PoCS What's The Story?

Principles of Complex Systems, Vols. 1 and 2
CSYS/MATH 6701, 6713
University of Vermont, Fall 2025
"This literally means nothing to me"
Assignment 09

It's Always Sunny in Philadelphia: The Gang Tries Desperately to Win an Award, S9E03
Episode links: IMDB, Fandom, TV Tropes.

Due: Friday, November 7, by 11:59 pm
https://pdodds.w3.uvm.edu/teaching/courses/2025-2026pocsverse/assignments/09/
Some useful reminders:
Deliverator: Prof. Peter Sheridan Dodds (contact through Teams)
Office: The Ether and/or Innovation, fourth floor
Office hours: See Teams calendar
Course website: https://pdodds.w3.uvm.edu/teaching/courses/2025-2026pocsverse
Overleaf: LaTeX templates and settings for all assignments are available at
https://www.overleaf.com/read/tsxfwwmwdgxj.

Some guidelines:

1. Each student should submit their own assignment.
2. All parts are worth 3 points unless marked otherwise.
3. Please show all your work/workings/workingses clearly and list the names of others with whom you ~~conspired~~ collaborated.
4. We recommend that you write up your assignments in LaTeX (using the Overleaf template). However, if you are new to LaTeX or it is all proving too much, you may submit handwritten versions. Whatever you do, please only submit single PDFs.
5. For coding, we recommend you improve your skills with Python. And it's going to be a no for the catachrestic Excel. Please do not use any kind of AI thing unless directed. The (evil) Deliverator uses (evil) Matlab.
6. There is no need to include your code but you can if you are feeling especially proud.

Assignment submission:

Via **Brightspace** (which is not to be confused with the death vortex of the same name, just a weird coincidence). Again: One PDF document per assignment only.

Please submit your project's current draft in pdf format via Brightspace.

Main thing this week: Continue with your projects.

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. (3 + 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.

(We are changing notation from the original papers where N is the length of the lattice. Terrible.)

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

Work on fire size rankings instead of probability distributions, whose exponents are connected with $\alpha = 1/(\gamma - 1)$.

Suggested simulation 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 design problems with $D = 1$, 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).
- Hint: Working on un-treed locations will make choosing the next location easier.

(a) For each value of D, plot the forest at (approximate) peak yield.

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

(c) For each value of D, plot Zipf (or size) 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) Plot size rankings 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].

Measure α for each value of ρ, and indicate these values in plots.

Perform the regression over all ranks except for the tail ranks for which $S = 1$.

(e) Plot your estimates of α as a function of $\rho = 0.10, 0.20, \ldots, 0.90$.

Is the exponent α the same across densities or does it change with ρ?

(f) Not seeing the forest for the trees:

However, the real density of these growing forests is not simply the number of trees normalized by L^2.

As shown in class, here's an example simulation run by David Matthews, PoCSologist in 2020,[1] that shows how the density is mismeasured.

https://pdodds.w3.uvm.edu/videos/2020-11-12L64D4096-david-matthews.mp4

The forests grow in from the corner least likely to be hit, and a large space of non-treed land remains in the top left corner.

So let's go back a few steps.

Plot the 9 forests you have produced for $\rho = 0.10, 0.20, \ldots, 0.90$.

(g) Now plot the density of the real forested section, $\rho_{\rm f}$ as a function of overall tree density ρ.

Do this by removing the largest contiguous non-treed area.

[1] When PoCS was 100% POX

(h) Using results from the previous two questions:
Plot α as a function of the measured density ρ_{f}.

(i) Not graded, open, and possibly impossible:
Can you deduce a form for how the exponent behaves with either density?
If so, how can we determine this relationship theoretically?

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