Perspectives on Small world: Will artificial models explain real networks?

DYNAMO COST 295

SMALL WORLD WORKING GROUP

N. Schabanel

Lisbonne, July 1st 2006

On “real” networks

- **Original properties of «real» networks:**
  - Linear number of edges
  - Abnormally high clustering coefficient (local density)
  - Heavy tailed degree distribution
  - Small diameter
  - Components of small doubling dimension

On “real” networks

- **Excellent performances in spite of:**
  - Extremely **partial** knowledge on the global network
  - **Undesigned** construction process

On "real" networks

- **Two central questions:**
  - Obtaining random models that reproduce these properties
  - How do their functionalities constrain the structures of real network?

On "real" networks

- **Real networks play precise roles under constraints:**
  - They fulfill a function
  - Information networks: design based on trade-off cost/performances
  - Social networks: trade-off sympathy / limited time & vision / efficiency

Small world property

- In spite of ...
  - a very large size
  - an extremely restricted local knowledge
  - a totally decentralized construction
- ... it is possible to reach very efficiently a previously unknown person in an globally unknown network

Small worlds Algorithmic

- **A central question in distributed algorithms:** routing efficiently with a extremely restricted decentralized knowledge
- Is it possible for all graphs?

Small worlds Algorithmic

- First model exhibiting algorithmic aspects of smallworld phenomemon (Kleinberg 2000) :
  - **2D mesh** (geographic positions)
  - **augmented with random links** (acquaintances met during lifetime)
- Smallworld phenomenon araises only for particular distribution of extra random links

Kleinberg's law

[K2000, BFKK2001, DHLS2005]

- Random contact of **u** should be chosen at a distance **r with probability 1/#B(u,r)** so that routing gets efficient, where #B(u,r) is the number of nodes at geographic distance ≤ **r** of **u**
- It garantees that one can access at all **distance scales**
- Has inspired a lot of peer-to-peer protocols (Chord, Papillon,...)

Small worlds Algorithmic

- **Generalization to other class of graphs :**
  - Bounded treewidth (F 2005)
  - Bounded doubling dimension (S 2005)
  - Bounded growth (DHLS 2005)
  - Minor excluding (IG 2006)
- **Non-“smallworldizable” metrics** exist (FLL 2006)
- These models do **not** provide **insights** on the emergence of phenomenon **(not decentralized)**

Very large graphs, slow connections

- **How can such a global property emerge in a decentralized manner?** even more with very small memory and very short time?
- Is there some "natural" process that generates the "Kleinberg's law"?

✓ It is possible to generate "Kleinberg's law"-like random links in **polylog time, polylog messages exchanges** per node and **polylog memory** only! [DHLS2006]

No convincing explanations yet

- Needs for realistics models
- Realistic generation: Could friends be met thru random walks? Other natural process?
- How to include temporal aspects?
- Is there really two types of friends: geographic & random? Is there really an underlying metric?
- Link with degree distribution? clustering coefficient?

