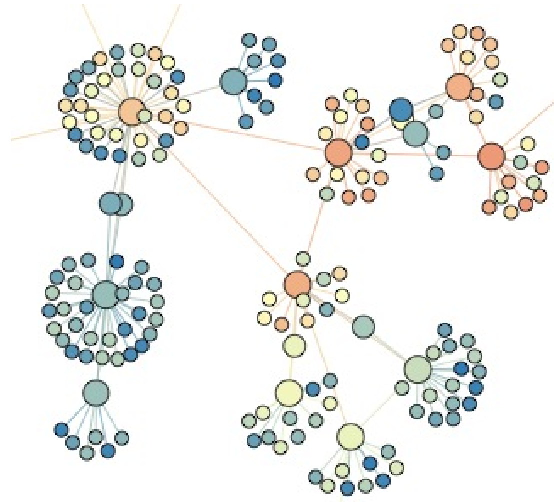


SNA 2A: ER graphs: Insights and realism

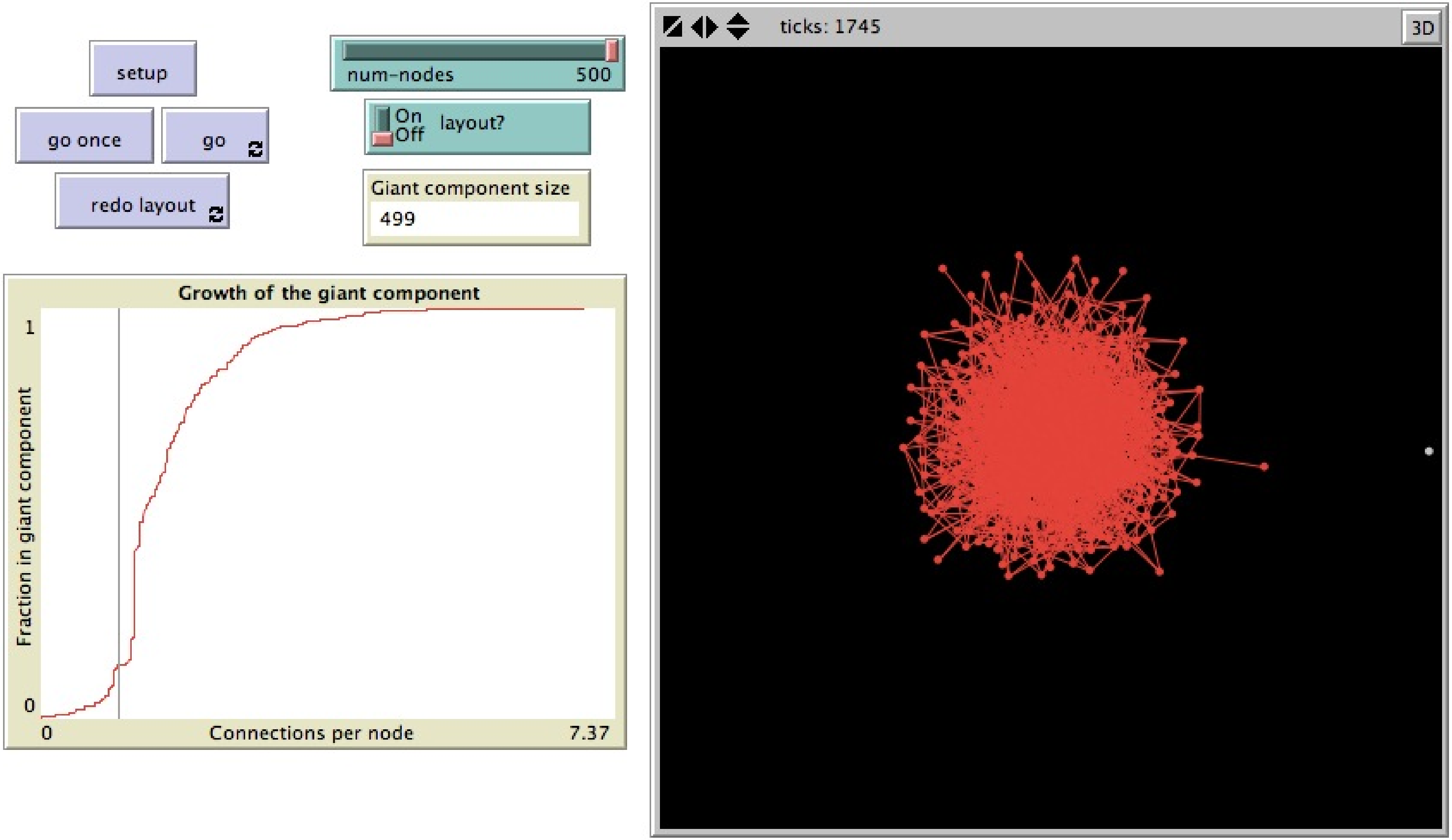
Lada Adamic

school of information

Insights

- Previously: degree distribution / absence of hubs
- Emergence of giant component
- Average shortest path

