

# Bak-Sneppen Evolutionary Model

Bak, PER AND Kim SNEPPEN, 1993, PUNCTUATED EQUILIBRIUM AND CRITICALITY  
IN A SIMPLE MODEL OF EVOLUTION: PHYSICAL REVIEW LETTERS, Vol. 71, No. 24, p 4083-4086  
PROGRAMMED by STEVE BAEDKE

Geology 200 - Evolutionary Systems  
James Madison University  
Lynn S. Fichter and Steven J. Baedke

## INTRODUCTION

Traditionally, evolution is imagined as gradual change taking place over long periods of time. This idea can be traced back to Darwin, . . .

*“Nothing can be effected, unless favourable variations occur, and variation itself is apparently always a very slow process. The process will often be greatly retarded by free intercrossing. Many will exclaim that these several causes are amply sufficient wholly to stop the action of natural selection. I do not believe so. On the other hand, I do believe that natural selection will always act very slowly, often only at long intervals of time, and generally on only a very few of the inhabitants of the same region at the same time.”* Charles Darwin, Origin of Species, 1<sup>st</sup> Edition, p 153

. . . and has been the mantra of the Modern Synthesis, the basis of *all* modern theories of evolution from the 1940's to the present day.

Then in 1972 Eldredge and Gould published their paper “Punctuated Equilibria: An Alternative to Phyletic Gradualism.<sup>1</sup>” The paper produced a rage of controversy generally split along discipline lines. Today the controversy has died down a little; each side has declared victory and gone about its business. Yet, a shaky middle-ground has also emerged, partly because people have gotten past their misconceptions of what they thought the other side was saying, and partly because of new discoveries. Yes, long term changes are punctuational, but when change does take place it takes place according to mechanisms worked out by population biologists. The difference is, biologists, who study evolution over short time spans (years to hundreds of years) using population genetics, see gradual and incremental changes, while paleontologists, who observe evolution over long spans of time (tens of thousands to billions of years), see long periods when no evolutionary change takes place - *equilibrium*, or stasis, *punctuated* by short intervals of very fast change - thus *punctuated equilibrium*.

But, if both gradualism and punctuated equilibrium are based on the same evolutionary theory then what is the difference? Eldredge and Gould based their theory on observations of the fossil record that show that species appear “suddenly” (by geological standards) and fully formed, and then undergo no further change until extinction. More specifically, Eldredge and Gould argue that most evolution occurs through speciation events, and since successful speciation events are episodic, evolution is episodic. There is more to it than this, but that is another story.

---

<sup>1</sup> Eldredge, N., & Gould, S. J. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In: Models In Paleobiology (Ed. by T. J. M. Schopf), p 82-115: San Francisco, Freeman Cooper

But, if the origin and evolution of species remains controversial the size and timing of extinction events is less so. Raup has shown that extinction events follow a power law distribution, and recall that power law relationships are one of the universal properties of complex systems. So, maybe building an evolutionary model that generates power law relationships can give us some insight into the larger dynamics of both evolution and extinction events. This led Bak and Sneppen to build the Bak-Sneppen model specifically to discover if a simple, random selection system produces, as their paper title says, “Punctuated Equilibrium and Criticality in a Simple Model of Evolution”. Definitely a top-down approach to the problem of evolution. Needless to say, the model does show criticality, otherwise we would not be studying it here. So, let’s get on with the experiment.