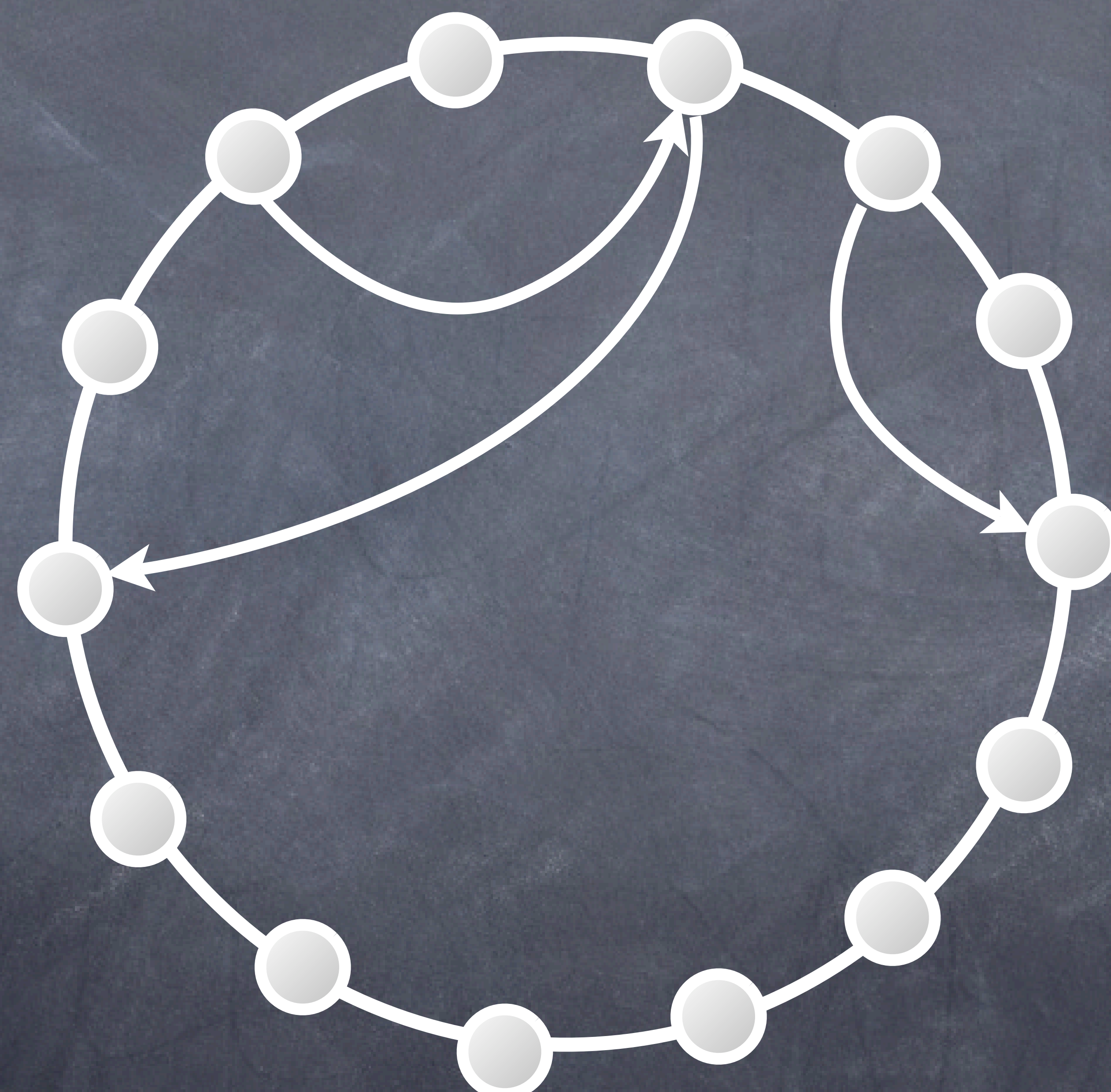
Some results in that direction

- [DHLS 2006] polylog time, memory & number of messages, add only one extra link per node... but
  - requires a lot of coordination between nodes (hierarchy)
  - not dynamic, nor really natural
- [CM2003] interesting empirical result based on rewiring
- Related to recent metric embedding results

