Network Optimization by Randomization

Exercise: Solutions

Task 1

Vertex Coloring

Vertex Coloring

In the lecture, a simple distributed algorithm ("Reduce") which colors an arbitrary graph with $\Delta + 1$ colors in n synchronous rounds was presented (Δ denotes the largest degree, n the number of nodes of the graph).

a) What is the message complexity, i.e., the total number of messages the algorithm sends in the worst case?

b) Does the algorithm also work in an asynchronous environment? If yes, formulate the asynchronous equivalent to the algorithm, if no, describe why.

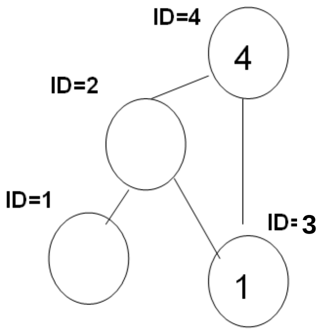
Vertex Coloring

Reduce? wait until all higher-ID neighbors chose color;
take first free and inform neighbors;

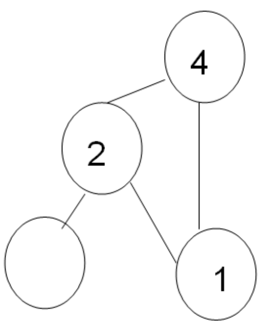
Recall:

First Free

Assume initial coloring (e.g., unique ID=color)
1. Each node uses smallest available color in neighborhood



Assume: two neighbors never choose color at the same time...



Reduce

Initial coloring = IDs
Each node v:
1. v sends ID to neighbors
2. while (v has uncolored neighbor with higher ID)
   1. v sends „undecided“ to neighbors
3. v chooses free color using **First Free**
4. v sends decision to neighbors

Analysis: Which messages are sent?

1. „*undecided*“ must not be sent...

2. So only two messages are sent by a given node over an edge: one at the beginning with ID, and one with chosen color when deciding

So in total? Sum over links (from both endpoints...)

4*|E| messages

Vertex Coloring

Vertex Coloring

In the lecture, a simple distributed algorithm ("Reduce") which colors an arbitrary graph with $\Delta + 1$ colors in n synchronous rounds was presented (Δ denotes the largest degree, n the number of nodes of the graph).

a) What is the message complexity, i.e., the total number of messages the algorithm sends in the worst case?

b) Does the algorithm also work in an asynchronous environment? If yes, formulate the asynchronous equivalent to the algorithm, if no, describe why.

Asynchronous Algorithm?

What assumptions do we need?

E.g.: assume nodes know how many neighbors they have!
(Otherwise you never know whether there will be one more later...)

Idea?

> Wait until all neighbors have replied with ID before starting to compute colors, and only choose color when all higher-ID neighbors have chosen color too (as before...)

Task 2

2 License to Match

In preparation of a highly dangerous mission, the participating agents of the gargantuan Charlottenburger secret service (CSS) need to work in pairs of two for safety reasons. All members in the CSS are organized in a tree hierarchy. Communication is only possible via the official channel: an agent has a secure phone line to his direct superior and a secure phone line to each of his direct subordinates.

Assume that at some time t_0, a subset of agents (from the whole tree hierarchy) learn that they need to take part in a mission. The goal for each "activated" agent is then to find a partner (chosen arbitrarily from the CSS tree), by using the official communication links mentioned above.

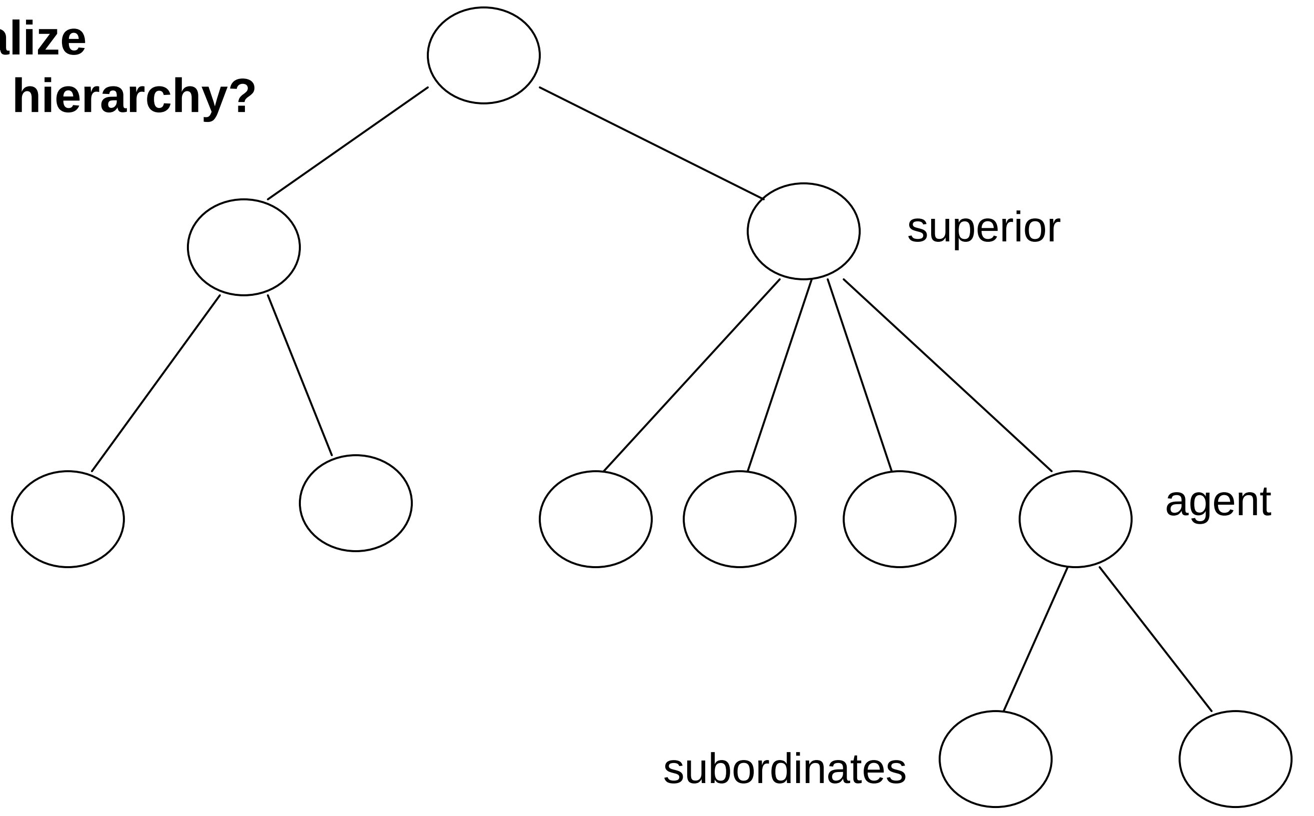
a) Devise an algorithm that will match up a participating agent with another participating agent given the constrained communication scenario. A "match" consists of an agent knowing the identity of his partner and the path in the hierarchy connecting them. Assume that there is an even number of participating agents so that each one is guaranteed a partner. Furthermore, observe that—in the case of an emergency where they lose contact—the phone links connecting two paired-up agents need to remain open at all times. Therefore, you cannot use the same link (i.e., an edge) twice when connecting agents with their partners.

b) What are the time and message (i.e., "phone call") complexities of your algorithm?

