3 things to do today

1. Revision of the apple policy (2’)
   – The social amendment

2. Understand “small world” phenomenon from an algorithmic viewpoint (50’)

3. Get introduced to power-law and reinforcement dynamics (45’)

CS@CU
Let’s make it more social.

First good news: I give you two apples!

Second good news: you receive from your friends
  – You may keep the first apple
  – You give the second one to a friend/neighbor
Outline

* Milgram’s “small world” experiment

* It’s a “combinatorial small world”

* It’s a “complex small world”

* It’s an “algorithmic small world”
Main idea: social networks follows a structure with a random perturbation

Formal construction:
1. Connect all nodes at distance in a regular lattice
2. Rewire each edge uniformly with probability $p$
   (variant: connect each node to $q$ neighbors, chosen uniformly)

Collective dynamics of ‘small-world’ networks.
Main idea: social networks follows a structure with a random perturbation

Collective dynamics of ‘small-world’ networks.
Outline

* Milgram’s “small world” experiment

* It’s a “combinatorial small world”
* It’s a “complex small world”
* It’s an “algorithmic small world”
Where are we so far?

Analogy with a cosmological principle

– Are you ready to accept a cosmological theory that does not predict life?

In other words, let’s perform a simple sanity check.
1. Consider a randomly augmented lattice
2. Perform “small world” Milgram experiment

Can you tell what will happen?
(a) The folder arrives in 6 hops
(b) The folder arrives in $O(\ln(N))$ hops
(c) The folder never arrives
(d) I need more information
(a) The folder arrives in 6 hops

NOT TRUE

* It actually does look like a naïve answer
* More precisely:
  – By previous result we know that shortest paths is of the order of ln(N), which contradicts this statement.
(b) The folder arrives in $O(\ln(N))$

**ACCORDING TO OUR PRINCIPLE, OUGHT TO BE TRUE BECAUSE IT WAS OBSERVED BY MILGRAM**

* A sufficient condition is:
  - Milgram’s procedure extract shortest path
* Answering this critical question boils down to an **algorithmic** problem
(c) The folder never arrives

SEEMS UNLIKELY

unless the procedure is badly designed (cycle)
or if the grid contains hole
(d) I need more information

* In particular, how to model Milgram’s procedure

  * “If you do not know a target, forward the folder to your friend or acquaintance that is most likely to know her.”
(d) I need more information

* In particular, how to model Milgram’s procedure
  * “If you do not know a target, forward the folder to your friend or acquaintance that is most likely to know her.”

* A mathematical formulation: a greedy routing
  - using grid information and shortcuts “incidentally”
  - N.B.: Grid dimensions can describe geography or other sociological property (occupation, language)
How does greedy routing perform?

* Does it extract the shortest path?
  – Not necessarily

* Case study: dimension k=1, target t
  – We introduce interval: \( I_l = \{ u \in V \mid |u - t| \leq l \cdot \sqrt{N} \} \)
  – The greedy routing constructs a path \( s = U_0, U_1, U_2 \)
    we denote the end-point of the ith shortcuts as \( X_i \)

\[
\mathbb{P} \left[ \bigcup_{i=1,\ldots,n} \{ X_i \in I_l \} \right] \leq \sum_{i=1,\ldots,n} \mathbb{P} [X_i \in I_l] \leq \frac{nl}{\sqrt{N}}
\]
How does greedy routing perform?

\[
P \left[ \bigcup_{i=1,\ldots,n} \{ X_i \in I_l \} \right] \leq \sum_{i=1,\ldots,n} P [X_i \in I_l] \leq \frac{nl}{\sqrt{N}}
\]

* Fixing \( n = l = \frac{1}{2} \sqrt{N} \), this event has proba \( \leq 1/2 \)
  
  – So with proba \( \geq 1/2 \), \( \{ X_1, X_2, \ldots, X_n \} \) are not in \( I_l \)

* On this event, assuming s not in \( I_l \)
  
  – Greedy routing needs more than \( n \) steps
  
  – Or it has to reach \( t \) from boundary of \( I_l \), using \( l \) steps
A thought experiment

* In a line Milgram’s uses $\frac{1}{2}\sqrt{N}$ steps
  - square root is not satisfying for small world
  - Not much better when $k>1$! $\frac{1}{4}N^{\frac{1}{k+1}}$
  - even worse, the proof applies to any distributed alg.