Emergence of the giant component

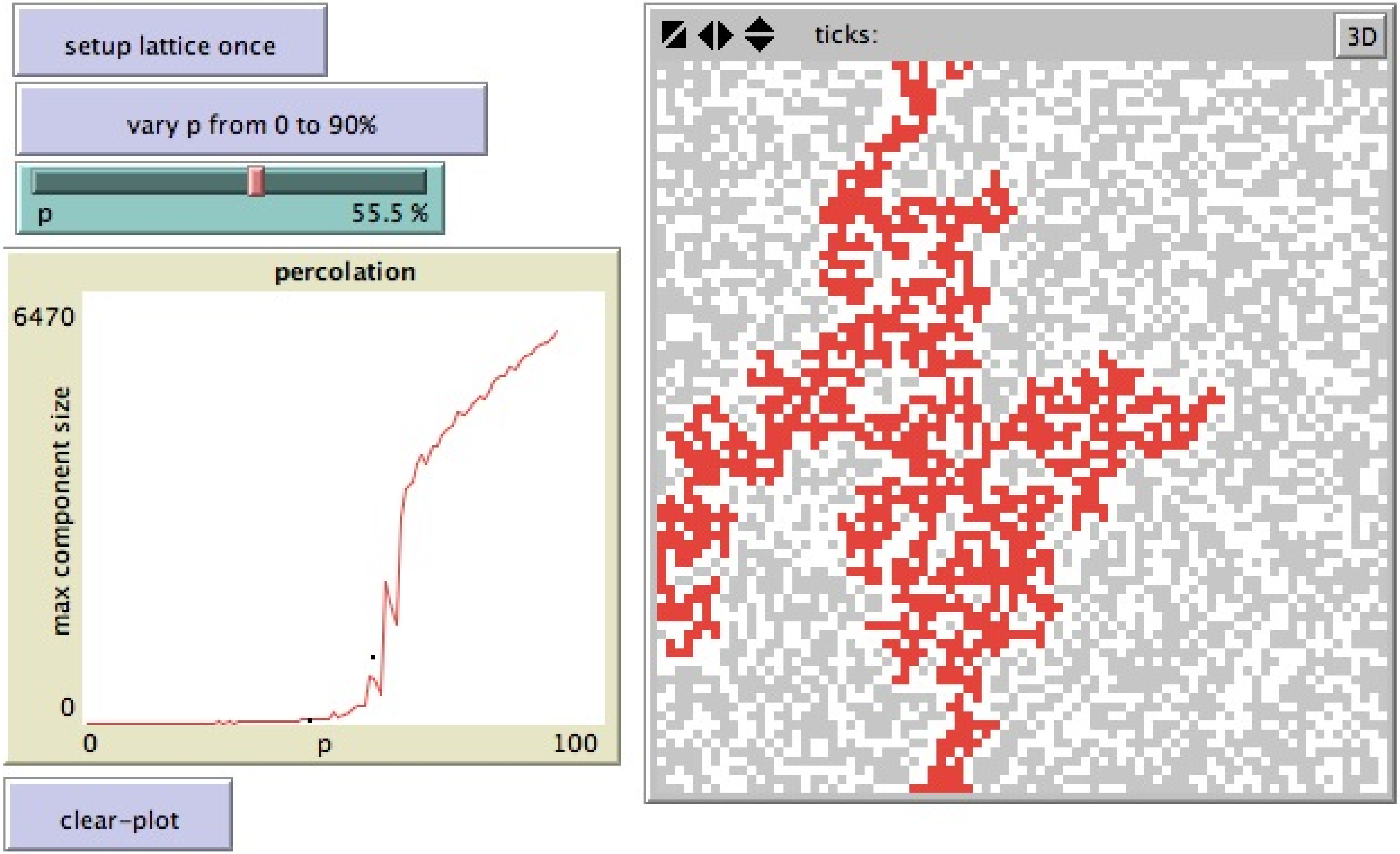


http://ccl.northwestern.edu/netlogo/models/GiantComponent

Quiz Q:

- What is the average degree z at which the giant component starts to emerge?
  - 0
  - 1
  - 3/2
  - 3

Percolation on a 2D lattice

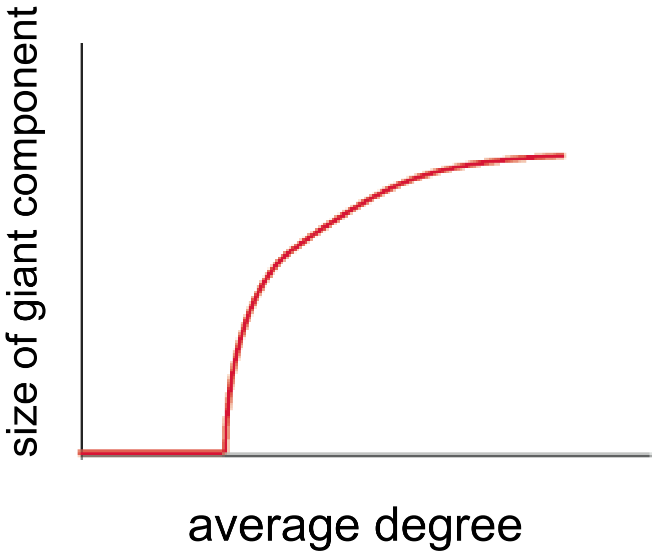


http://www.ladamic.com/netlearn/NetLogo501/LatticePercolation.html

Quiz Q:

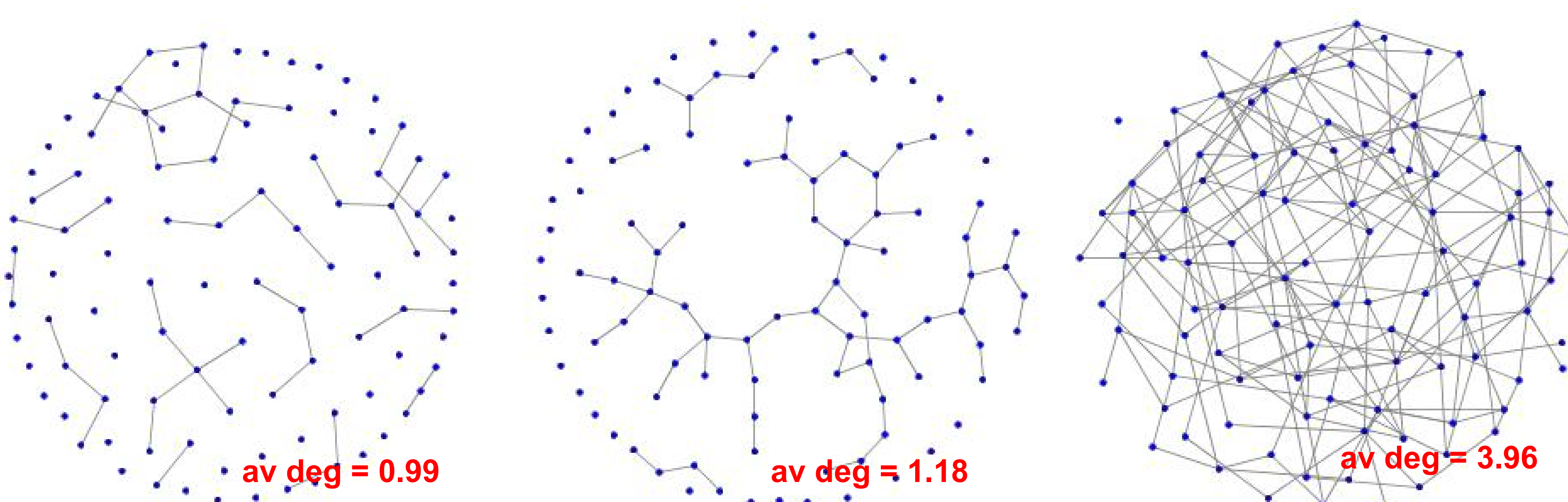
- What is the percolation threshold of a 2D lattice: fraction of sites that need to be occupied in order for a giant connected component to emerge?
  - 0
  - ¼
  - 1/3
  - 1/2

Percolation threshold



Percolation threshold: how many edges need to be added before the giant component appears?