Communication over trees

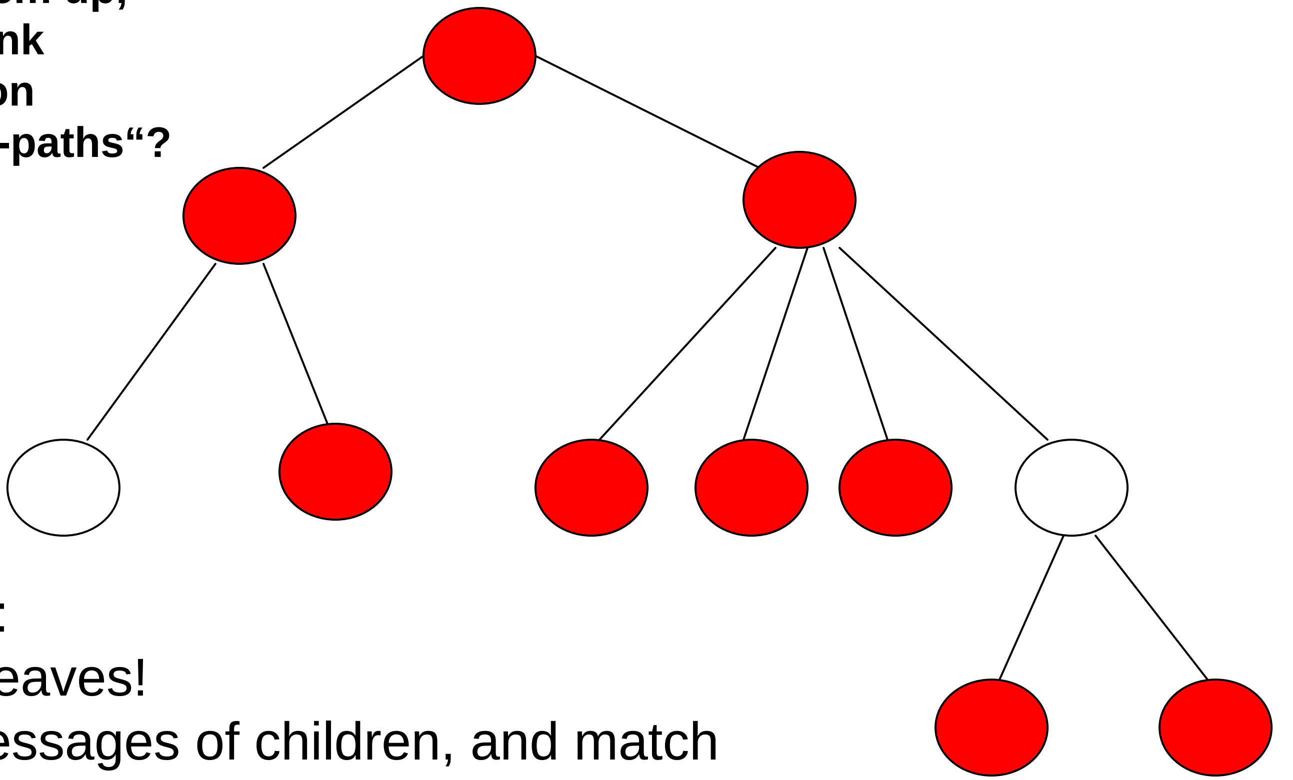
How to visualize organization hierarchy?

Tree!



Agents activated!

How to pair them up, such that no link is used twice on these „partner-paths"?

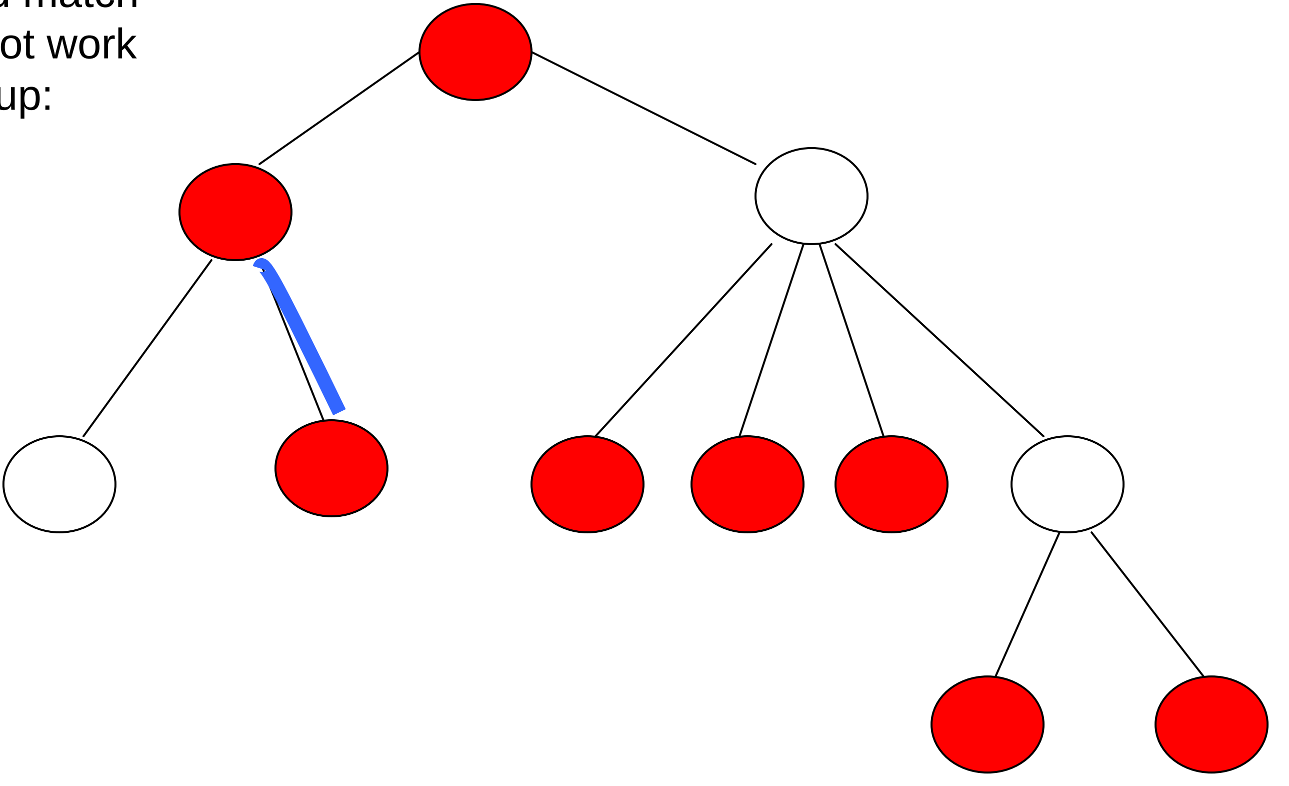


Idea „Echo":
- start with leaves!
- wait for messages of children, and match them; if does not work out propagate up: at most one!

Example here?