* Our sanity check test has grandly failed!
  - “Small world” results explain that short paths exist
    … finding them remains a daunting algorithmic task
Milgram’s “small world” experiment

* It’s a “combinatorial small world”
* It’s a “complex small world”
* It’s an “algorithmic small world”
  – Beyond uniform random augmentation
**Autopsy of “Small-world” failure**

* In a uniformly augmented lattice shortcuts do exist
  – About $\sqrt{N}$ shortcuts leads to $I_l$ when $l = \sqrt{N}$.

* But they are dispersed among $N$ nodes
* Moreover, previous shortcuts do not progress

* Is there another augmentation?
* What if the augmentation exhibits a bias
  - You may know someone outside your usual circles of acquaintances, but it remains likely that this person is closer to you than an arbitrary person

* Does this break the lower bound proof?
  (a) Yes, finding a neighborhood of $t$ becomes easier
  (b) Yes, but for another reason
  (c) I need more information
How to model augmentation bias

* Formal construction:
  1. Connect all nodes at distance p in a regular lattice
  2. Connect each node to q other nodes, chosen with a biased probability

The small-world phenomenon: An algorithmic perspective.
How to model augmentation bias

* Formal construction:
  1. Connect all nodes at distance $p$ in a regular lattice
  2. Connect each node to $q$ other nodes, chosen with a biased probability

\[
\mathbb{P} [u \rightsquigarrow v] = \frac{1}{\sum_{v \neq u} \frac{1}{\|u-v\|^r}} \frac{1}{\|u-v\|^r}
\]

* $r$ may be called the clustering coefficient
  * If a node is twice further, probability is $2^r$ times less

The small-world phenomenon: An algorithmic perspective.
Impact of clustering coefficient

Small values of r
Approaches uniform augmentation

Large values of r
Approaches original lattice
(a) Yes, finding a neighborhood of $t$ becomes easier

**A PRIORI NOT TRUE**

– It is easier *only if* you are already near the target
– In general, it can take a larger number of steps

* The above argument is missing a key aspect
  – All positions are not equal, hence *progress* is possible
  – As shortcut are used recursively, probability increases
Augmented lattice

Navigable small world
dist. alg need $O(\log^2(N))$ steps

<table>
<thead>
<tr>
<th>Combinatorial Small world</th>
<th>Not a small world</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Short paths exist)</td>
<td>(Short paths do not exist)</td>
</tr>
<tr>
<td>dist. alg. need $N^{(k-r)/(k+1)}$ steps</td>
<td>alg. need $N^{(r-k)/(r-(k-1))}$ steps</td>
</tr>
</tbody>
</table>

The small-world phenomenon: An algorithmic perspective.
The critical case

* Assume $r=k$ (dimension of the grid)
  - A neighborhood of $t$ of radius $d/2$
  - Contains $(d/2)^k$ nodes
  - Each may be chosen with probability roughly $1/(3d/2)^k$
  - Growth of ball compensates probability decreases!

* Harmonic distribution.

Is the analysis of greedy routing tight?
- Yes, greedy routing performs in $\Omega(\log^2 n)$

Can we find path as short as $\log(n)$ (shortest path)?
- Yes, with extra information on neighboring nodes
- Or another augmentation

Can we build augmentation for an infinite lattice?
- See homework exercise (check tomorrow night)
Can we augment other graphs?

- $G=(V,E)$ (i.e. a lattice) with distance known
- Random augmentation adds one shortcut per node

Is routing on $G$ + shortcuts used incidentally efficient?

Indeed all these graphs are polylog augmentable:

- Bounded ball growth, Doubling dimensions
- Bounded “width” (Trees, bounded treewidth graphs)

What about all graphs? Lower Bound $O(n^{1/\ln(n)})$
Practical follow up

Can we observe harmonic distribution?
- Yes, using closeness rank instead of distance

Can we prove it emerge?
- Recent results
- Through rewiring, mobility

Geographic routing in social networks.
D. Liben-Nowell et. al. PNAS (2005)
Milgram’s experiment prove that social networks are navigable
- individuals can take advantage of short paths
- with basic information

This is at odds with uniform random graphs

The key ingredients to explain navigability
- A space easy to route (e.g. grid, trees, etc.).
- A subtle harmonic augmentation (e.g. ball radius).