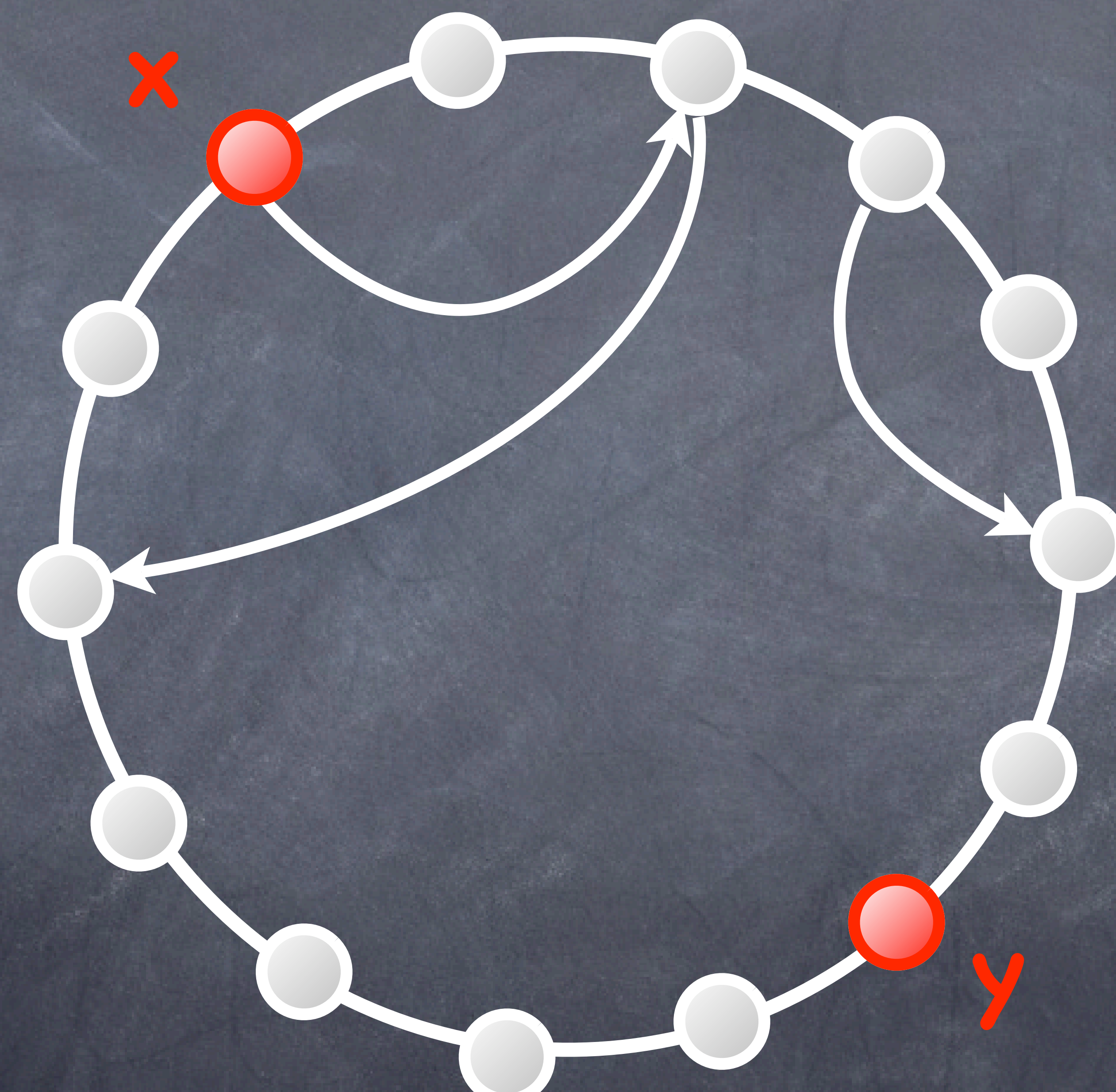
[CM2003] Rewiring scheme

- **Start** with a cycle where each node **x** has an extra link pointing to a arbitrary other node **l(x)**
- **Repeat:**
  - **Select** a random pair **x,y**
  - **Select** a random threshold **T** **uniformly** at random in **{1,..,d(x,y)}**
  - **Go** greedily from **x** to **y** **for at most T steps**
  - **If stop** at **z ≠ y**, **x** "**bookmarks**" **z**, i.e. **l(x) := z**