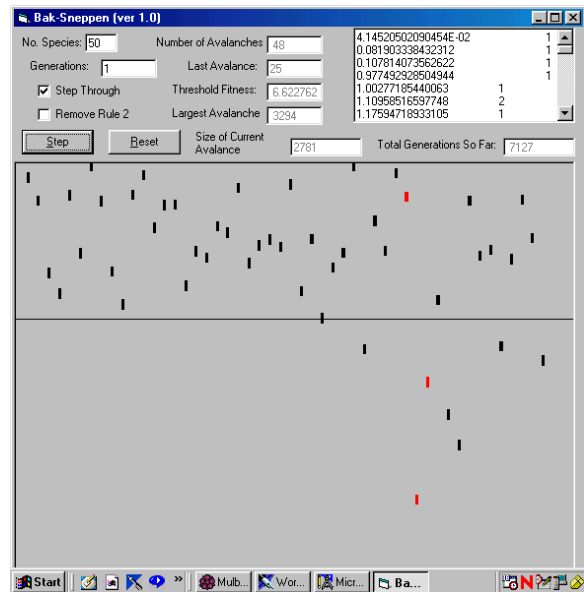
### THE BAK-SNEPPEN MODEL

The Bak-Sneppen (B-S) evolutionary model is an “ecosystem” in which the fitness of each “species” changes because of its relationships with other “species”, following two simple rules.

‘ Click the Bak-Sneppen” icon in the desktop Evolutionary Systems folder and the program to the right opens.

‘ Observe the following:

1. The model contains some number of species, (changeable in the upper left) each represented only by its total fitness, a number ranging from 0.0 (bottom) to 1.0 (top).
2. Visually, each species is represented by a vertical line. The higher the fitness the longer the line, up to maximum fitness (length = 1.0). Only a dot marking the top of the fitness line is visible. The red species are the ones undergoing change each generation.
3. The species fitness lines are laid out side-by-side in a row, but the row ends wrap; that is, an event that moves off the left side appears on the right side, and vice versa. Imagine it as a circle of vertical lines, and the entire circle represents the “ecosystem.”
4. The model acts out two rules: **Rule One** - find the species with the lowest fitness and randomly change its fitness. **Rule Two** - at the same time the lowest fit species is changed, also randomly change the fitness of the species to the immediate left and right.
5. The horizontal line is the threshold fitness line - the highest level the lowest fitness for the entire ecosystem has reached. That is, every generation the model keeps track of the lowest fitness species and if the lowest fitness species happens to fall higher than the last lowest fitness species the threshold fitness line ratchets up to the new value. The fitness of individual species can fall below the threshold fitness, but the threshold fitness only goes up. The interesting questions are: how far up, how fast, why, and does real evolution behave similarly?



## Experiment One - Bak-Sneppen Model

### Exploring the Model

#### ' *Exploration One - Stepping through the Generations*

- < Set number of species to 10.
- < Check the **Step Through** button; this allows you to move the simulation ahead one generation at a time. Notice the "Resume" button toggles from "Step" to "Resume" by clicking "Step Through".
- < Clicking the **Reset** button randomly resets all the fitnesses; i.e. it's a whole new simulation.
- <

Normally the BS model runs very fast, hundreds or thousands of generations a second; too fast to see what is going on. So, begin by stepping through a simulation one generation at a time. Observe:

1. That the species undergoing change each generation are indicated in red.
2. That it is in fact the lowest fitness species that changes. Find the lowest fitness species and observe to see if it in fact it is the one that changes, turns red if it is not red, in the next generation.
3. That it is always the two species on either side of the lowest fitness species that also change.
4. That the fitness changes are random. Of course, without mathematically analyzing the changes it is not possible to be sure that fitness changes at random, but click through many generations. Can you see any systematic patterns to the changing fitness of the red species?
5. Click through the generations until the **fitness threshold line** goes up. Note that the threshold line rises only when there are no species below the line. Can you see in the simulation the event that caused the fitness threshold to go up? Keep stepping through fitness thresholds.
  - < Does the line ever go up when any of red species are below the threshold line?
  - < Hint: for the fitness threshold to move up the lowest fitness red species must fall above the old threshold fitness value.

#### ' *Exploration Two - Threshold Fitness*

- < Set number of species to 10.
  - < Click **Reset** for a new simulation and then click **Step** for the observations below.
1. Click through the first simulation. Observe how fast the threshold fitness line moves up.
    - < Note the box "Total Generations So Far" keeping track of how many steps.
    - < Big box, upper right, tells you the threshold fitness value and the size of each avalanche.
  2. Record in the table below the generation numbers in which the threshold fitness moves up for 10 rises in threshold level.
  3. Do it twice more. Is there any common pattern to how the threshold fitnesses rise?
    - < How deterministic is this system?
    - < How predictable is the behavior of this system?