As the average degree increases to z = 1, a giant component suddenly appears



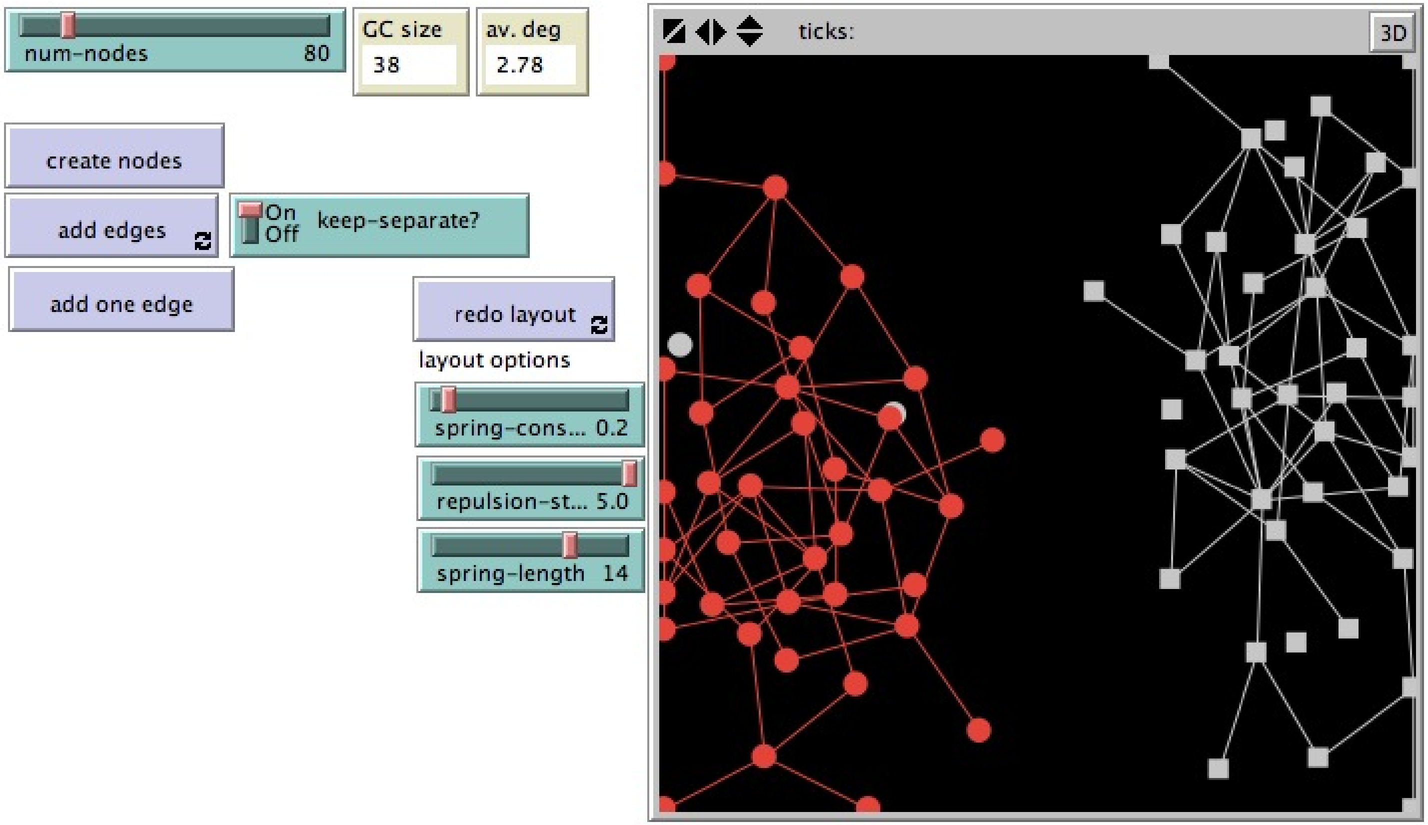
Giant component – another angle

- How many other friends besides you does each of your friends have?
- By property of degree distribution
  - the average degree of your friends, you excluded, is z
  - so at z = 1, each of your friends is expected to have another friend, who in turn have another friend, etc.
  - the giant component emerges

Giant component illustrated

Why just one giant component?

- What if you had 2, how long could they be sustained as the network densifies?



http://www.ladamic.com/netlearn/NetLogo501/ErdosRenyiTwoComponents.html

Quiz Q:

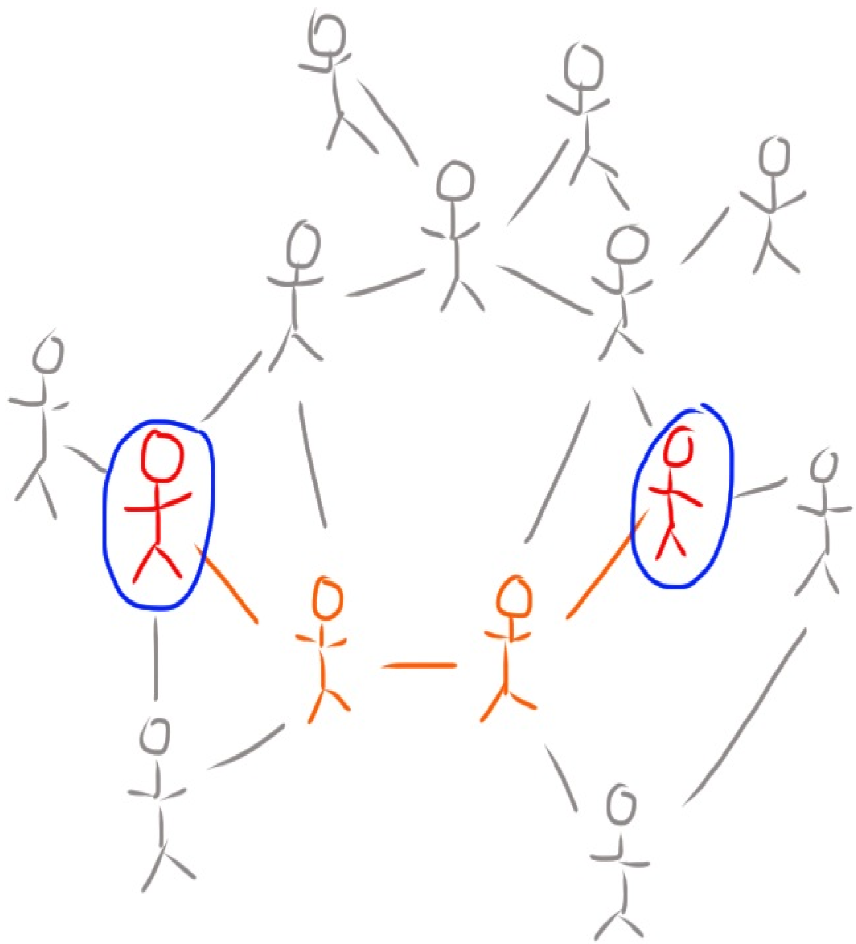
- If you have 2 large-components each occupying roughly 1/2 of the graph, how long does it typically take for the addition of random edges to join them into one giant component
  - 1-4 edge additions
  - 5-20 edge additions
  - over 20 edge additions

