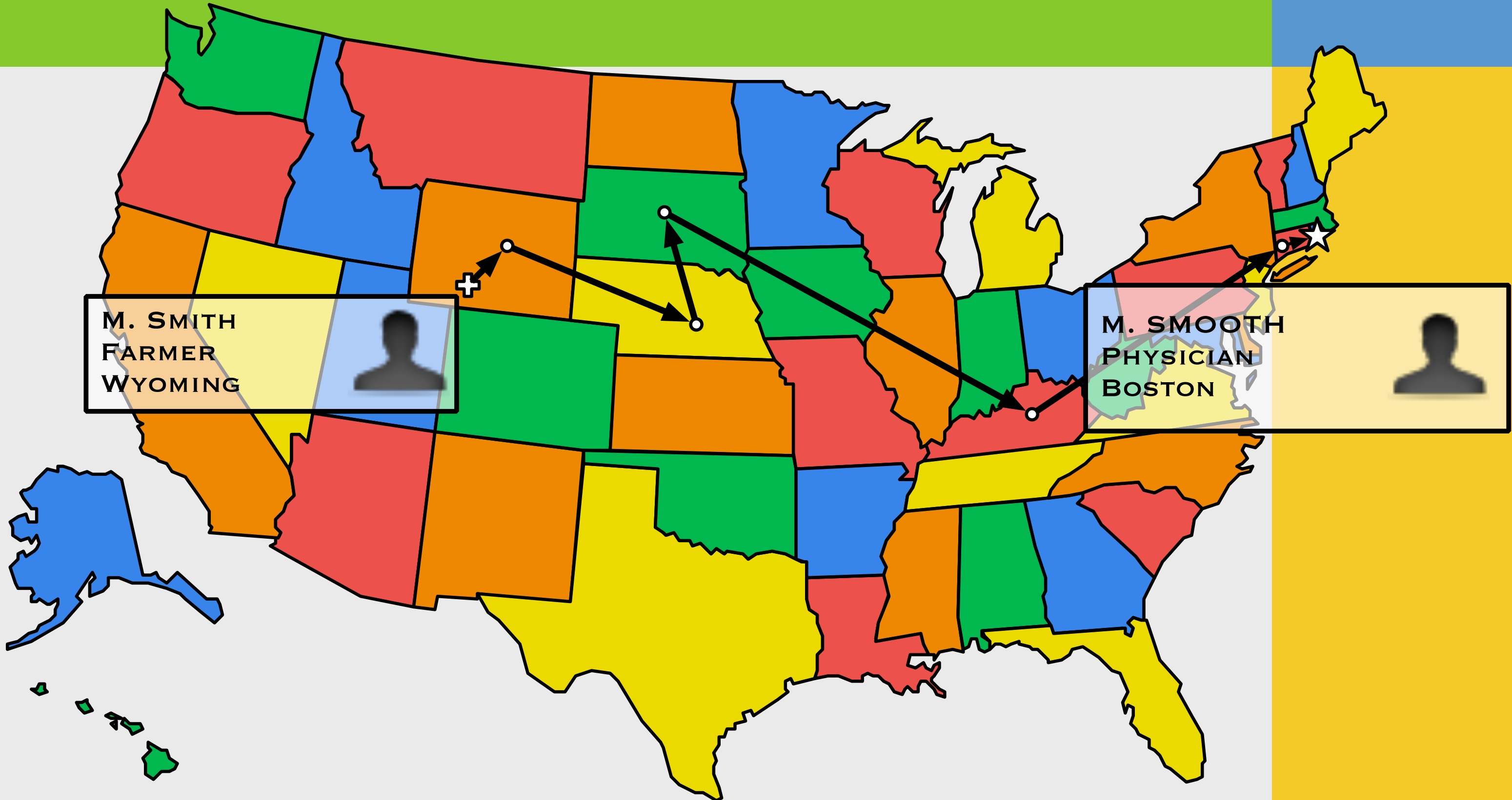
Navigable graph metrics

P. Duchon, N. Hanusse (Bordeaux, France)
E. Lebhar, N. Schabanel (Lyon, France)

Milgram (1967)