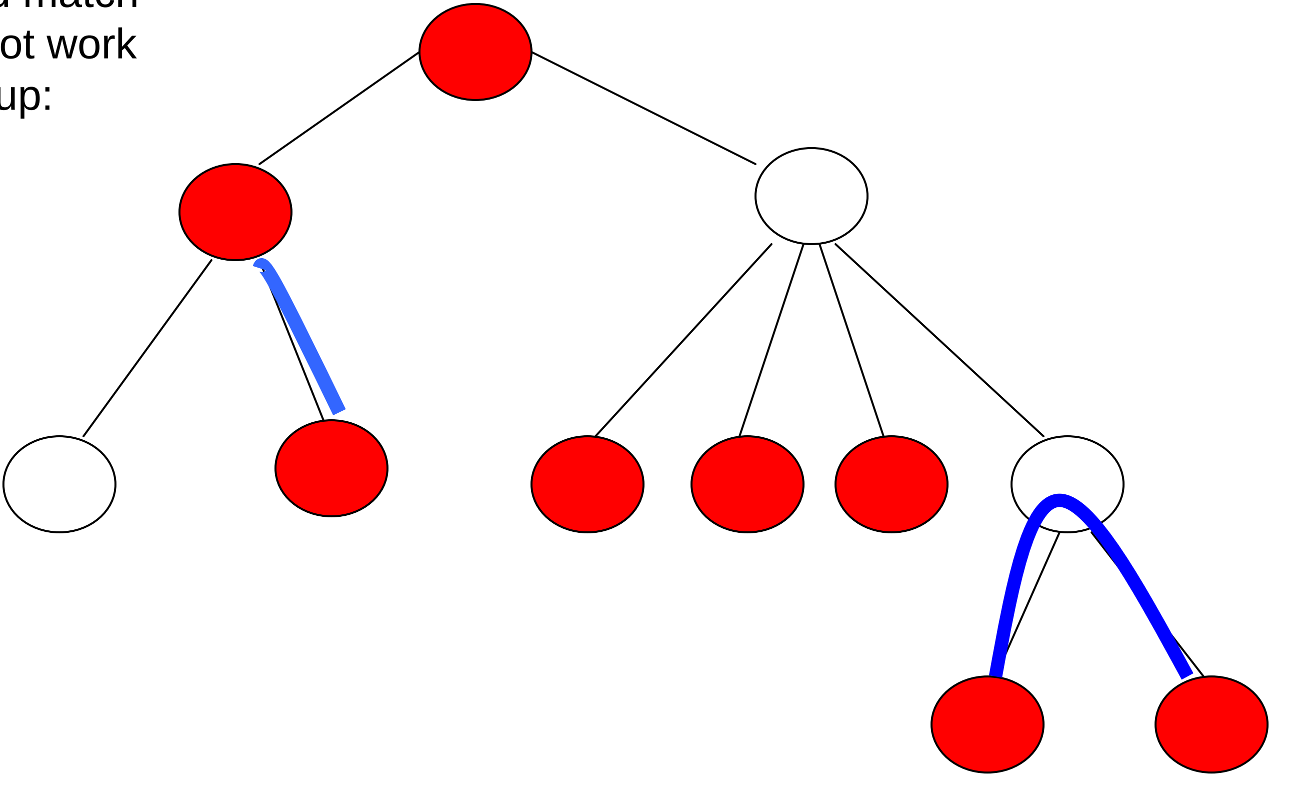
Echo Algorithm

Echo: wait for messages of children, and match them; if does not work out propagate up: at most one!



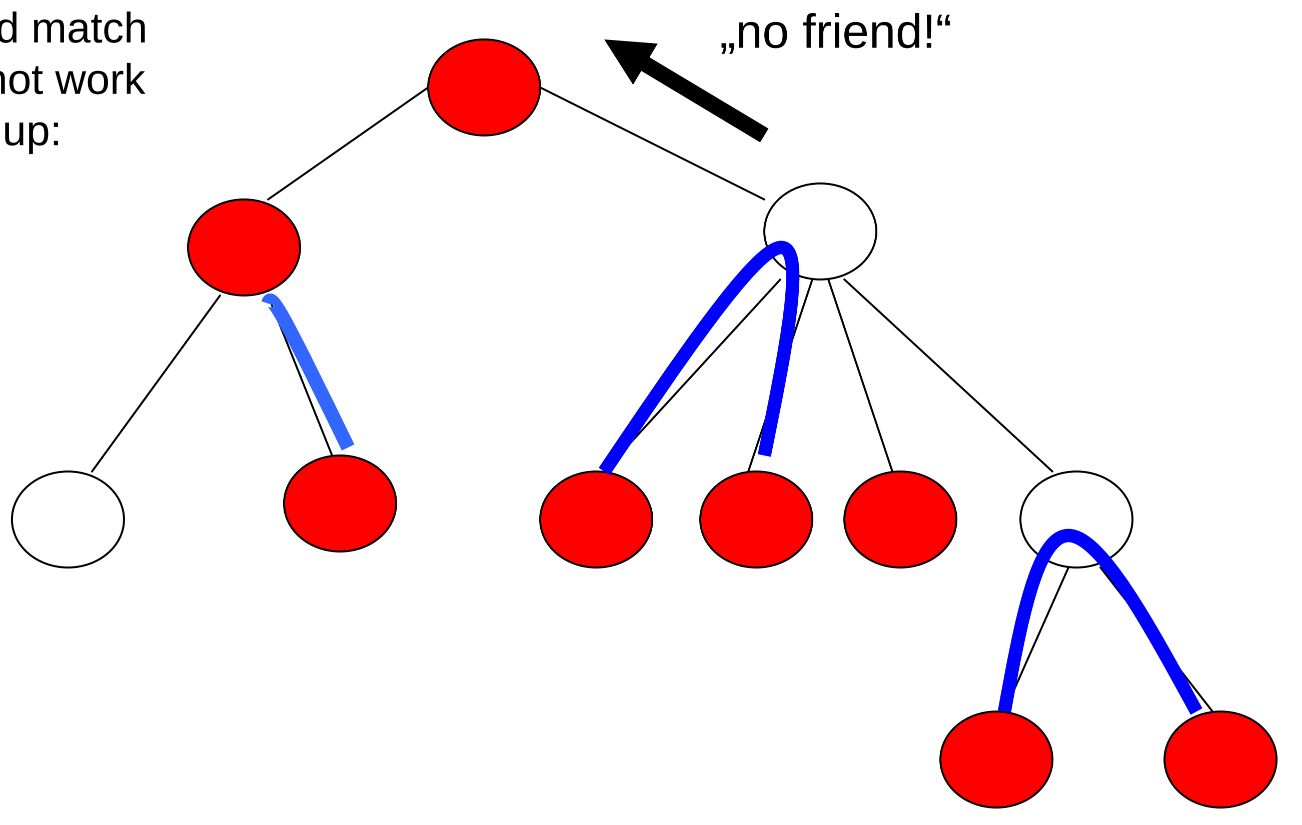
Echo Algorithm

Echo: wait for messages of children, and match them; if does not work out propagate up: at most one!