Average shortest path

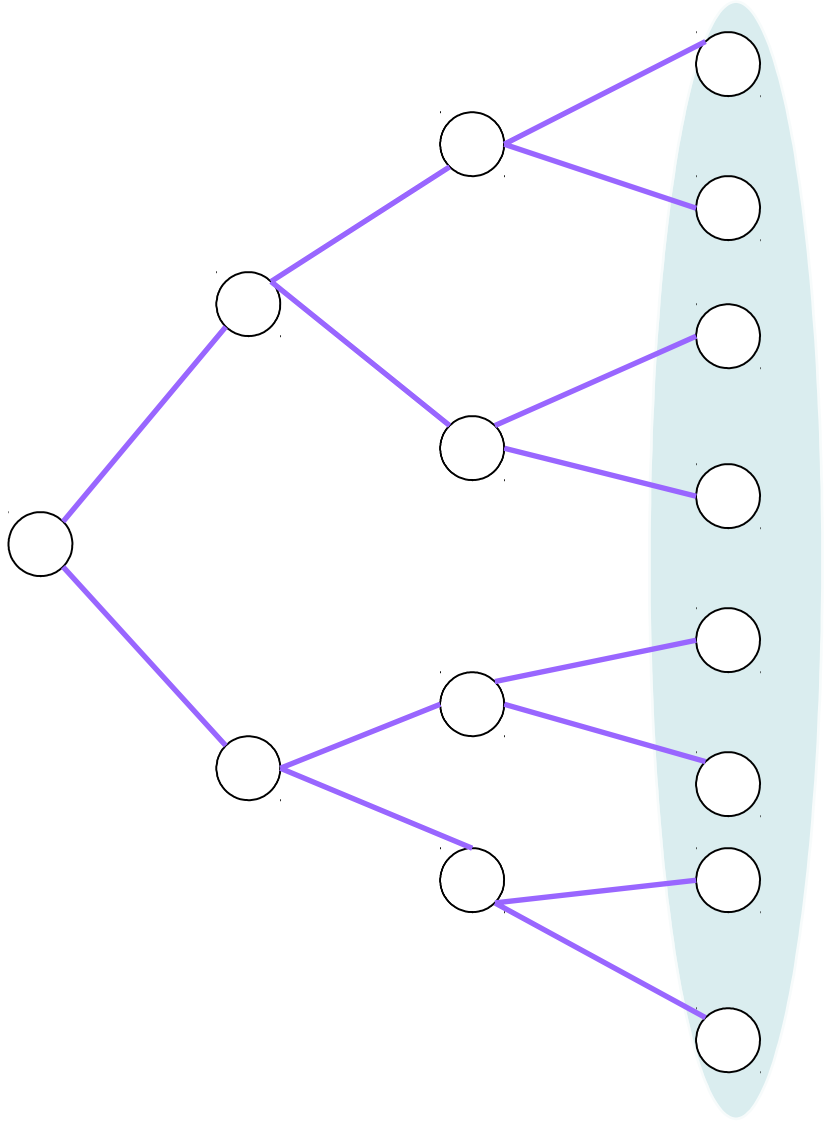
- How many hops on average between each pair of nodes?
- again, each of your friends has ***z = avg. degree*** friends besides you
- ignoring loops, the number of people you have at distance l is