Every one is at hop distance ≤ 7 from every other... and is able to <u>find</u>, based on its <u>local</u> view, a <u>short</u> path to a <u>random unknown</u> person

Interaction networks

... share a lot of properties:

- Large (billions of nodes, a linear number of edges)
  $\Rightarrow$ every one has only a local view of the network
- Small diameter
- Power law distributions (on degrees...)
- Local clustering (high clustering coefficient...)

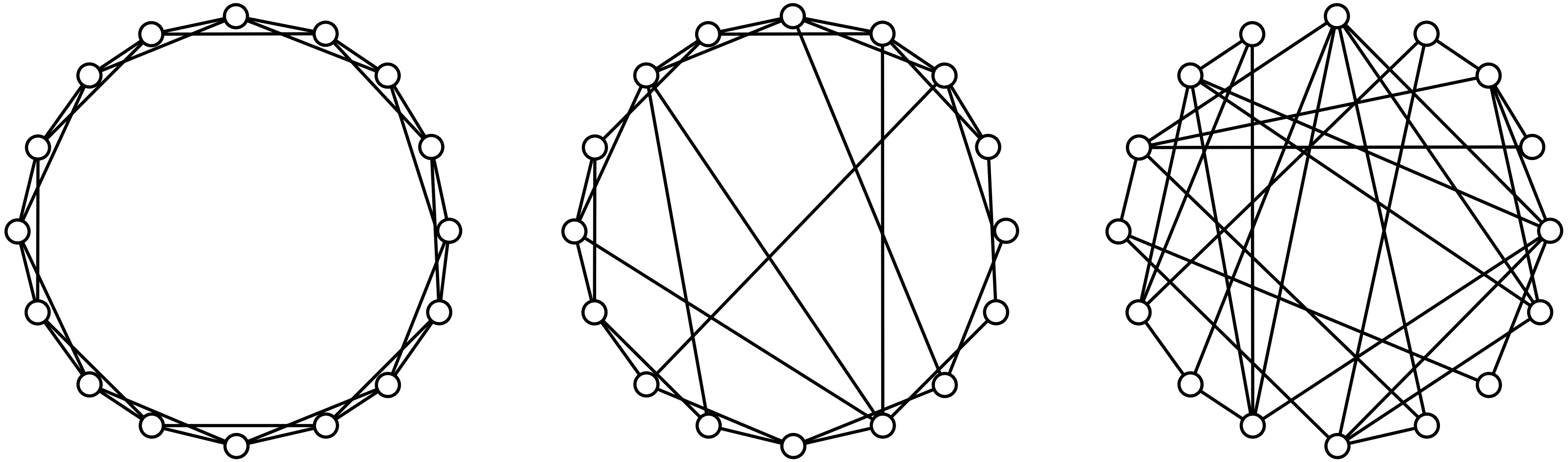
Smallworld phenomenon

- Navigability: Between any pair of nodes, there exists a very short path (of polylogarithmic length) that can be found efficiently, in spite of a very partial knowledge of the global structure of this « spontaneous » networks. (no global design)
- What kind of underlying global knowledge on the structure allow this phenomenon to arise?

