Packard's Snowflakes

Janko Gravner, UC Davis
David Griffeath, UW Madison

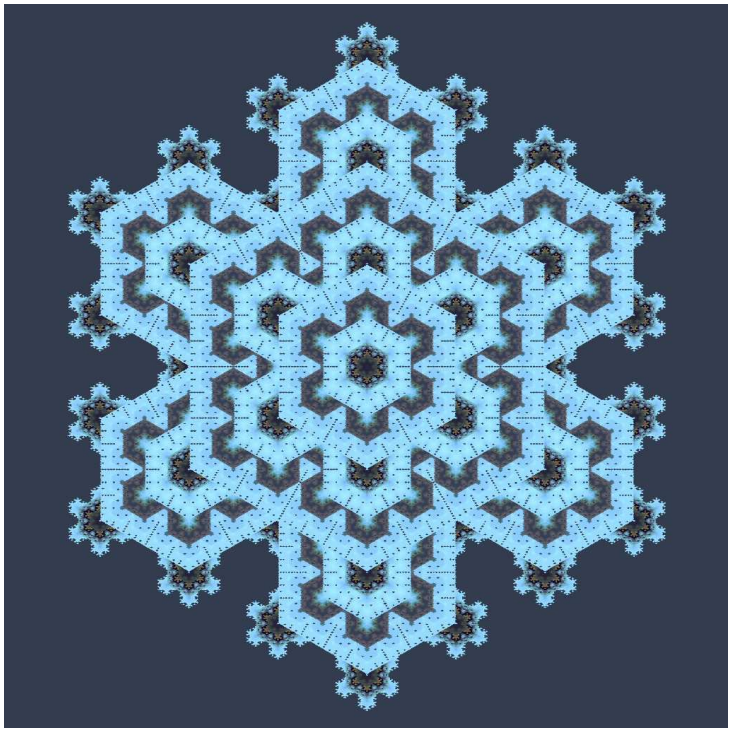
October 2007

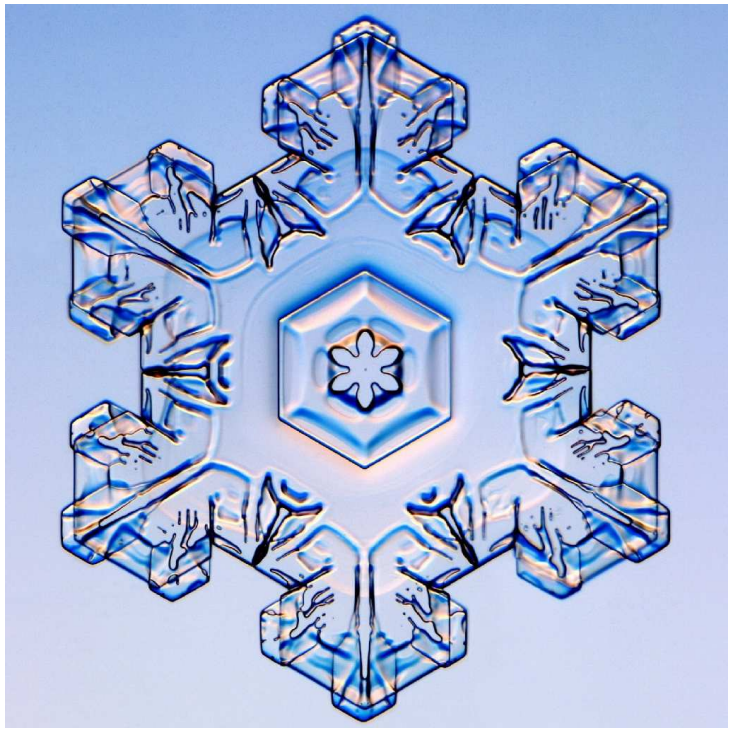
Packard's Snowflakes

In a 1984 paper, N. Packard proposed two models for snowflake growth:

- A variant of diffusion limited aggregation (DLA).
- A class cellular automata on a triangular lattice $\mathbb{T}$ with the property that a site having exactly one occupied (i.e., frozen) neighbor always becomes occupied at the next time, but a site with exactly two occupied neighbors does not. Idea: in real snowflakes growth favors the tips of the crystal.

A more realistic-looking DLA dynamics was developed by Reiter (2005), while the CA approach was popularized by Wolfram in various papers and his 2002 book.





Comparison between a Packard's CA and a real snowflake.