$$z^l$$

Average shortest path



friends at distance l



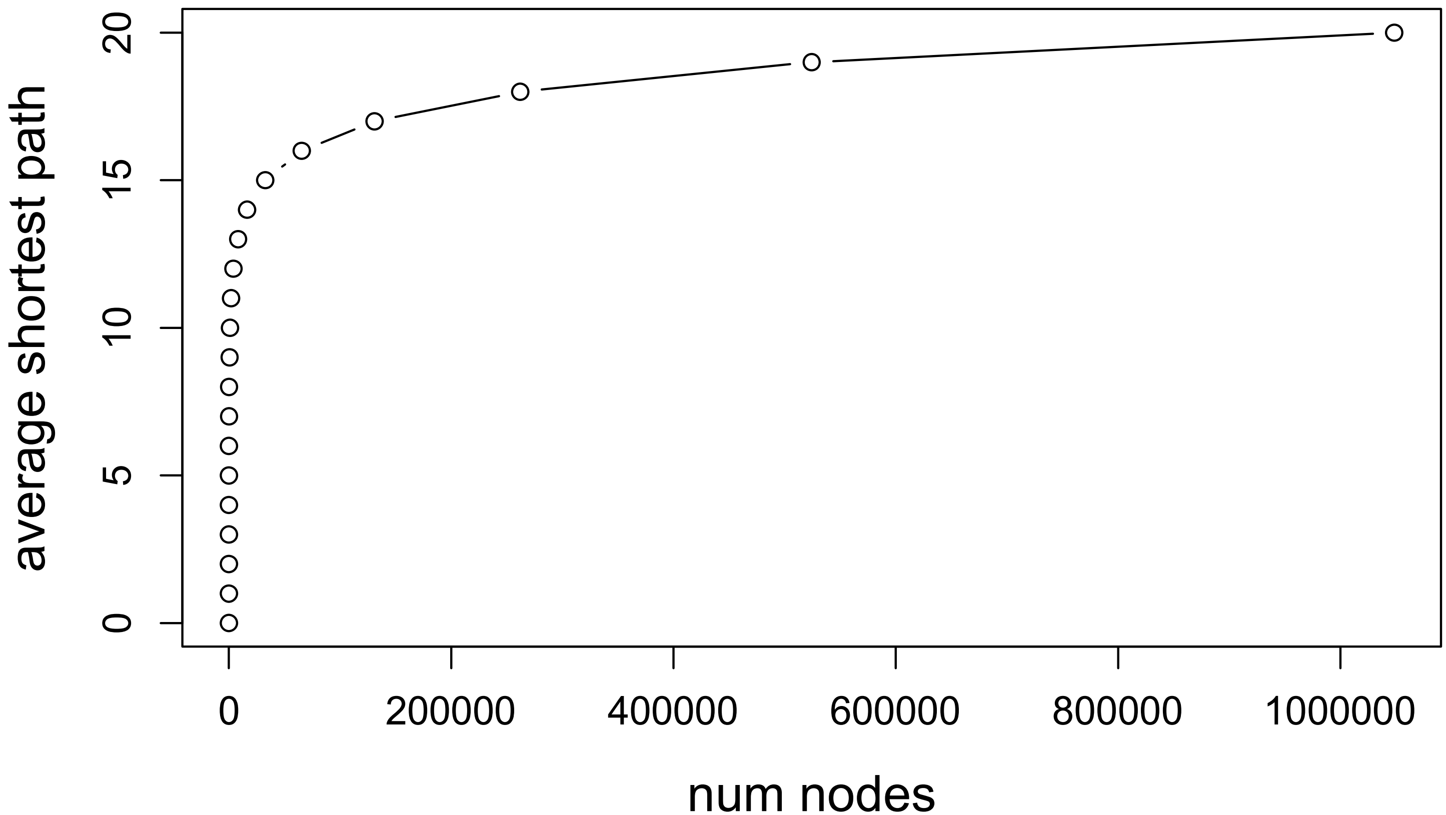
$$N_l = z^l$$

scaling:
average shortest path l_{av}

$$l_{av} \sim \frac{\log N}{\log z}$$

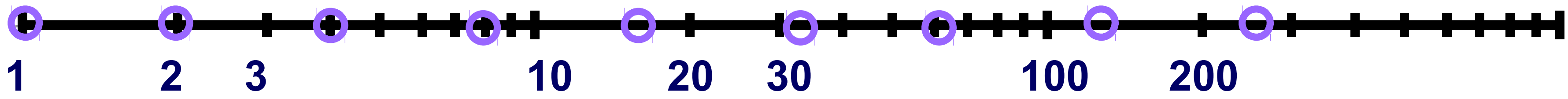
What this means in practice

- Erdös-Renyi networks can grow to be very large but nodes will be just a few hops apart



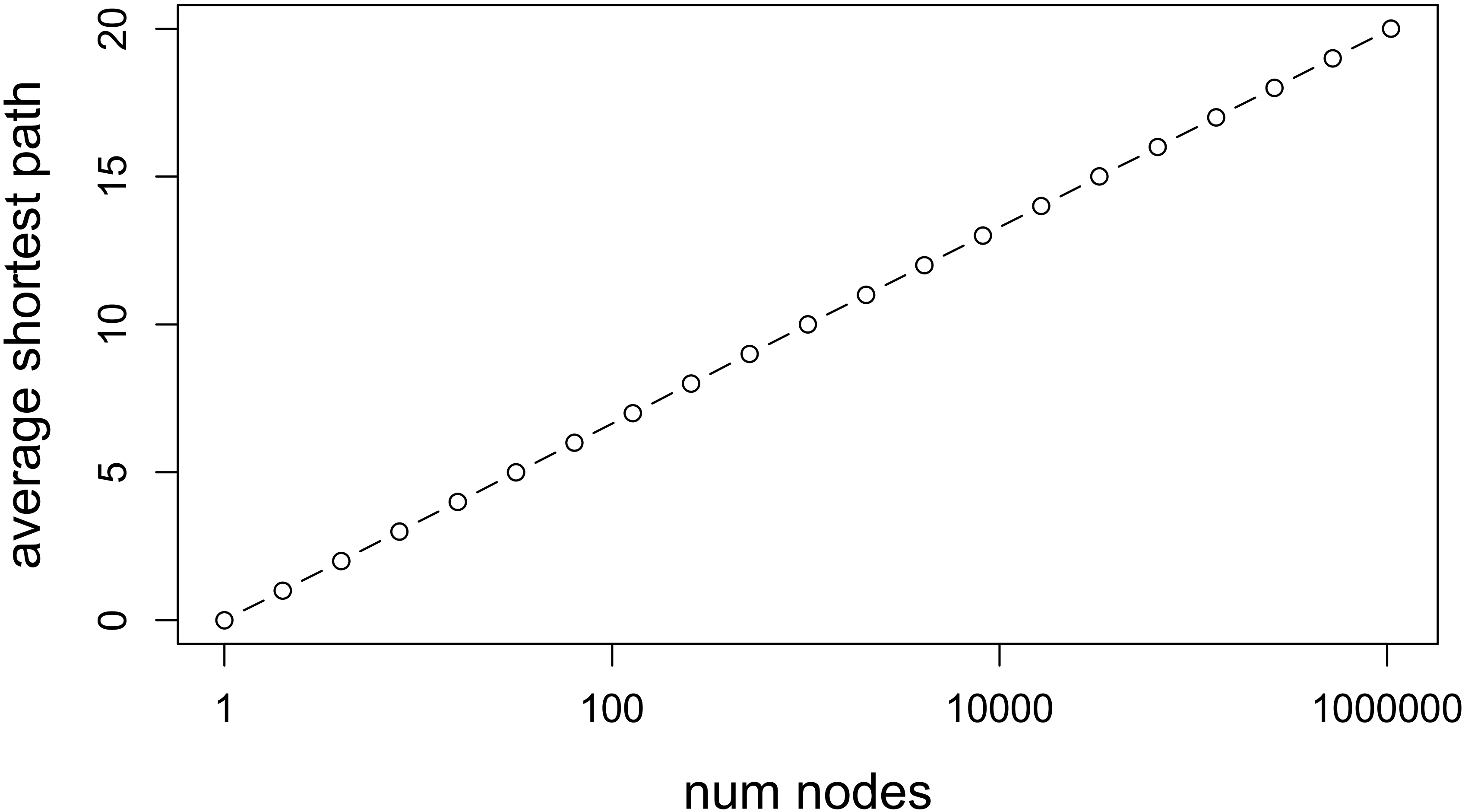
Logarithmic axes

- powers of a number will be uniformly spaced



- $2^0=1$, $2^1=2$, $2^2=4$, $2^3=8$, $2^4=16$, $2^5=32$, $2^6=64$,….