Smallworld models

- Watts & Strogatz (1998)
- Kleinberg (2000)

Watts & Strogatz (1998)



No randomness	**Some randomness**	Lot of randomness
Large diameter	**Small diameter**	Small diameter
High clustering	**High clustering**	Low clustering

Watts & Strogatz (1998)

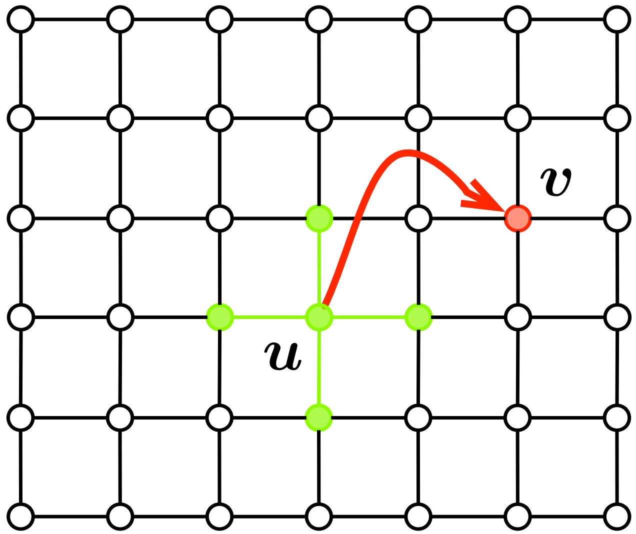
- Kleinberg (2000): this model fails to capture navigability

In spite of a polylogarithmic diameter, any decentralized routing on these augmented rings yields polynomial length paths

No randomness	**Some randomness**	Lot of randomness
Large diameter	**Small diameter**	Small diameter
High clustering	**High clustering**	Low clustering

Kleinberg (2000)

Toric d-dimensionnal n x ... x n grid augmented by random links



$$\Pr\{u \rightarrow v\} = \frac{\mathrm{Cte}}{\mathrm{dist}(u,v)^s}$$

- Grid = Global geographic knowledge
- Long range links/contacts = Extra local knowledge at each node
  = random contact met someday

Kleinberg (2000)

<u>Individual's model:</u> decentralized algorithms

- routes only to known contacts (local or long range), i.e. contacts of previously visited nodes,
- using only the geographic position, i.e. the underlying metric (the grid).

Kleinberg (2000)

- Greedy routing (send the message to the closest neighbor) routes in $O(\log^2 n)$ steps on expectation if $s = d$.
- If $s \neq d$, then any decentralized algorithm takes $\geq \text{poly}(n)$ steps.
- Note that:
  - if $s \leq d$: logarithmic diameter
  - if $d < s < 2d$: polylogarithmic diameter
  - if $s > 2d$: polynomial diameter

Smallworldizable metrics

Given a graph G, augmented into a random graph G^+ by adding one random long range link per node:

- We say that a routing algorithm is decentralized in G^+ if it finds a path between any pair of nodes in the augmented graph G^+, using only the metric of G, i.e., all long range links rooted on non previously visited nodes are unknown.

Smallworldizable metrics

Given a graph G, augmented into a random graph G^+ by adding one random long range link per node:

- An infinite graph G (or a graph family (G_n)) is smallworldizable if there exists a distribution assigning one long range link per node, for which there exists a decentralized routing algorithm that finds short paths, i.e., of expected length uniformly bounded:

  $\leq$ polylog(| ball in G of radius $dist_G$(source,target) |).