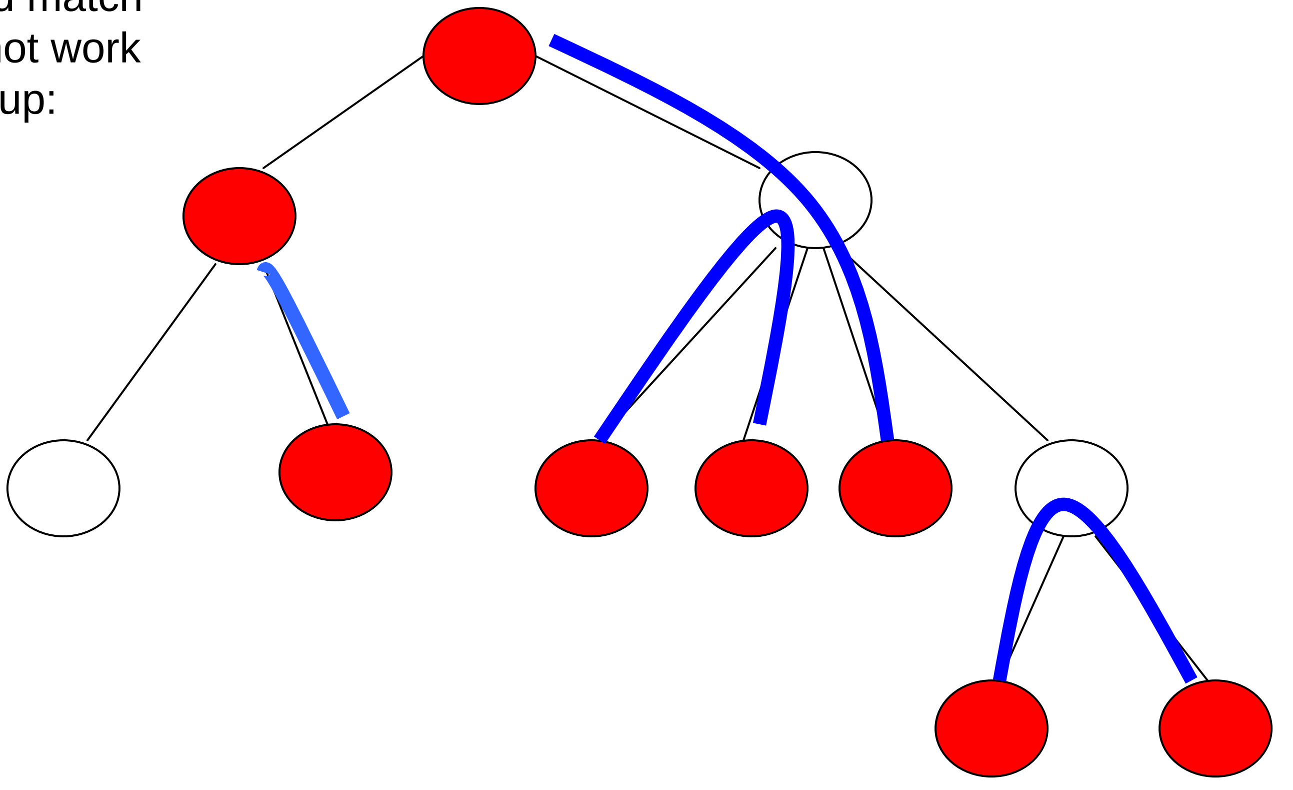
Echo Algorithm

Echo: wait for messages of children, and match them; if does not work out propagate up: at most one!



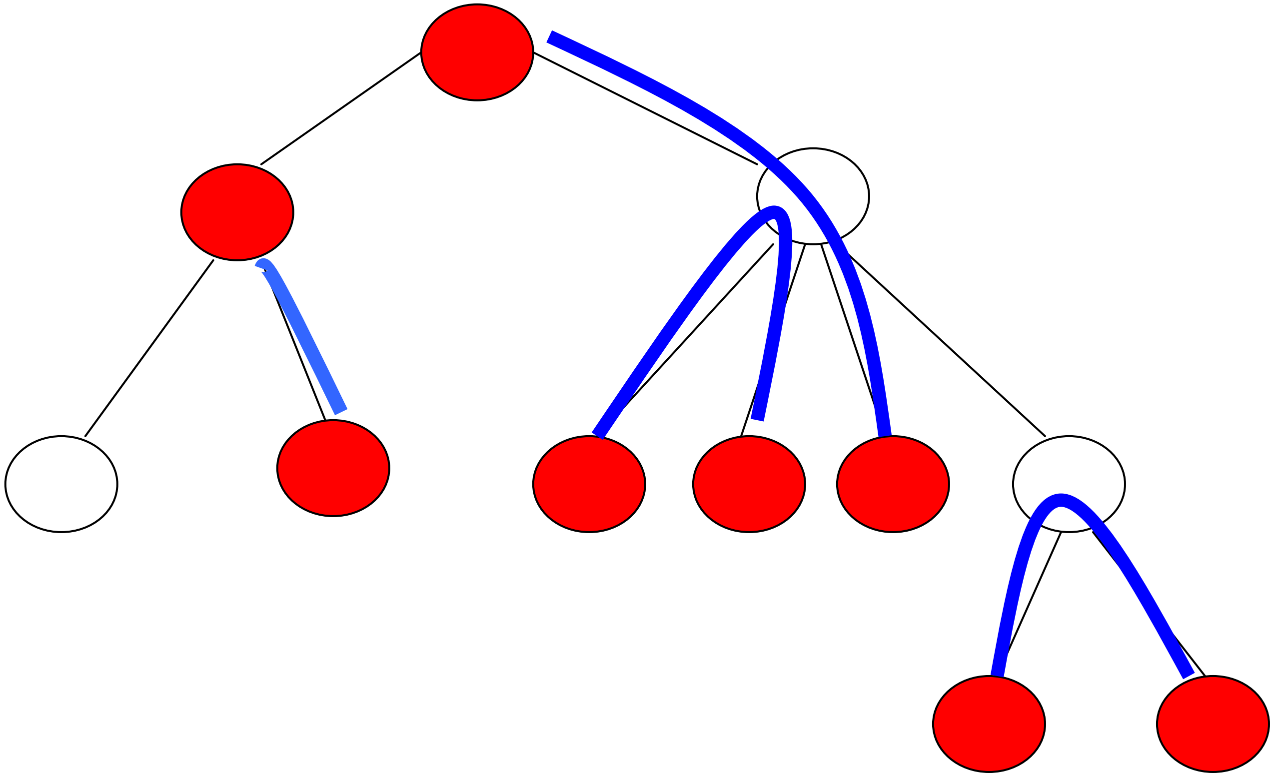
Echo Algorithm

Echo: wait for messages of children, and match them; if does not work out propagate up: at most one!



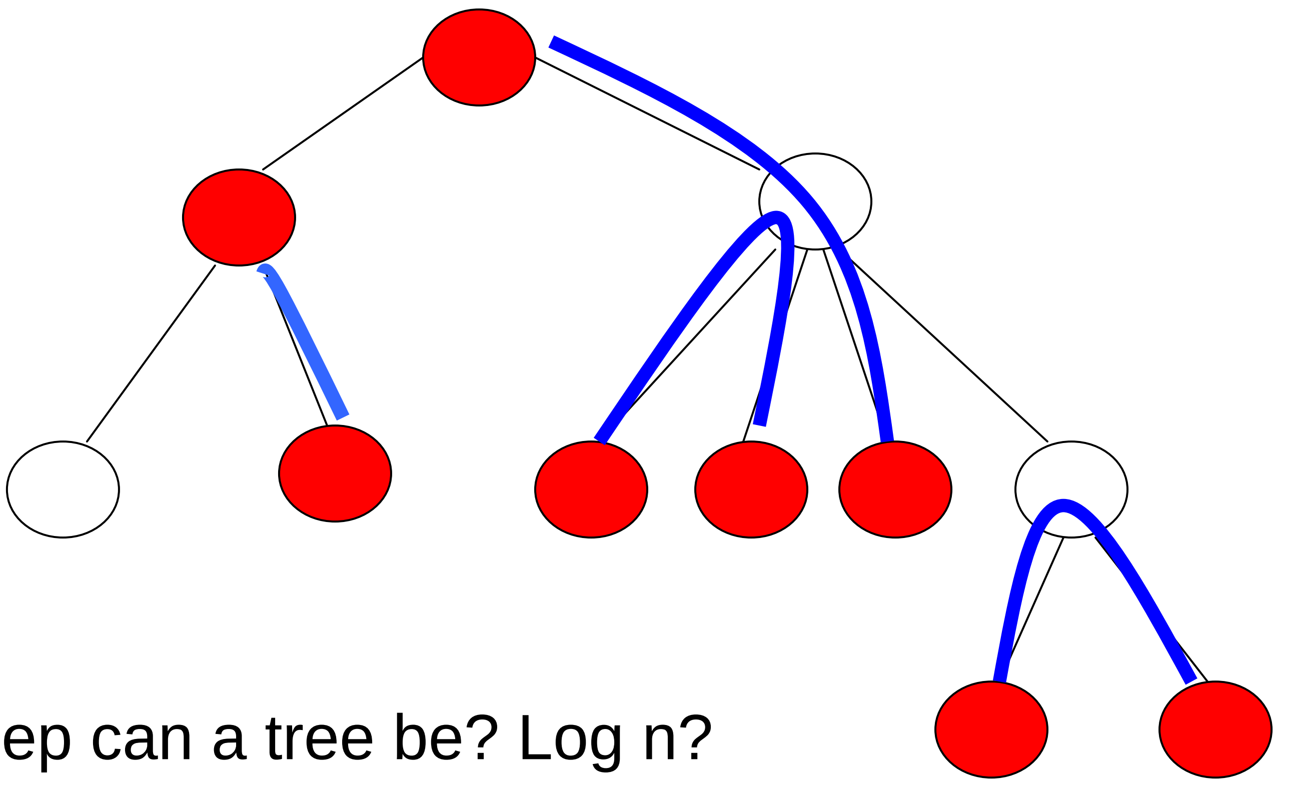
It works! ☺

Each link used only once: at most one node in subtree unmatched, so link available!