Simulations

A fortuitous correspondence between $\mathbb{T}$ and $\mathbb{Z}^2$, with neighborhood of 0

$$\mathcal{N} = \begin{array}{ccc} & \bullet & \bullet \\ \bullet & 0 & \bullet \\ \bullet & \bullet & \end{array} \quad ,$$

lets us use Cartesian coordinates, and easily simulate the dynamics. Here we use *MCell*.

Examples:

- *Hex 1*. A site joins the crystal iff it has exactly one occupied neighbor.
- *Hex 135*. A site joins the crystal iff the number of occupied neighbors is odd.
- *Hex 134*.
- *Hex 1456*.

Realism?

Packard's snowflakes have been widely publicized to illustrate how very simple algorithms can emulate complex natural phenomena. For example:

> "An elementary schoolchild could look at any of the gorgeous pictures of computer screens in Packard's collection and instantly identify it as a snowflake."
>
> – Steven Levy, *Artificial Life, 1992*

Mathematical framework

Represent $\mathbb{T}$ by $\mathbb{Z}^2$, with the neighborhood $\mathcal{N}$ of $(0,0)$ consisting of itself and the six sites $(\pm 1, 0)$, $(0, \pm 1)$, and $\pm(1,1)$. The neighborhood of x is then $x + \mathcal{N}$.

$A_t \subset \mathbb{Z}^2$ is the set of occupied sites (or 1's) at time $t = 0, 1, 2 \ldots$.

Solidification: $A_t \subset A_{t+1}$.

Whether $x \notin A_t$ belongs to A_{t+1} depends only on the number of sites it sees in A_t, that is, on $|(x + \mathcal{N}) \cap A_t|$. Thus there is a function $\pi : \{1, 2, 3, 4, 5, 6\} \to \{0, 1\}$ such that, for $x \notin A_t$, $\pi(|(x + \mathcal{N}) \cap A_t|) = 1$ iff $x \in A_{t+1}$. We specify π by listing all n for which $\pi(n) = 1$, e.g. *Hex 13*.

Initial and final sets

Our canonical choice is $A_0 = \{0\}$, although we are interested in arbitrary finite initial sets.

For any solidification dynamics and every A_0, the final set A_∞ exists as a sitewise limit of A_t, and $A_\infty = \cup_{t \geq 0} A_t$.

We assume that $\pi(1) = 1$, $\pi(2) = 0$ and analyze the growth of the resulting 16 snowflakes.

Asymptotic density

For $S \subset \mathbb{Z}^2$, let $\mu_\epsilon = \epsilon^2 \cdot$ (counting measure on $\epsilon\, S$). S has *asymptotic density* ρ if $\mu_\epsilon \to \rho \cdot \lambda$ as $\epsilon \to 0$. Here λ is Lebesgue measure on $\mathbb{R}^2$ and the convergence holds in the usual sense:

$$\int f \, d\mu_\epsilon \to \rho \cdot \int f \, d\lambda$$

for any $f \in \mathcal{C}_c(\mathbb{R}^2)$.

Theorem 1. *To each of the 16 digital snowflakes there corresponds a $\rho \in (0, 1]$, the asymptotic density of A_∞, which is independent of the finite seed A_0.*

Unlike stochastic systems, cellular automata rarely lose memory of their initial states.

Density values

Theorem 2. *The densities are exactly computable in 8 cases:*

$$\rho_{13} = \rho_{135} = 5/6 \approx 0.8333,$$
$$\rho_{134} = \rho_{1345} = 21/22 \approx 0.9545,$$
$$\rho_{136} = \rho_{1356} = \rho_{1346} = \rho_{13456} = 1.$$

In six other cases, one can estimate, within ± 0.0008*,*

$$\rho_1 \approx 0.6353,$$
$$\rho_{14}, \rho_{145} \approx 0.9689,$$
$$\rho_{15} \approx 0.8026,$$
$$\rho_{16} \approx 0.7396,$$
$$\rho_{156} \approx 0.9378.$$

Finally,

$$\rho_{146} \in (0.995, 1), \rho_{1456} \in (0.9999994, 1).$$

Note: $\rho_{14} > \rho_{134}$.

Open problem: Are $\rho_{14} = \rho_{145}$ and $\rho_{146} = \rho_{1456}$?