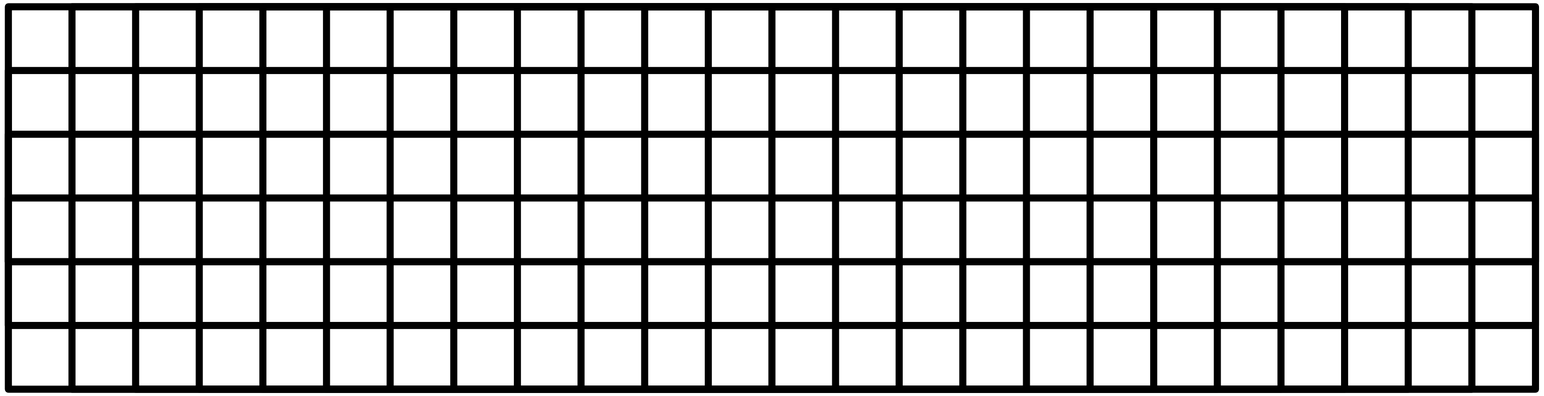
Smallworldizable metrics?

- Any balanced toric grid n x ... x n of dimension d are smallworldizable by adding one single long range link per node whose length r is distributed as $1/r^d$.

What about elongated grids?

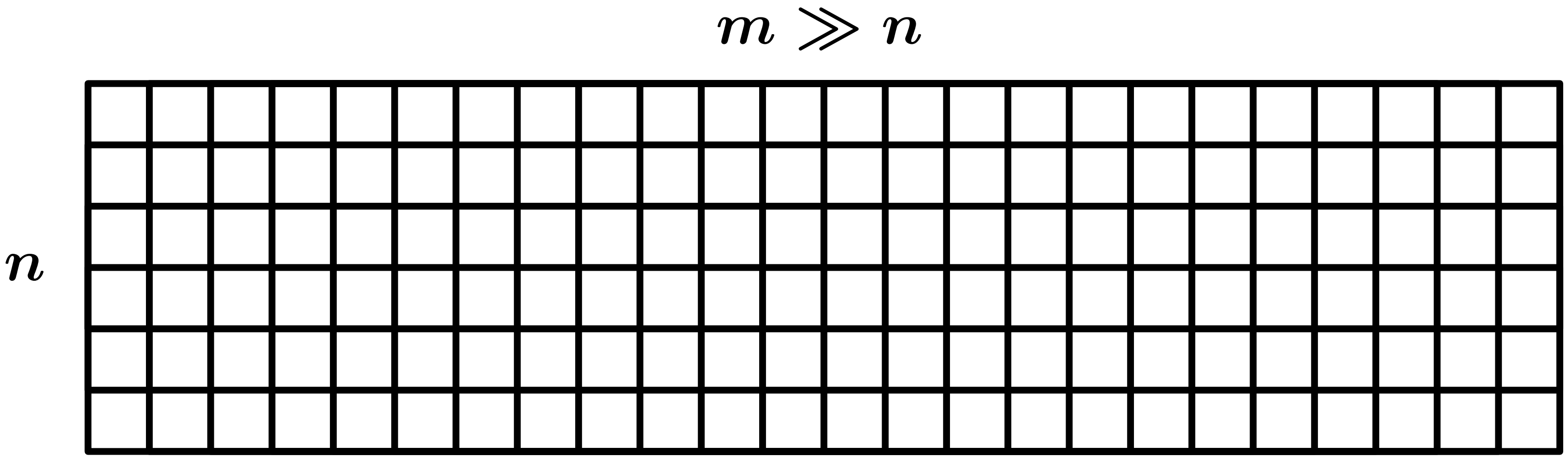
$m \gg n$

n



- Here the dimension decreases as the distance grows from 2 to 1.
- $1/r^2$ distribution does <u>not</u> yield a navigable smallworld

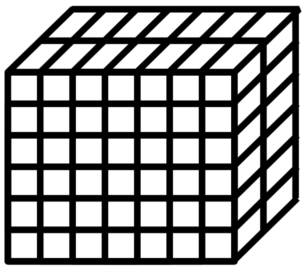
What about elongated grids?