Example

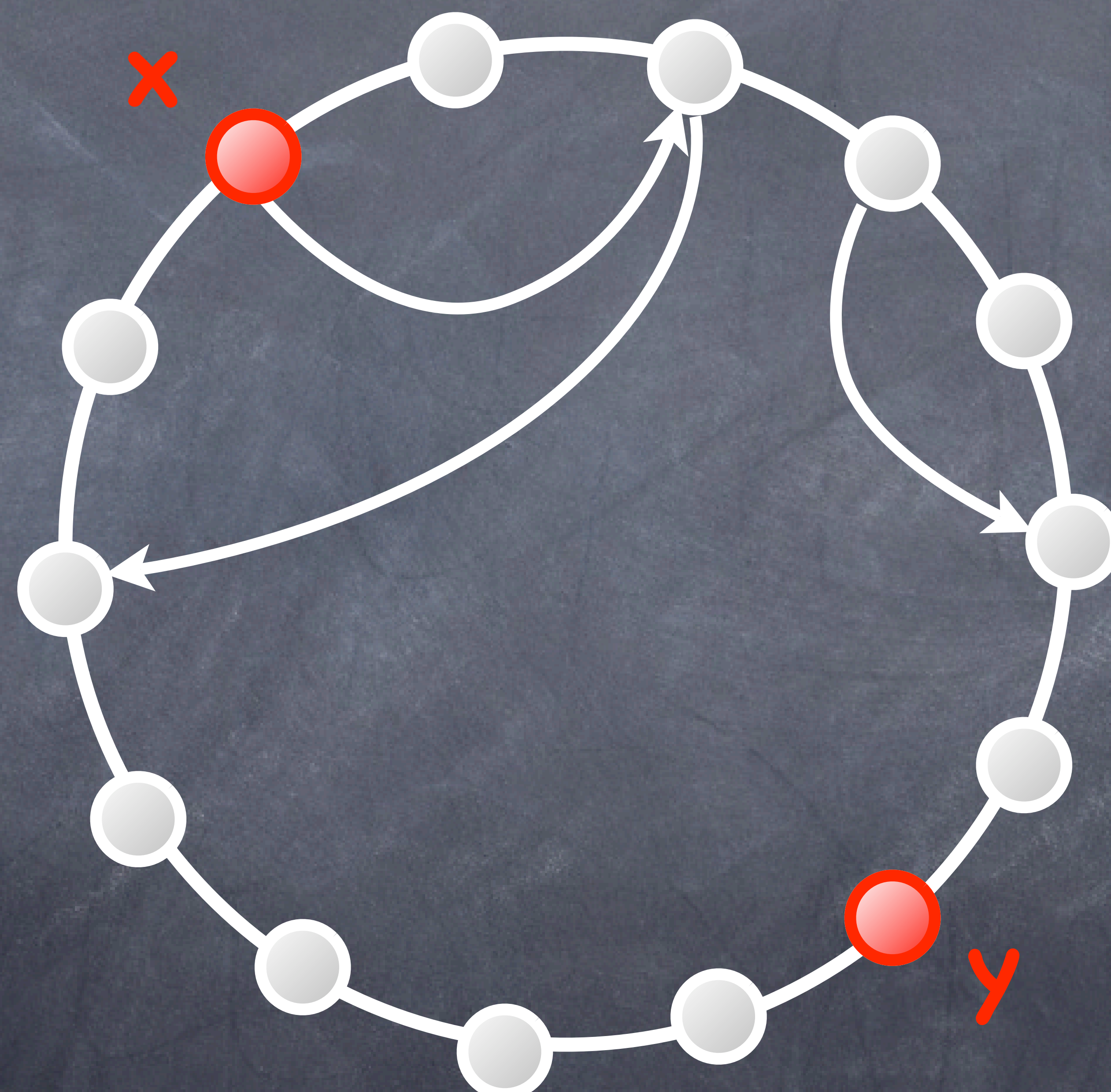


Example



Example