In computer simulations the two dynamics of each pair appear *identical* starting from $A_0 = \{0\}$. This is not persuasive: one would conclude, from observing *Hex 1456* from $A_0 = \{0\}$ on $10^6 \times 10^6$ array, that $A_\infty = \mathbb{Z}^2$!

Macroscopic dynamics

Theorem 3. *Assume that* $t_n = a \cdot 2^n$*, for* $a \in [0,1]$*. Then:*

$$2^{-n} A_{t_n} \to \mathcal{S}_a,$$

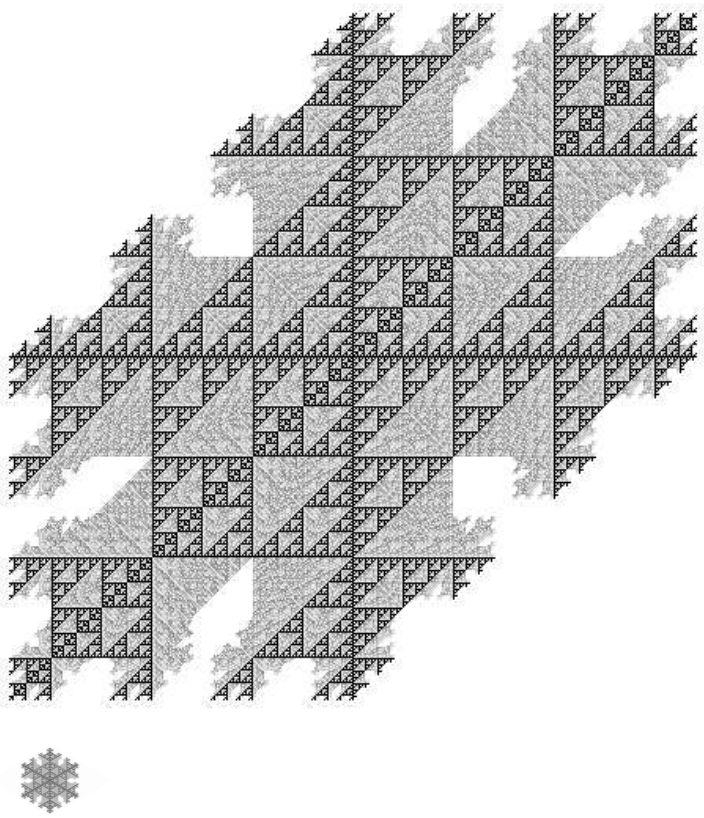
in the Hausdorff metric, as $n \to \infty$*, uniformly in* a*. Here* $\mathcal{S}_a$ *is a strictly increasing family of closed subsets of the hexagon* $\mathrm{co}(\mathcal{N}) \subset \mathbb{R}^2$*.*

The 12 rules that do not have both $\pi(3) = 1$ and $\pi(4) = 1$ have a single $\mathcal{S}_a$. In the remaining 4 cases, $\mathcal{S}_a$ depends on the initial state, but not on the rule!

$\partial \mathcal{S}_a$ is often fractal, e.g. $\dim_H(\partial \mathcal{S}_{14/15}) = 5/4$. Fractal dimension is not known for irrational a.

Why does this all work?

Starting from a single occupied cell, the *light cone CA* forms an impenetrable web of occupied sites that divides further solidification into independent finite domains with simple boundary conditions.

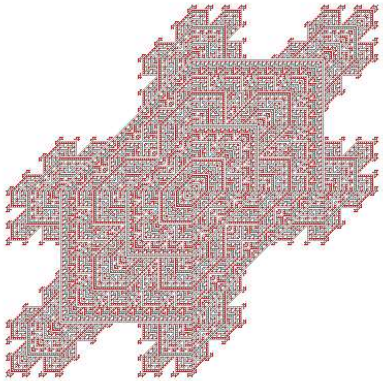


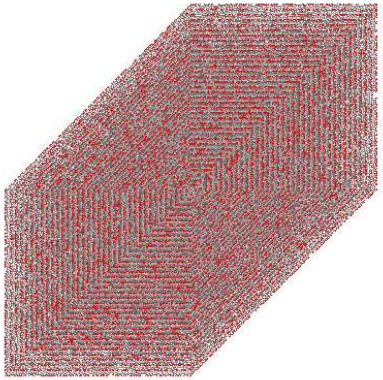
Boundary effects within each domain are controlled. Finally, the light cone CA is additive, so the web from a general finite seed is representable as a superposition of webs from each of its individual cells.