- Here the dimension decreases as the distance grows from 2 to 1.

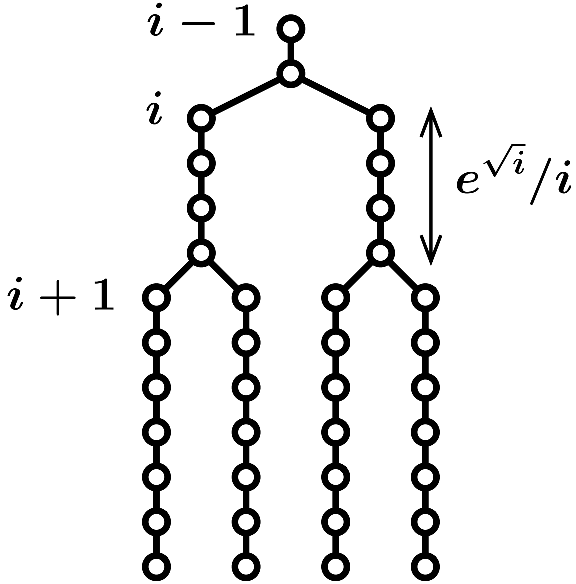
➡ Define the long range link distribution in terms of ball growth, instead of in terms of distance.

Ball growths



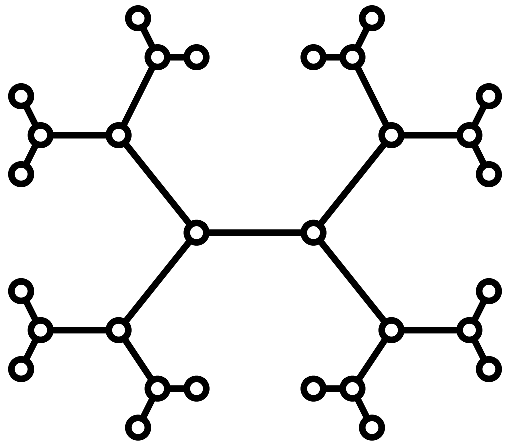
r^d

polynomial growth



$e^{\log^2 r}$

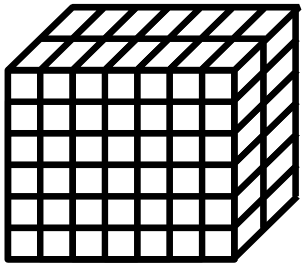
intermediate growth

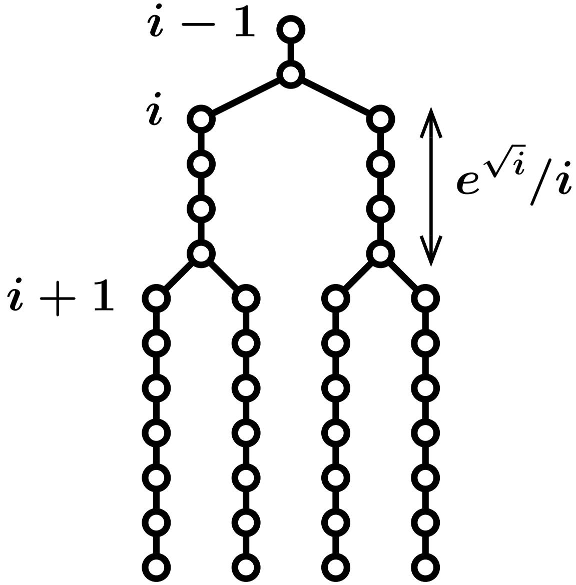


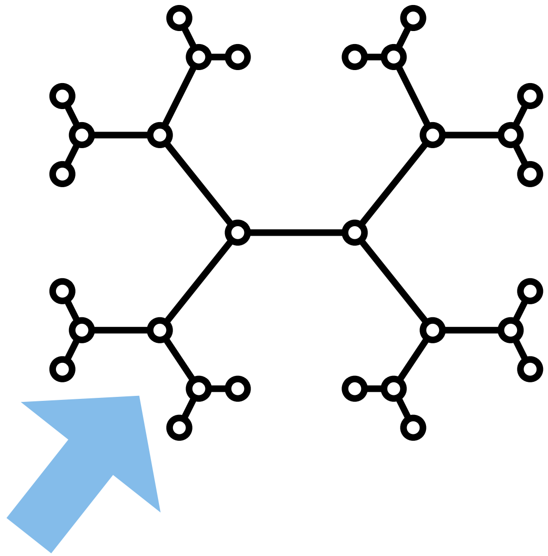
$3 \cdot 2^r$

exponential growth

Ball growths







r^d — polynomial growth

$e^{\log^2 r}$ — intermediate growth

already a smallworld $3 \cdot 2^r$ — exponential growth

Moderate growth

$$B_u(r) = r^{d_u(r)}, \text{ with } \frac{\partial d_u(r)}{\partial r} \leq \frac{\mathrm{C}}{r \log r}$$

Moderate growth

- Theorem [DHLS05] Every infinite graph G (or family of graphs (G_n)) for which the size of the balls centered on every node u satisfies:

$$B_u(r) = r^{d_u(r)}, \text{ with } \frac{\partial d_u(r)}{\partial r} \leq \frac{C}{r \log r}$$

are smallworldizable by adding to every node u, a long range link of length r distributed as $\frac{1}{B_u(r) \log^{1+\epsilon} r}$.