Complexity?

Time O(Depth), and at most two messages along link (at most one request and one reply per subtree).



How deep can a tree be? Log n?

Task 3

Diameter of the Augmented Grid

Diameter of the Augmented Grid

Recall the network from the lecture where nodes were arranged in a grid and each node had an additional directed link to an uniformly and independently at random drawn node in the network (i.e., $\alpha = 0$). In the lecture, a proof of the fact that such a network has diameter $O(\log n)$ w.h.p. was sketched. We will now fill in the details.

a) Show that $O(cn/\log n)$ many nodes are enough to guarantee with high probability that at least one of their random links connects to a given set of $\Omega(\log^2 n)$ nodes. Prove this *(i)* by direct calculation and *(ii)* using Chernoff's bound.

Hint: Use that $1 - p \leq e^{-p}$ for any p.

b) Suppose for some node set S we have that $|S| \in \Omega((\log n)^2) \cap o(n)$ and denote by H the set of nodes hit by their random links. Prove that H and together with its grid neighbors contains w.h.p. $(5 - o(1))|S|$ nodes!

Hint: Observe that *independently* of all previous random choices, each new link has at least a certain probability p of connecting to a node whose complete neighborhood has not been reached yet. Then use Chernoff's bound on the sum of $|S|$ many variables.

c) Infer from **b)** that starting from $\Omega(\log^2 n)$ nodes, with each hop the number of reached nodes w.h.p. more than doubles, as long as we have still $O(n/\log n)$ nodes (regardless of the constants in the O-notation).

Hint: Play with the constant c in the definition of w.h.p. and use the union bound.

d) Conclude that the diameter of the network is w.h.p. in $O(\log n)$.

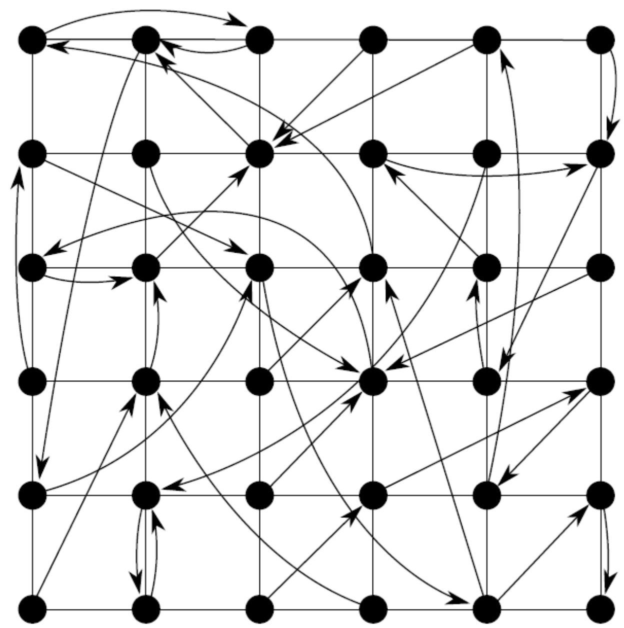
Recall

Augmented Grid

Consider an (m × m) grid of $n=m^2$ nodes, where each node has a directed edge to each lattice neighbor (local contacts). In addition, each node has an additional random link (long-range contact).
For all u and v, the long-range contact of u to points to node v with probability $d(u,v)^{-\alpha}/\sum_{w \in V \setminus \{u\}} d(u,w)^{-\alpha}$, where d() is the distance in the grid and α is a parameter.

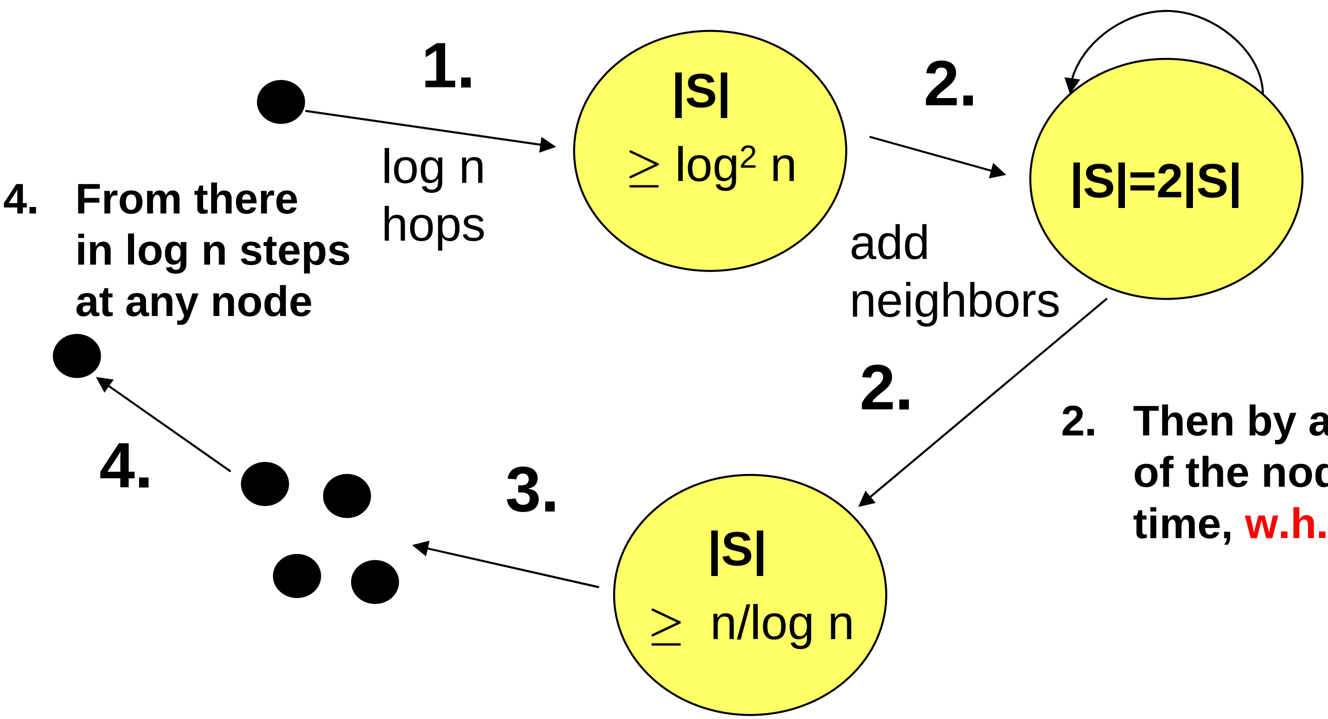
$\alpha=0$?

Means uniform (indep. of distance)!