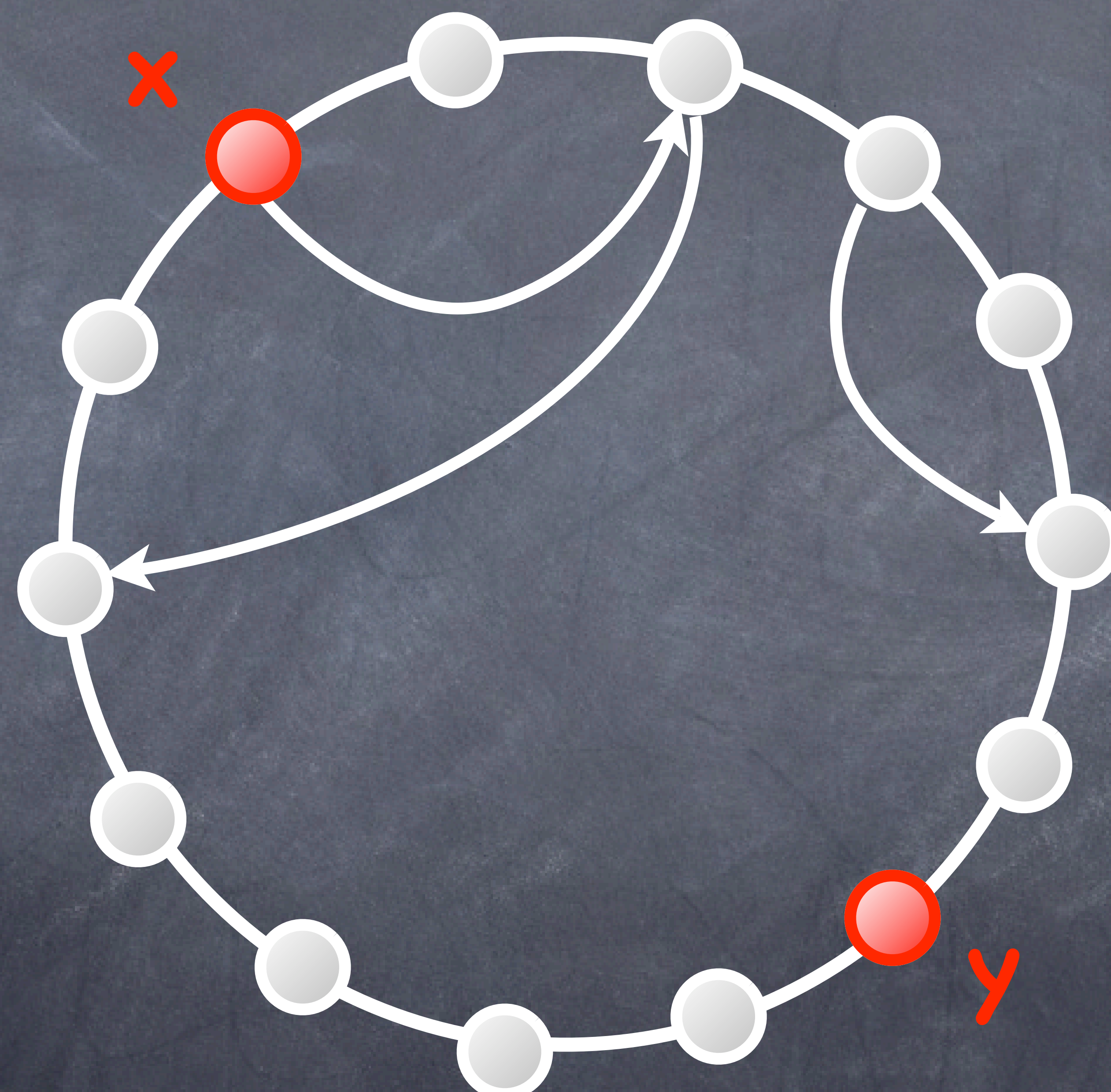
d(x,y) = 7



Example

d(x,y) = 7

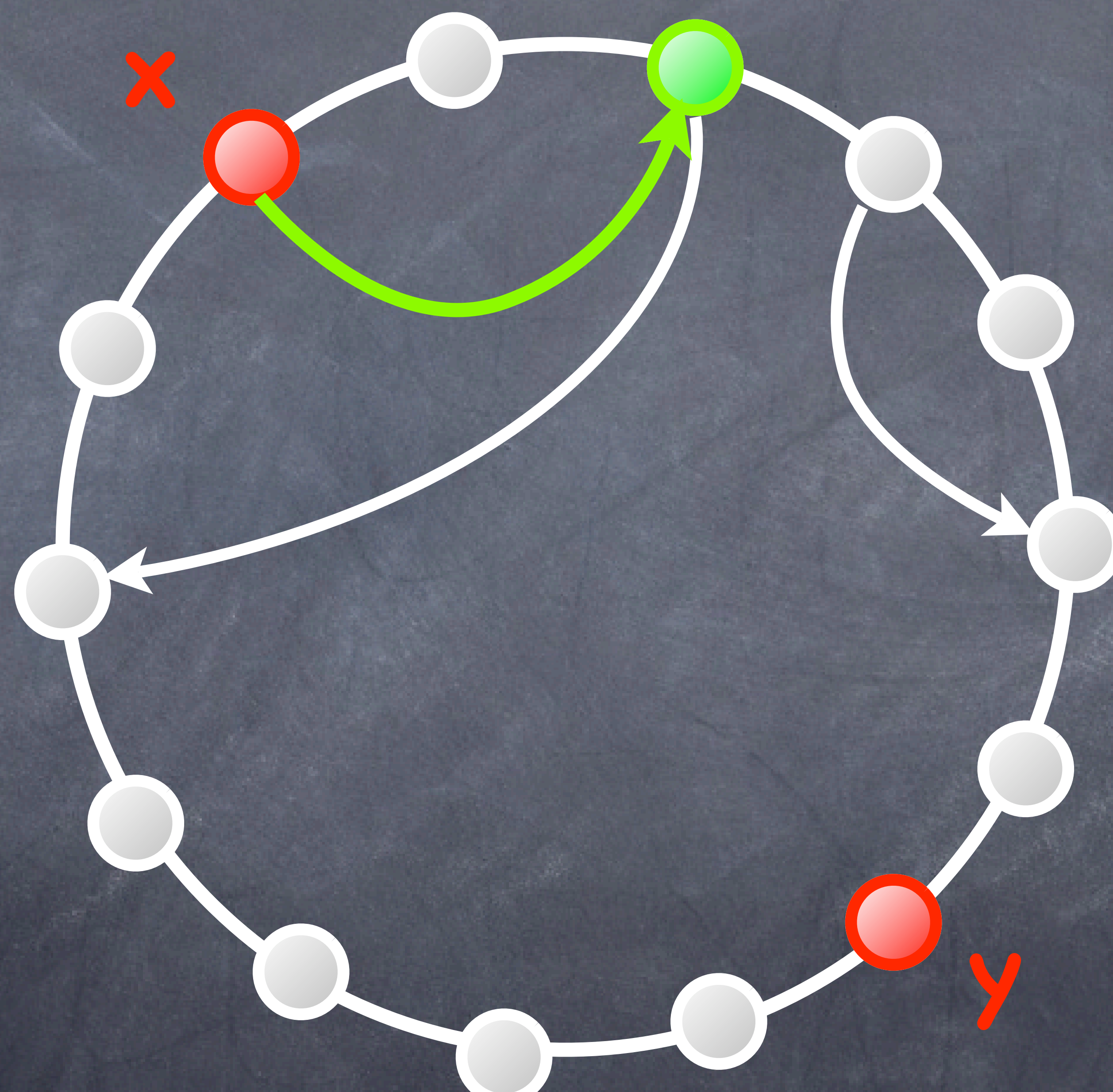