Moderate growth

- Theorem [DHLS05] Every infinite graph G (or family of graphs (G_n)) for which the size of the balls centered on every node u satisfies:

$$B_u(r) = r^{d_u(r)}, \text{ with } \frac{\partial d_u(r)}{\partial r} \leq \frac{C}{r \log r}$$

are smallworldizable by adding to every node u, a long range link of length r distributed as $\frac{1}{B_u(r) \log^{1+\epsilon} r}$.

✓ Example: graphs with "decreasing dimension", or with ball growth upto $r^{O(\log \log r)}$.

Moderate growth

Lemma 1: The distribution is properly defined

- The condition on the derivative of $d_u(r)$ guarantees that the spheres grow smoothly.

$$S_u(r) \leqslant \frac{O(\log\log r)}{r} B_u(r)$$

- The normalizing constant is then:

$$Z_u = \sum_r \frac{S_u(r)}{B_u(r)\log^{1+\epsilon} r} \leqslant \sum_r \frac{O(\log\log r)}{r \log^{1+\epsilon} r} < \infty$$

Moderate growth

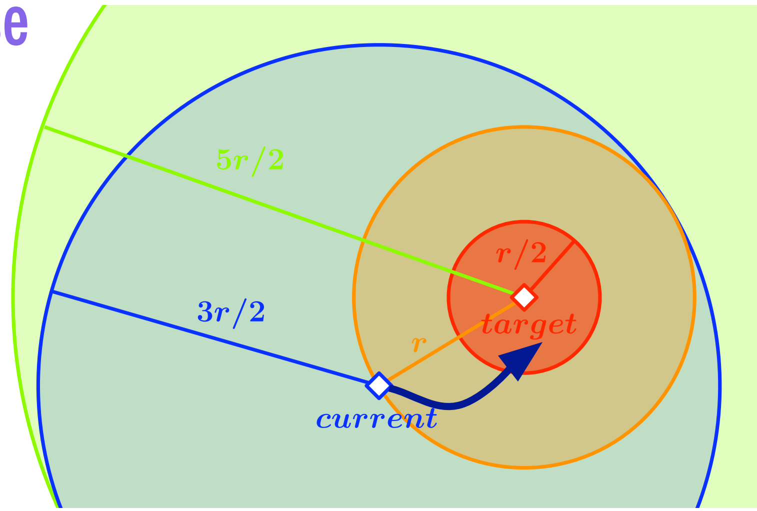
Lemma 2: Greedy routing computes path of expected length $O(\log^{2+\epsilon+C'} r)$ between any pair of nodes at distance r.