THRESHOLDS REACHED AT 400 GENERATIONS					
	Threshold Fitness		Threshold Fitness		Threshold Fitness
Run One		Run Six		Run Eleven	
Run Two		Run Seven		Run Twelve	
Run Three		Run Eight		Run Thirteen	
Run Four		Run Nine		Run Fourteen	
Run Five		Run Ten		Run Fifteen	

**Exploration Four - Avalanches in 10,000 Generations**

An *avalanche* is a cascade of fitness changes below the threshold (i.e. all the blinking dots), although this behavior also results in random fitness changes above the line. An avalanche is over when the lowest fitness species rises above the old threshold, and the line rises to the new value. The size of each avalanche, and the threshold fitness at which it ended are shown in the upper right box.

- < Set number of species to 50.
- < Set generations to 10,000.

1. Run the simulation for 10,000 generations and when done record in the table below the final threshold fitness.
  - < Observe how fast the early avalanches occur. As soon as you click **Reset** notice that the upper right fitness threshold box fill faster than you can observe it.
  - < Watch the **Threshold Fitness** box in the last half of the run. How much, how fast does it move up?
2. Repeat for fifteen simulations.
  - < Circle the smallest and largest threshold fitnesses.
3. Examine the final threshold values. How do they compare with threshold values at 400 generations? If the patterns are similar why are they similar? If they are different why are they different?

THRESHOLDS REACHED AT 10,000 GENERATIONS					
	Threshold Fitness		Threshold Fitness		Threshold Fitness
Run One		Run Six		Run Eleven	
Run Two		Run Seven		Run Twelve	
Run Three		Run Eight		Run Thirteen	
Run Four		Run Nine		Run Fourteen	
Run Five		Run Ten		Run Fifteen	

**' Exploration Five - Avalanches in 100,000 Generations**

< Set generations to 100,000.

1. One hundred thousand is an order of magnitude increase, and takes correspondingly long to run, but run the experiment again at these longer times.
2. Compare the threshold fitnesses from 400 generations with those at 10,000 and those at 100,000. Are there any patterns? What are they?

THRESHOLDS REACHED AT 100,000 GENERATIONS					
	Threshold Fitness		Threshold Fitness		Threshold Fitness
Run One		Run Six		Run Eleven	
Run Two		Run Seven		Run Twelve	
Run Three		Run Eight		Run Thirteen	
Run Four		Run Nine		Run Fourteen	
Run Five		Run Ten		Run Fifteen	

**' Exploration Six - Avalanches With Different Population Sizes**

Theoretically there is a lot of debate about the stability of ecosystems. Specifically does the number of species in an ecosystem affect its stability? We cannot get at this question directly with the Bak-Sneppen model, but we can ask if changing population size affects the behavior of the model.

< Set generations at 10,000.

< Select a variety of population sizes and observe if the system changes behavior. Data tables below.

< While the simulations are running observe for the following.

- C Observe how fast the threshold fitness climbs. Are there any patterns to this? Is it a smooth, steady climb, or does it climb quickly and then slow, or climb slowly and end quickly, or is every run uniquely different? This may be hard to observe but do the best you can. We will work this out mathematically later.
- C Does it matter how many species are present? Does the speed of rise vary significantly with the number of species?

1. Does the number of species affect the threshold fitness?
2. Does the size of the population determine how high, or how fast, the threshold fitness rises, or is the rise and the stability independent of population size?

