in lectures.

   Main goal: extract yield curves as a function of the design D parameter as described below.

   Suggested simulations elements:

   - Take $L = 32$ as a start. Once your code is running, see if $L = 64$, 128, or more might be possible. (The original sets of papers used all three of these values.) Use a value of L that's sufficiently large to produced useful statistics but not prohibitively time consuming for simulations.
   - Start with no trees.
   - Probability of a spark at the (i,j)th site: $P(i,j) \propto e^{-i/\ell} e^{-j/\ell}$ where (i,j) is tree position with the indices starting in the top left corner ($i, j = 1$ to L). (You will need to normalize this properly.) The quantity ℓ is the characteristic scale for this distribution. Try out $\ell = L/10$.
   - Consider a design problem of $D = 1$, 2, L, and L^2. (If L and L^2 are too much, you can drop them. Perhaps sneak out to $D = 3$.) Recall that the design problem is to test D randomly chosen placements of the next tree against the spark distribution.
   - For each test tree, compute the average forest fire size over the full spark distribution:

     $$\sum_{i,j} P(i,j) S(i,j),$$

where $S(i, j)$ is the size of the forest component at (i, j). Select the tree location with the highest average yield and plant a tree there.

- Add trees until the 2-d forest is full, measuring average yield as a function of trees added.
- Only trees within the cluster surrounding the ignited tree burn (trees are connected through four nearest neighbors).

(a) Plot the forest at (approximate) peak yield.

(b) Plot the yield curves for each value of D, and identify (approximately) the peak yield and the density for which peak yield occurs for each value of D.

(c) Plot distributions of tree component sizes S at peak yield. Note: You will have to rebuild forests and stop at the peak yield value of D to find these distributions. By recording the sequence of optimal tree planting, this can be done without running the simulation again.

(d) Extra level: Plot size distributions for $D = L^2$ for varying tree densities $\rho = 0.10, 0.20, \ldots, 0.90$. This will be an effort to reproduce Fig. 3b in [2].

Hint: Working on un-treed locations will make choosing the next location easier.

References

[1] J. M. Carlson and J. Doyle. Highly optimized tolerance: A mechanism for power laws in designed systems. *Phys. Rev. E*, 60(2):1412–1427, 1999. pdf

[2] J. M. Carlson and J. Doyle. Highly optimized tolerance: Robustness and design in complex systems. *Phys. Rev. Lett.*, 84(11):2529–2532, 2000. pdf