Erdös-Renyi avg. shortest path



Quiz Q:

- If the size of an Erdös-Renyi network increases 100 fold (e.g. from 100 to 10,000 nodes), how will the average shortest path change
  - it will be 100 times as long
  - it will be 10 times as long
  - it will be twice as long
  - it will be the same
  - it will be 1/2 as long

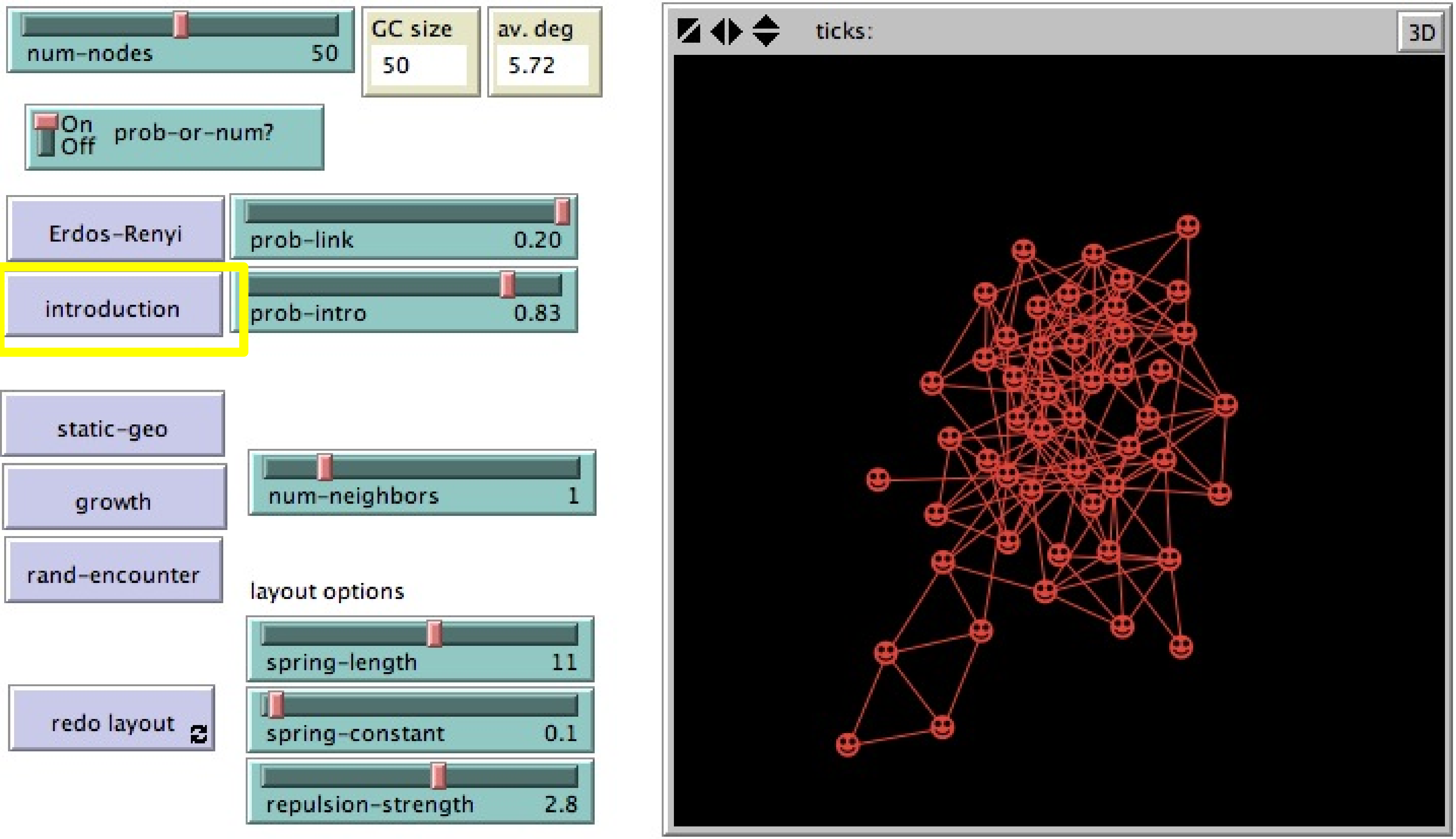
Realism

- Consider alternative mechanisms of constructing a network that are also fairly "random".
- How do they stack up against Erdös-Renyi?
- http://www.ladamic.com/netlearn/nw/RandomGraphs.html

Introduction model

- Prob-link is the p (probability of any two nodes sharing an edge) that we are used to
- But, with probability prob-intro the other node is selected among one of our friends' friends and not completely at random