Proof Strategy

1. **Starting from one node, we can reach $\log^2 n$ many nodes in log n hops (along grid...)**

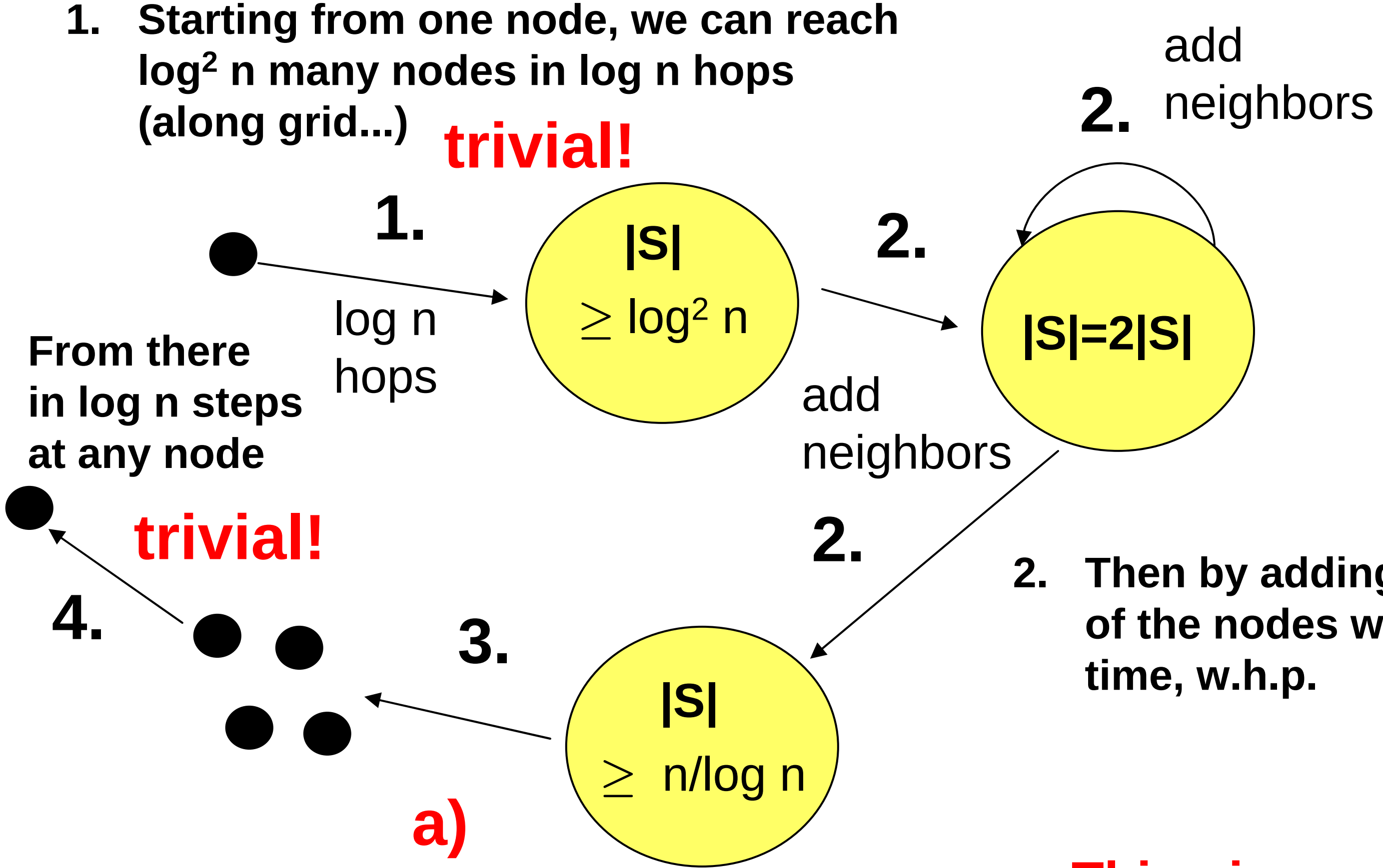


2. **Then by adding the neighborhood of the nodes we double the size each time, w.h.p.**

3. **From there we can reach any log n neighborhood ($\log^2 n$ many nodes) of any node, w.h.p.**

Proof Strategy

1. **Starting from one node, we can reach $\log^2 n$ many nodes in log n hops (along grid...)**



4. **From there in log n steps at any node**

2. **Then by adding the neighborhood of the nodes we double the size each time, w.h.p.**

b)+c)

3. **From there we can reach any log n neighborhood ($\log^2 n$ many nodes) of any node, w.h.p.**

This gives a path between two given nodes, w.h.p.: Diameter = all paths w.h.p.!
=> Exercise d)

Diameter of the Augmented Grid

Diameter of the Augmented Grid

Recall the network from the lecture where nodes were arranged in a grid and each node had an additional directed link to an uniformly and independently at random drawn node in the network (i.e., $\alpha = 0$). In the lecture, a proof of the fact that such a network has diameter $O(\log n)$ w.h.p. was sketched. We will now fill in the details.

a) Show that $O(cn/\log n)$ many nodes are enough to guarantee with high probability that at least one of their random links connects to a given set of $\Omega(\log^2 n)$ nodes. Prove this (i) by direct calculation and (ii) using Chernoff's bound.

Hint: Use that $1 - p \leq e^{-p}$ for any p.

b) Suppose for some node set S we have that $|S| \in \Omega((\log n)^2) \cap o(n)$ and denote by H the set of nodes hit by their random links. Prove that H and together with its grid neighbors contains w.h.p. $(5 - o(1))|S|$ nodes!

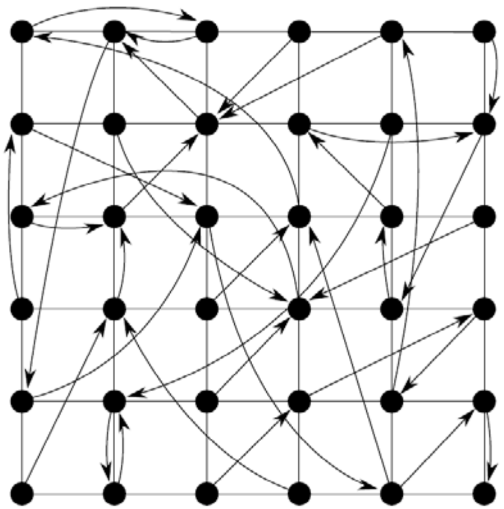
Hint: Observe that *independently* of all previous random choices, each new link has at least a certain probability p of connecting to a node whose complete neighborhood has not been reached yet. Then use Chernoff's bound on the sum of $|S|$ many variables.

c) Infer from **b)** that starting from $\Omega(\log^2 n)$ nodes, with each hop the number of reached nodes w.h.p. more than doubles, as long as we have still $O(n/\log n)$ nodes (regardless of the constants in the O-notation).

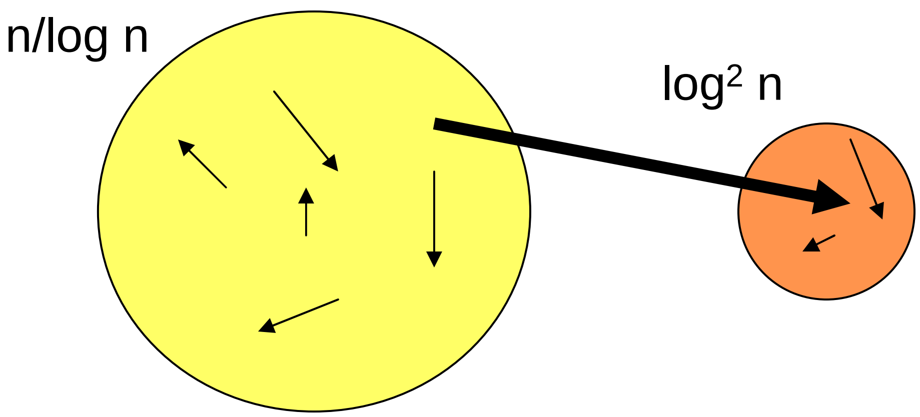
Hint: Play with the constant c in the definition of w.h.p. and use the union bound.

d) Conclude that the diameter of the network is w.h.p. in $O(\log n)$.