THRESHOLDS REACHED AT VARIOUS POPULATION SIZES FOR 10,000 GENERATIONS								
	Population Size	Threshold Fitness		Population Size	Threshold Fitness		Population Size	Threshold Fitness
Run 1			Run 6			Run 11		
Run 2			Run 7			Run 12		
Run 3			Run 8			Run 13		
Run 4			Run 9			Run 14		
Run 5			Run 10			Run 15		

### ' *Exploration Seven - Avalanche Patterns*

All those blinking fitness dots can look like just so much random business, but is it? Are there any patterns emerging from these random changes in fitness values? Lets find out.

< Set number of species to 100.

< Click **Begin**, then **Resume** will run the simulation until you shut it off.

1. Watch the avalanche behavior over a long period of time looking for patterns in the avalanches through the ecosystem as a whole. How do the avalanches move through the ecosystem? Do the number of species affected remain the same, or change?
2. Describe the patterns in the avalanches. How many different distinctive avalanche behaviors can you observe.
3. If the threshold fitness is the highest level the lowest fitness has reached why do the avalanches take place below the threshold line?

**' Exploration Eight - Up Close and Personal**

It is wonderful to have a God's eye view of the world and life. Everything is so abstract and impersonal. It really means nothing to us. But, there are personal lessons for us.

- < Set number of species to 50.
  - < Set generations to unlimited.
1. Watch the species above the threshold. How stable are they?
    - < How much are they able to change on their own?
  
    - < How much do they contribute to raising the threshold line to the next level?
  
    - < Imagine the "species" in the model are not species but people. Can you think of anyone, or any group, in our society analogous to the "people" above the threshold?
  
  2. Get personal. Pick out one species above the threshold line and identify with it; imagine it is you.
    - < How safe are you in this avalanche prone world?
  
    - < How much control do you have over your destiny? Why or why not?
  
    - < Are there any innocent victims?
  
    - < Is there any way to protect yourself in such a world.
  
  3. Is there any part of this ecosystem that is isolated from the rest, sitting in a protected niche, independent and self sufficient.
    - < What would be required for a part of this ecosystem to exist protected, unconnected, and independent of the rest?
  
  4. What theoretical models have we explored that apply here?

## Experiment Two - Bak-Sneppen Model

### Changing the Rules

Bak and Sneppen specifically set up the model with two rules:

- < **Rule One** - find the species with the lowest fitness and randomly change its fitness.
- < **Rule Two** - at the same time the lowest fit species is changed, also randomly change the fitness of the species to the immediate left and right.

The question is, are these rules critical to the model. What would happen if we dropped Rule One, or used different rules? You can get some insight into what might happen by recalling the behavior of cellular automata when their birth and survival rules were changed. How did changes in those rules affect the behavior of the system?

#### ' **Exploration Ten - Dropping Rule Two**

- < Set number of species to 50.
  - < Set generations to unlimited.
1. **PREDICTION:** Imagine Rule Two is shut off; write a description of how the model's behavior will change.
  
  2. Click button **Remove Rule Two** and run the simulation until it stabilizes. Describe what happens.
  
  3. Explain why the model behaves as it does.
    - < Now, do it again, only this time after it has reached what appears to be maximum threshold uncheck the **Remove Rule Two** box and observe what happens.
      - i Toggle back and forth with **Remove Rule Two** and observe the behavior.
  
  4. Any explanations?

### ' ***Exploration Eleven - Changing to Rule Three***

Bak and Sneppen devised their model following the ecological principle that it is the species closest to the one undergoing change that are also going to change. But, what happens if we change Rule Two to ***Rule Three: after lowest fitness species is changed, change fitness of two other species selected at random.***

- < Set number of species to 50.
- < Set generations to unlimited.

1. **PREDICTION:** Write a description of how you think the model will behave under Rule Three.
  
2. Run the simulation, observe, and describe how the model really does behave with Rule Three.

### ' ***Exploration Twelve - Rule Four***

Bak and Sneppen devised their model following insight from Kaufmann's NK fitness models that show that epistasis influences the ability of fitness to rise. Epistasis refers to the number of species that are coupled such that a change in one causes changes in the other. So,  $N$  = number of genes.  $K$  = epistasis: number of connections among the  $N$  number of genes. This can range from no connections to  $N-1$  connections (since a species cannot connect with itself).