Introduction model



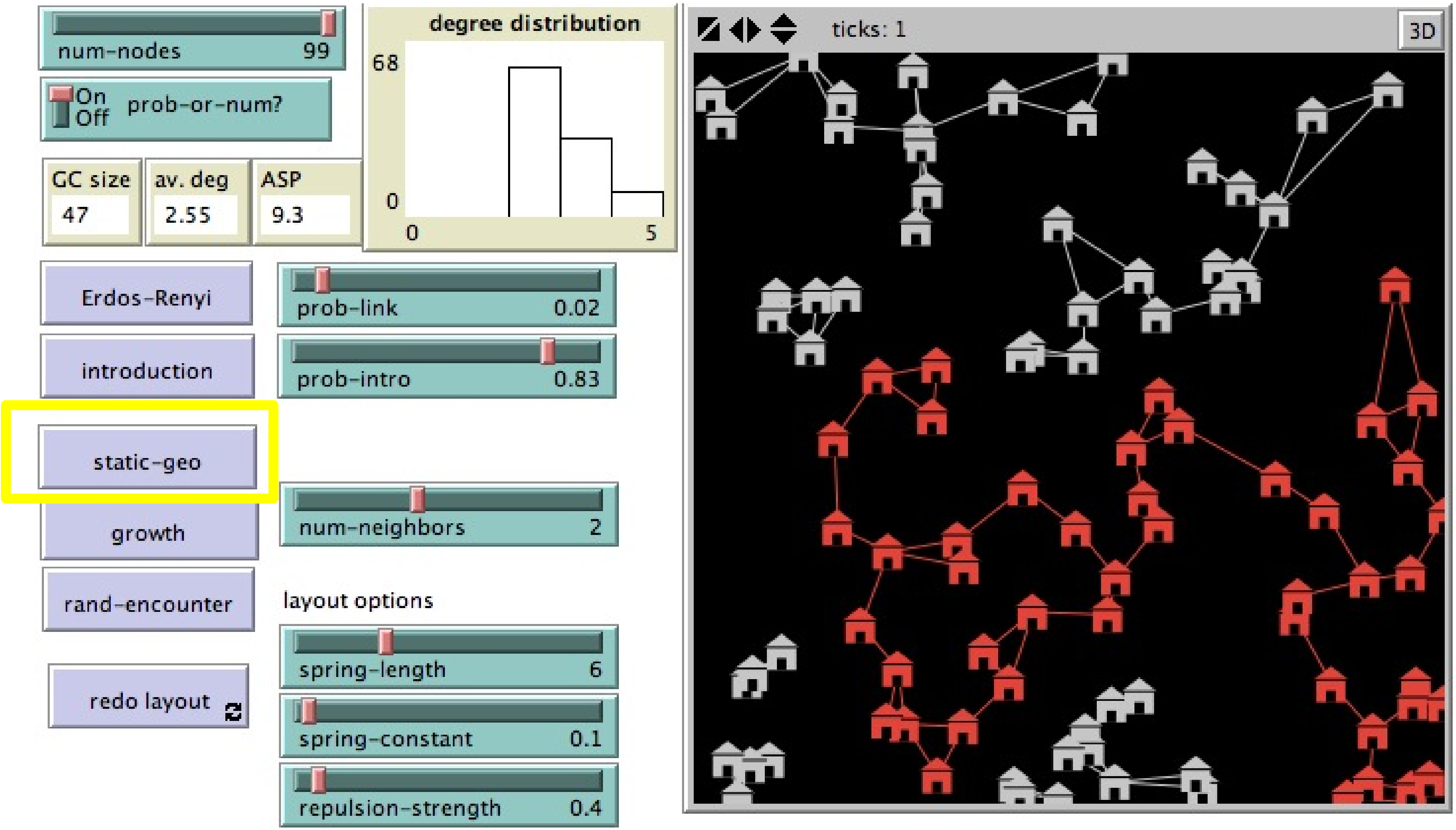
Quiz Q:

- Relative to ER, the introduction model has:
  - more edges
  - more closed triads
  - longer average shortest path
  - more uneven degree
  - smaller giant component at low p

Static Geographical model

- Each node connects to num-neighbors of its closest neighbors
- use the num-neighbors slider, and for comparison, switch PROB-OR-NUM to ‘off’ to have the ER model aim for num-neighbors as well
- turn off the layout algorithm while this is running, you can apply it at the end

static geo


num-nodes 99
On Off prob-or-num?
GC size 47
av. deg 2.55
ASP 9.3
degree distribution
68
0
0
5
ticks: 1
3D
Erdos-Renyi
introduction
static-geo
growth
rand-encounter
redo layout
prob-link 0.02
prob-intro 0.83
num-neighbors 2
layout options
spring-length 6
spring-constant 0.1
repulsion-strength 0.4

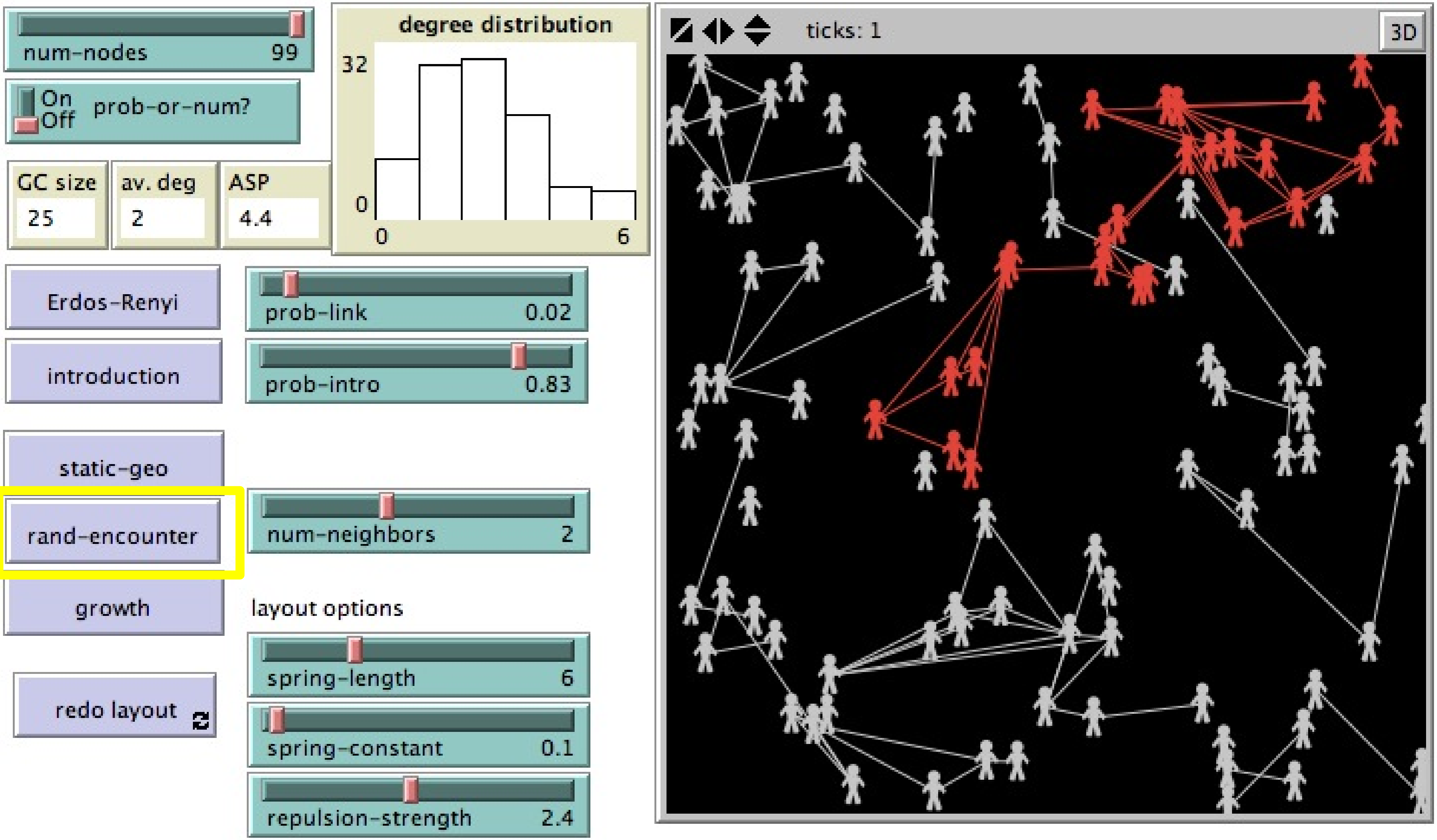
Quiz Q:

- Relative to ER, the static geographical model has :
  - longer average shortest path
  - shorter average shortest path
  - narrower degree distribution
  - broader degree distribution
  - smaller giant component at a low number of neighbors
  - larger giant component at a low number of neighbors

Random encounter

- People move around randomly and connect to people they bump into
- use the num-neighbors slider, and for comparison, switch PROB-OR-NUM to ‘off’ to have the ER model aim for num-neighbors as well
- turn off the layout algorithm while this is running (you can apply it at the end)

random encounters



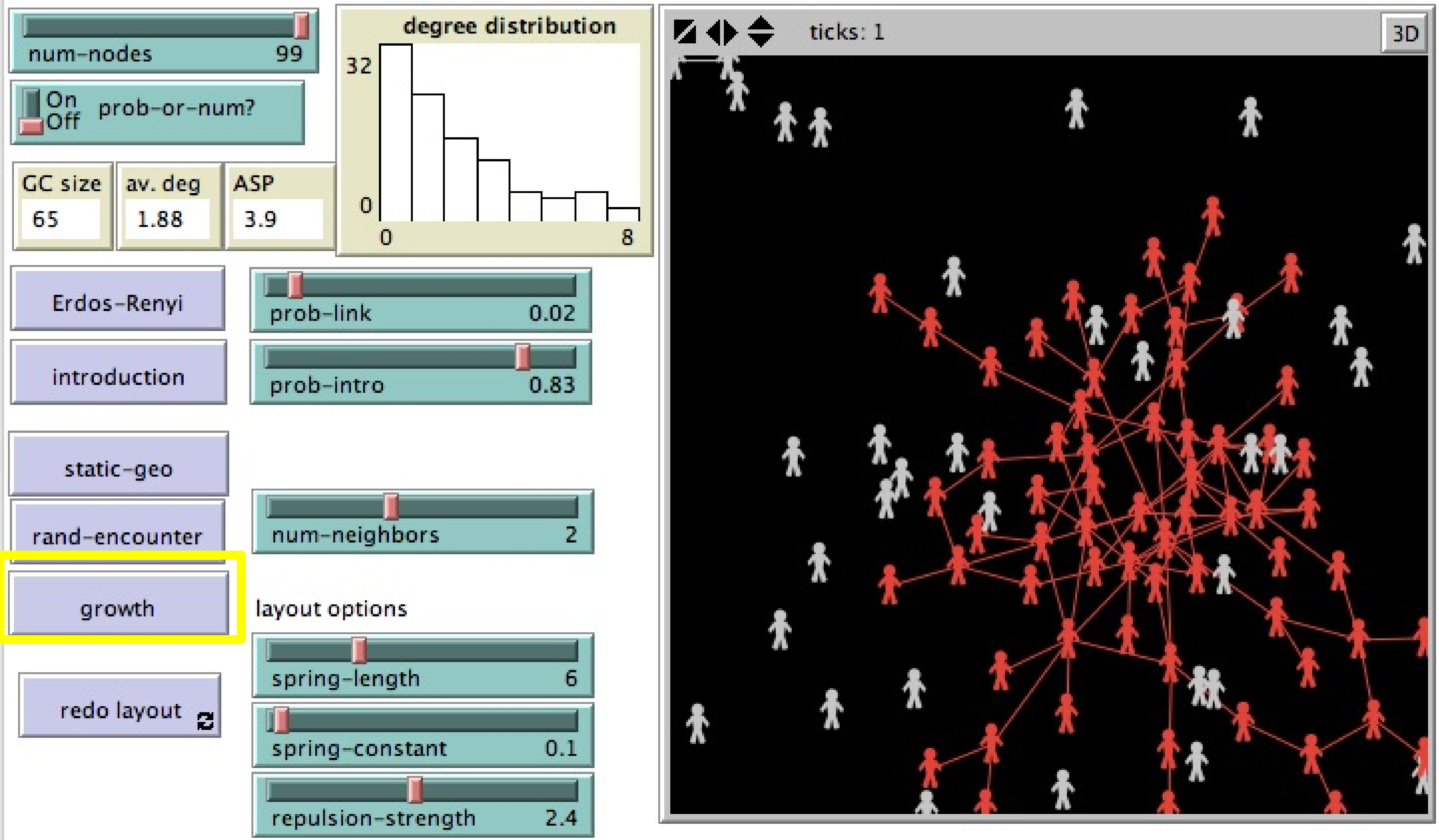
Quiz Q:

- Relative to ER, the random encounters model has :
  - more closed triads
  - fewer closed triads
  - smaller giant component at a low number of neighbors
  - larger giant component at a low number of neighbors

Growth model

- Instead of starting out with a fixed number of nodes, nodes are added over time
- use the num-neighbors slider, and for comparison, switch PROB-OR-NUM to ‘off’ to have the ER model aim for num-neighbors as well

growth model



Quiz Q:

- Relative to ER, the growth model has :
  - more hubs
  - fewer hubs
  - smaller giant component at a low number of neighbors
  - larger giant component at a low number of neighbors

other models

- in some instances the ER model is plausible
- if dynamics are different, ER model may be a poor fit