Important steps in the proofs

- In the 8 non-exactly-solvable cases, the Renewal Theorem is necessary for the existence of densities, as well as their approximation.
- For *Hex 1456*, one needs a rescaling argument, very reminiscent of those used in percolation theory, to show that the density is not 1.
- Extensive computer computations and mathematical arguments are both necessary components of such research. Ignore either one at your peril!

Sensitivity to random perturbations





The set A_t now solidifies according to a random rule:

- A site $x \notin A_t$ with exactly one occupied neighbor becomes occupied automatically.
- A site $x \notin A_t$ with at least 2 occupied neighbors becomes occupied with probability $p > 0$.

Theorem 4. *With probability 1,* $(t - R\log t) \cdot \mathrm{co}(\mathcal{N}) \subset A_t$, *eventually. Here,* $R = R(p) > 0$ *is a constant.*

A stronger result should be true!

Proof of Theorem 5

Assume that the light cone CA has, in the first quadrant, above $y = x$, a horizontal interval of exactly L 0's, e.g.,

$$\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 1 \\ \cdot & \cdot & & & & \end{array}$$

Proof of Theorem 5

Assume that the light cone CA has, in the first quadrant, above $y = x$, a horizontal interval of exactly L 0's, e.g.,

$$\begin{array}{cccccccc} & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ . & . & & & & & \end{array}$$

Proof of Theorem 5

Assume that the light cone CA has, in the first quadrant, above $y = x$, a horizontal interval of exactly L 0's, e.g.,

$$\begin{array}{cccccccc} & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & \end{array}$$

Proof of Theorem 5

Assume that the light cone CA has, in the first quadrant, above $y = x$, a horizontal interval of exactly L 0's, e.g.,

$$\begin{array}{cccccc} & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & \end{array}$$

For $a \in \mathbb{Z}^2$ within the above portion of the light cone

$$P(d(a, A_t) \geq k) \leq \sum_{L \geq k} L \sum_{n \geq 0} (1-p)^{L+n} \leq C \cdot e^{-pk/2}.$$

Now use that A_t grows at least as fast as f.p.p. with geometric (p) passage time.

Proof of Theorem 5, continued

Pick $a \in (t - R\log t) \cdot \mathrm{co}(\mathcal{N})$. We need to show that uniformly in such a,

$$P(a \notin A_t) \leq C \cdot t^{-4},$$

provided R is large enough.

To do this, we first bound

$$P\left(d(a, A_{t-R\log t}) \geq \beta R \log t\right),$$

by choosing $R \geq 4/(\beta p)$. (Small $\beta > 0$ to be chosen later.)

Then, we need to deal with

$$P\left(d(a, A_{t-R\log t}) \leq \beta R \log t, a \notin A_t\right),$$

which is bounded by

$$P\left(\sum_{k=1}^{\beta R\log t}\eta_k \geq R\log t\right),$$

where η, η_k are i.i.d. geometric(p). This, by exponential Chebyshev, is at most

$$\left[(Ee^{\alpha\eta})^{\beta}\cdot e^{-\alpha}\right]^{R\log t}.$$

Here $\alpha > 0$ is another parameter, and we only need to show that we can choose α,β so that $\log[\cdot] < 0$. However,

$$Ee^{\alpha\eta} = \sum_{k=1}^{\infty} e^{\alpha k}p\cdot(1-p)^{k-1} = \frac{e^{\alpha}p}{1-e^{\alpha}(1-p)}$$

and so

$$\begin{aligned}\log[\cdot] &= -\alpha + \beta\cdot\alpha - \beta\cdot\log(1-(p^{-1}-1)(e^{\alpha}-1)) \\ &\sim -\alpha + \beta\cdot\alpha + \beta\cdot(p^{-1}-1)\cdot\alpha\end{aligned}$$

if αp^{-1} is small. This equals

$$\alpha(\beta p^{-1}-1),$$

so $\beta = p/2$ works.

What could be improved?

Our result is

$$(t - R\log t) \cdot \mathrm{co}(\mathcal{N}) \subset A_t,$$

if t is above some random time T.

From the proof we have $R = 8/p^2$. Whereas it is clear that the distributio of T should depend on p, this is probably not true for the optimal R! The boundary of A_t appears to approach an invariant measure whose density is bounded away from 0 for $p > 0$.

Possible idea: large intervals of 0 require large intervals of 1's. What makes large intervals of 1's?