High Hitting Probability



O(Cn/log n) nodes enough to guarantee random link into a set of $\Omega(\log^2 n)$ nodes.



Probability p that a given link connects into this set S?

$\alpha=0$, so uniform, so $p \in \Omega(\log^2 n/n)$.

The Cn/log n nodes have a random link each. Probability that none hits S?

$$(1-p)^{cn/\log n} \leq e^{-pcn/\log n} \in e^{-\Omega(c \log n)} = \frac{1}{n^{\Omega(c)}}$$

whp!

QED

Alternative Proof

Definition 2.13 (Chernoff's Bound). *Let* $X = \sum_{i=1}^{k} X_i$ *be the sum of* k *independent* $0-1$ *random variables. Then* Chernoff's bound *states that w.h.p.*

$$|X - \mathbb{E}[X]| \in O\left(\log n + \sqrt{\mathbb{E}[X]\log n}\right).$$

O(Cn/log n) nodes enough to guarantee random link into a set of $\Omega(\log^2 n)$ nodes.

Let X_i be the indicator variable whether the i-th link hits the set S, for $i \in \{1,...,l\}$ for some $l \in O(n/\log n)$. Let X be the sum of the X_i. Let $p \in \Omega(\log^2 n/n)$ denote probability that a link hits the set (see before). So E[X]?

$$\mathbf{E[X] = p \cdot l \geq C \log n}, \text{ for some constant C... So?}$$

So according to Chernoff: **$| X - E[X] | < E[X]$**, w.h.p.

Since **$P[X>0] = P[\, | X - E[X]| < E[X]\,]$ = „w.h.p.“**

the claim follows by choosing the constant accordingly.

QED

Diameter of the Augmented Grid

Diameter of the Augmented Grid

Recall the network from the lecture where nodes were arranged in a grid and each node had an additional directed link to an uniformly and independently at random drawn node in the network (i.e., $\alpha = 0$). In the lecture, a proof of the fact that such a network has diameter $O(\log n)$ w.h.p. was sketched. We will now fill in the details.

a) Show that $O(cn/\log n)$ many nodes are enough to guarantee with high probability that at least one of their random links connects to a given set of $\Omega(\log^2 n)$ nodes. Prove this (i) by direct calculation and (ii) using Chernoff's bound.

Hint: Use that $1 - p \leq e^{-p}$ for any p.

b) Suppose for some node set S we have that $|S| \in \Omega((\log n)^2) \cap o(n)$ and denote by H the set of nodes hit by their random links. Prove that H and together with its grid neighbors contains w.h.p. $(5 - o(1))|S|$ nodes!

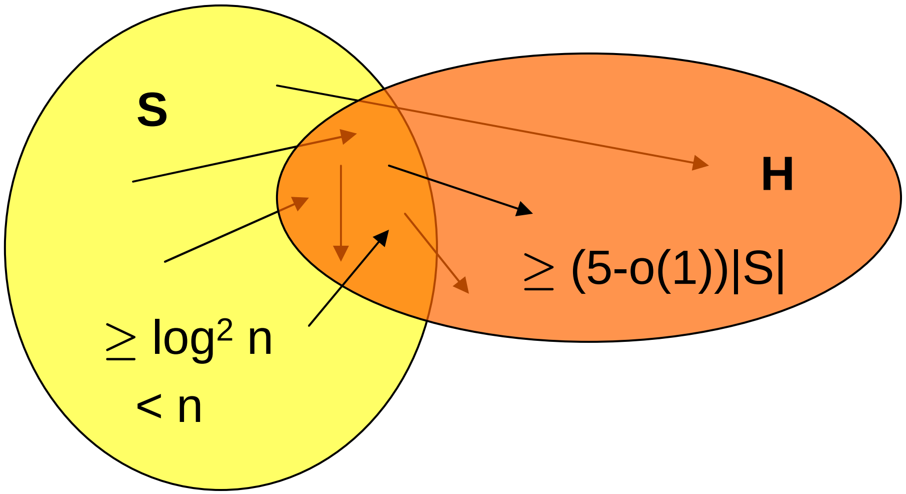
Hint: Observe that *independently* of all previous random choices, each new link has at least a certain probability p of connecting to a node whose complete neighborhood has not been reached yet. Then use Chernoff's bound on the sum of $|S|$ many variables.

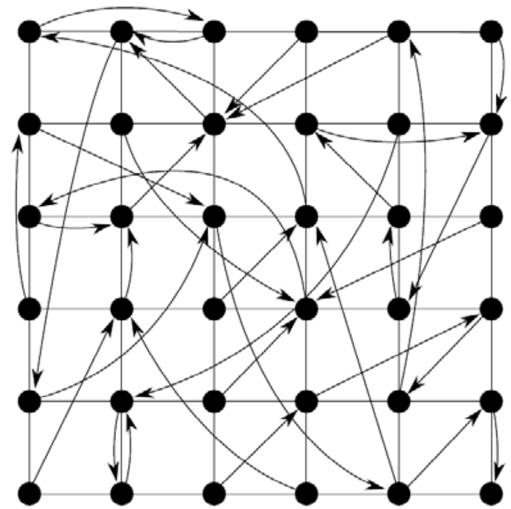
c) Infer from **b)** that starting from $\Omega(\log^2 n)$ nodes, with each hop the number of reached nodes w.h.p. more than doubles, as long as we have still $O(n/\log n)$ nodes (regardless of the constants in the O-notation).

Hint: Play with the constant c in the definition of w.h.p. and use the union bound.

d) Conclude that the diameter of the network is w.h.p. in $O(\log n)$.

High Diversity





Since $|S| \in o(n)$, the union of S with all grid neighbors and random neigbors of nodes in S is also in o(n).

So there are (1-o(1))n nodes (less than a linear part is missing!) that were not visited (and nor did their neighbors).

So with **probability $p \in 1-o(1)$** any link will give 5 new nodes (incl. grid neighbors).

We can apply Chernoff bounds to these random variables! Let X be the sum of such „good choices" with 5 new nodes.

Then Chernoff bound: since $E[X] > \log n$ (as $|S|$ is $> \log^2 n$ and each has constant and independent probability) it holds w.h.p.

QED

Diameter of the Augmented Grid

Diameter of the Augmented Grid

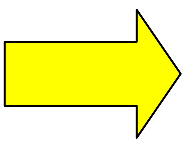
Recall the network from the lecture where nodes were arranged in a grid and each node had an additional directed link to an uniformly and independently at random drawn node in the network (i.e., $\alpha = 0$). In the lecture, a proof of the fact that such a network has diameter $O(\log n)$ w.h.p. was sketched. We will now fill in the details.

a) Show that $O(cn/\log n)$ many nodes are enough to guarantee with high probability that at least one of their random links connects to a given set of $\Omega(\log^2 n)$ nodes. Prove this (i) by direct calculation and (ii) using Chernoff's bound.

Hint: Use that $1 - p \leq e^{-p}$ for any p.

b) Suppose for some node set S we have that $|S| \in \Omega((\log n)^2) \cap o(n)$ and denote by H the set of nodes hit by their random links. Prove that H and together with its grid neighbors contains w.h.p. $(5 - o(1))|S|$ nodes!