What Kaufmann discovered is that epistasis = 2 is optimum, and that the more genes are coupled the harder it is for fitness to rise. Without exploring NK fitness models here we can still ask, "What happens to the Bak-Sneppen ecosystem if more and more species are coupled?"

***Rule Four: after the lowest fitness species is changed, change the fitness of X number of species selected at random.***

- < Set number of species to 50.
- < Set generations to 2000.

1. Rule Four allows you to change the number of species that are coupled ; that is, the number of species that will change when the lowest fitness species changes. Run a variety of simulations with a variety of epistatic couplings, find a rule, and write it down below.

## Experiment Three - Bak-Sneppen Model

# Self Organized Criticality and Universality

*“Self-organized criticality refers to the tendency of large dynamical systems to organize themselves into a ‘poised’ state far out of equilibrium with propagating avalanches of activity of all sizes.”* Bak has demonstrated that in any SOC system these avalanches follow a power law distribution, a straight line on a log-log graph. Indeed, a straight line log-log plot is prima facie evidence that the system self evolves to criticality. Thus, our last step is to demonstrate the existence of this power law relationship.

When examining the avalanches what we are interested in are the *sizes of the avalanches* - the number of generations from the end of one avalanche to the beginning of the next avalanche, plotted against the *frequency with which each avalanche size appears*. Of course each avalanche represents a rise in the threshold fitness and by now you should know the higher the threshold fitness gets the longer it takes for the next rise to precipitate the next avalanche.

It turns out that in a small run, tens of thousands of generations, the data looks noisy; there is a plot but it is sometimes hard to see the straight line. So, we need hundreds of thousands to millions of generations - a lot of computer time and number crunching.

## HOW IT WORKS

Imagine an un-coupled selection system where only a single agent changes its fitness at random, and it is always the lowest fitness agent that changes (only Rule One is operating). After the lowest fitness agent changes, its new fitness may be either higher or lower than where it started. If higher, another species's fitness becomes lowest and activity shifts to it. But, if the fitness of shifts lower it will precipitate another random shift in fitness. As long as the fitness remains the lowest the avalanche of change is confined to the single species, and it remains active until it is no longer has the lowest fitness, at which time activity shifts to the new lowest-fitness agent, which enters a new avalanche of behavior.

Fitness shifts are, of course, random, and it is equally likely that fitness will shift up or down. However, fitness changes cannot drop below 0.0 or rise above 1.0, so when the threshold fitness is low there is more "room" at the higher fitnesses, and it is easier for the fitness to move up farther than it can move down. Thus, at first, when the threshold fitness is low (say below 0.5) it is easier and faster to raise it. Once above 0.5 it becomes progressively harder and harder for the threshold to climb. Conversely, the higher the threshold the easier it is for an avalanche of random changes to stay below the threshold fitness, and thus the higher the threshold climbs the longer it takes for a random fitness change to reach it - exponentially long, in fact, as plotting the avalanches shows empirically.

Rises in threshold fitness also become exponentially more difficult. Imagine one agent undergoing an avalanche. Its avalanche stops only when its fitness climbs higher than the next lowest agent's, which begins the next avalanche. And this agent's avalanche ends only when its fitness climbs, at random, above the next lowest agent's, and so on. As threshold fitness climbs closer and closer to 1.0 it becomes more and more difficult to randomly climb above the threshold. In fact, and empirically, the optimal fitness is about two thirds (0.66); fitness climbs rapidly to this value and then slows dramatically (exponentially) in climbing higher.

Adding Rule Two to the system does not change in principle how the system operates, but because there are three agents changing fitness each generation, and the fitness of all three must rise above the threshold to end an avalanche, it is more difficult than an uncoupled, single-agent model. Coupled fitness is, of course, closer to the real world, which is one reason Bak and Sneppen designed their model as they did.

( F ( i ( n ( i ( s (