- The bound on $d_u(r)$ gives:

$$B(5r/2) \leqslant O(\log^{1+\epsilon+C'} r) B(r/2)$$

- The probability to get twice as close to the target in one step is:

$$\geqslant \frac{B_t(r/2)}{Z \log^{1+\epsilon}(r) B_t(5r/2)}$$
$$= \Omega(1/\log^{1+\epsilon+C'} r)$$



Product of moderate growth

- One particularity of Kleinberg's networks, is their relative independence of the dimension.

Theorem [DHLS05] The cartesian product of moderate growth graphs is smallworldizable by adding one long range link per node of length r distributed as

The expected length of the greedy path is then:

$$O(\log^{2+\epsilon+C'_1+C'_2} r)$$

Product of moderate growth

- One particularity of Kleinberg's networks, is their relative independence of the dimension.

Theorem [DHLS05] The cartesian product of moderate growth graphs is smallworldizable by adding one long range link per node of length r distributed as

$$\frac{1}{B_u(r)\log^{1+\epsilon} r}$$

The expected length of the greedy path is then:

$$O(\log^{2+\epsilon+C_1'+C_2'} r)$$

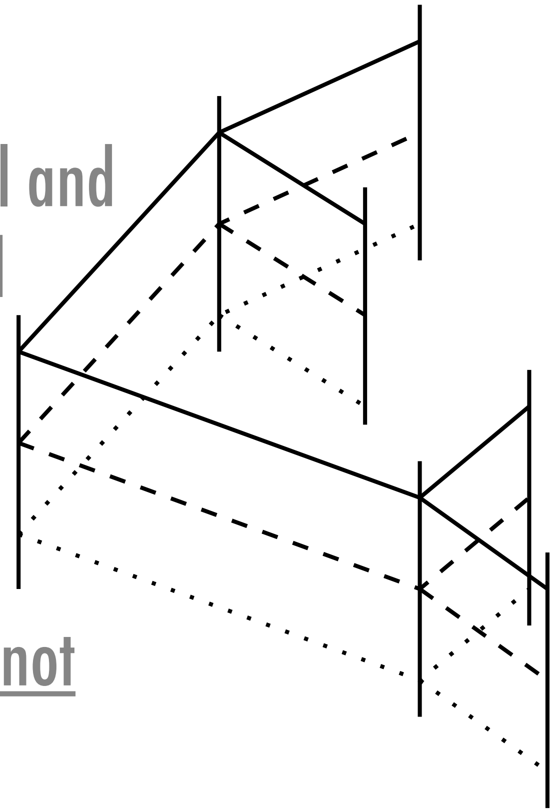
Other smallworldizations

Theorem [Fraigniaud05] Graphs with bounded treewidth are smallworldizable.

Theorem [Martel, NGuyen 05] If their exists a function L, s.t. for all u,v,w, $L(u,w) \leq a(L(u,v)+L(v,w))$, and if $L(u,v) < \log n$ implies that u and v lie in the same connected component, then the graph is smallworldizable.

Conclusion & open questions

- Could $\dfrac{B(\alpha \cdot r)}{B(r)} \leq \mathrm{polylog}(r)$ be sufficient?
- Some graphs are mixture of polynomial and exponential growths. Can we detect and improve the polynomial part?



- Does there exist graph metrics that are not smallworldizable?

Thank you