Hint: Observe that *independently* of all previous random choices, each new link has at least a certain probability p of connecting to a node whose complete neighborhood has not been reached yet. Then use Chernoff's bound on the sum of $|S|$ many variables.



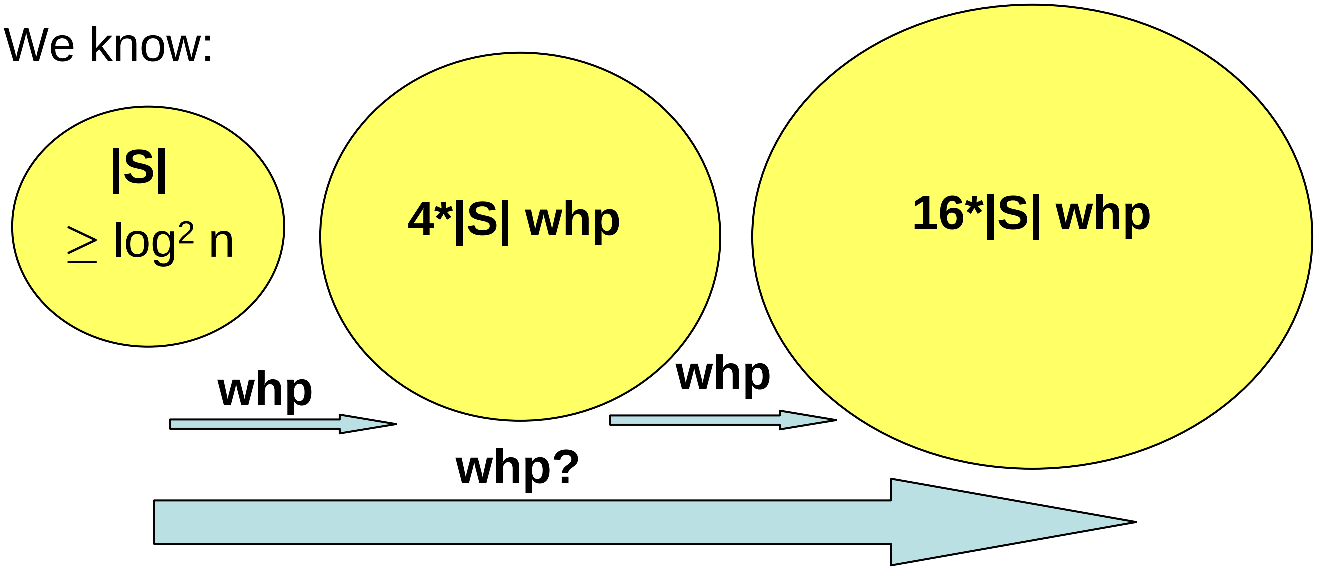
c) Infer from **b)** that starting from $\Omega(\log^2 n)$ nodes, with each hop the number of reached nodes w.h.p. more than doubles, as long as we have still $O(n/\log n)$ nodes (regardless of the constants in the O-notation).

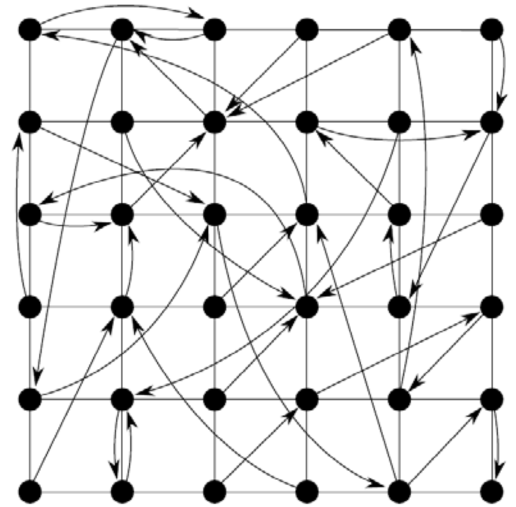
Hint: Play with the constant c in the definition of w.h.p. and use the union bound.

d) Conclude that the diameter of the network is w.h.p. in $O(\log n)$.

Fast Growth

We know:





We know that the set grows at least by a factor (5-o(1)) in each step w.h.p., so only a fraction of $1/n^c$ goes wrong (growth factor less than 4).

We can double the size at most log n many times. The total fraction that goes wrong is bounded by?

$$\log n / n^c < 1 / n^{c'}$$

for some other constant c‘. So it‘s still w.h.p.!

QED

Diameter of the Augmented Grid

Diameter of the Augmented Grid

Recall the network from the lecture where nodes were arranged in a grid and each node had an additional directed link to an uniformly and independently at random drawn node in the network (i.e., $\alpha = 0$). In the lecture, a proof of the fact that such a network has diameter $O(\log n)$ w.h.p. was sketched. We will now fill in the details.

a) Show that $O(cn/\log n)$ many nodes are enough to guarantee with high probability that at least one of their random links connects to a given set of $\Omega(\log^2 n)$ nodes. Prove this (i) by direct calculation and (ii) using Chernoff's bound.

Hint: Use that $1 - p \leq e^{-p}$ for any p.

b) Suppose for some node set S we have that $|S| \in \Omega((\log n)^2) \cap o(n)$ and denote by H the set of nodes hit by their random links. Prove that H and together with its grid neighbors contains w.h.p. $(5 - o(1))|S|$ nodes!

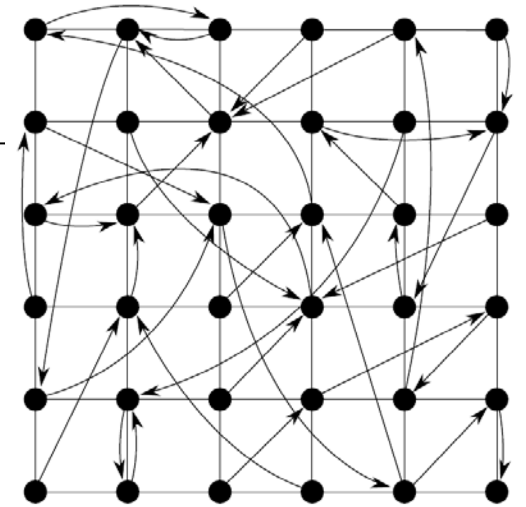
Hint: Observe that *independently* of all previous random choices, each new link has at least a certain probability p of connecting to a node whose complete neighborhood has not been reached yet. Then use Chernoff's bound on the sum of $|S|$ many variables.

c) Infer from **b)** that starting from $\Omega(\log^2 n)$ nodes, with each hop the number of reached nodes w.h.p. more than doubles, as long as we have still $O(n/\log n)$ nodes (regardless of the constants in the O-notation).

Hint: Play with the constant c in the definition of w.h.p. and use the union bound.

d) Conclude that the diameter of the network is w.h.p. in $O(\log n)$.

Diameter



We know that:

1. Each node can reach $\Omega(\log^2 n)$ nodes in the grid (w/o random links) within **log n steps**;
2. Starting from these nodes, $\Theta(n/\log n)$ nodes can be reached within **log n more hops** w.h.p. (just seen in c)!)
3. From these nodes we can reach (log n)-hop neighborhood of every node (on grid), because that‘s $\log^2 n$ given nodes
4. From there we can reach all other nodes in **log n hops** on the grid!
5. So total path between two given nodes has logarithmic length w.h.p.! Are we done?
6. No, we must prove that for all pairs of nodes such a path exists w.h.p.! How many paths are there? At most a polynomial number!
   So the „w.h.p.“ holds for all paths (exponential always swallows polynomial with the right exponent...): in O(log n)!

QED