Pick T = 3 u.a.r in {1..7}



Example

d(x,y) = 7

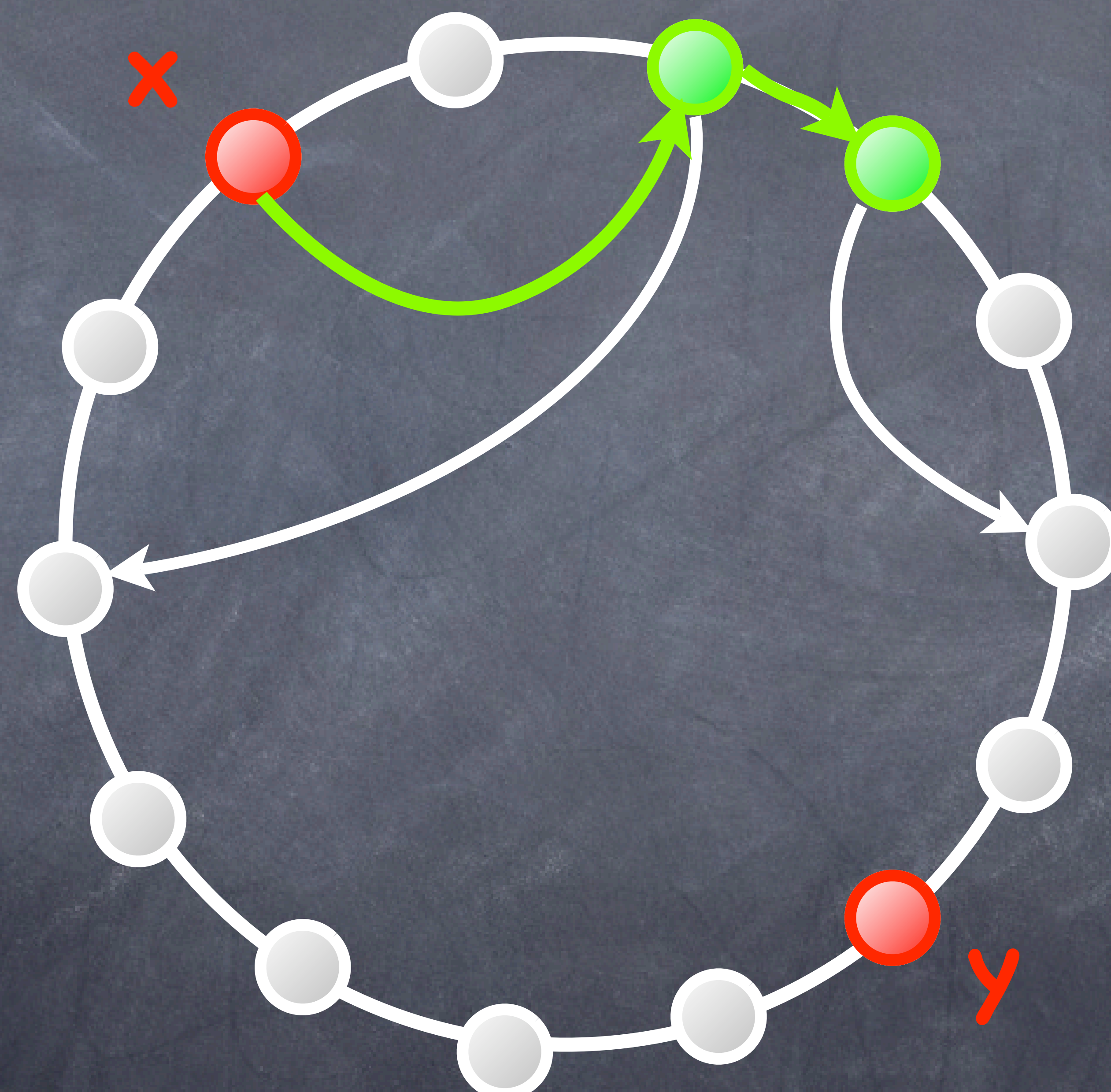
Pick T = 3 u.a.r in {1..7}



Example

d(x,y) = 7

Pick T = 3
u.a.r in {1..7}

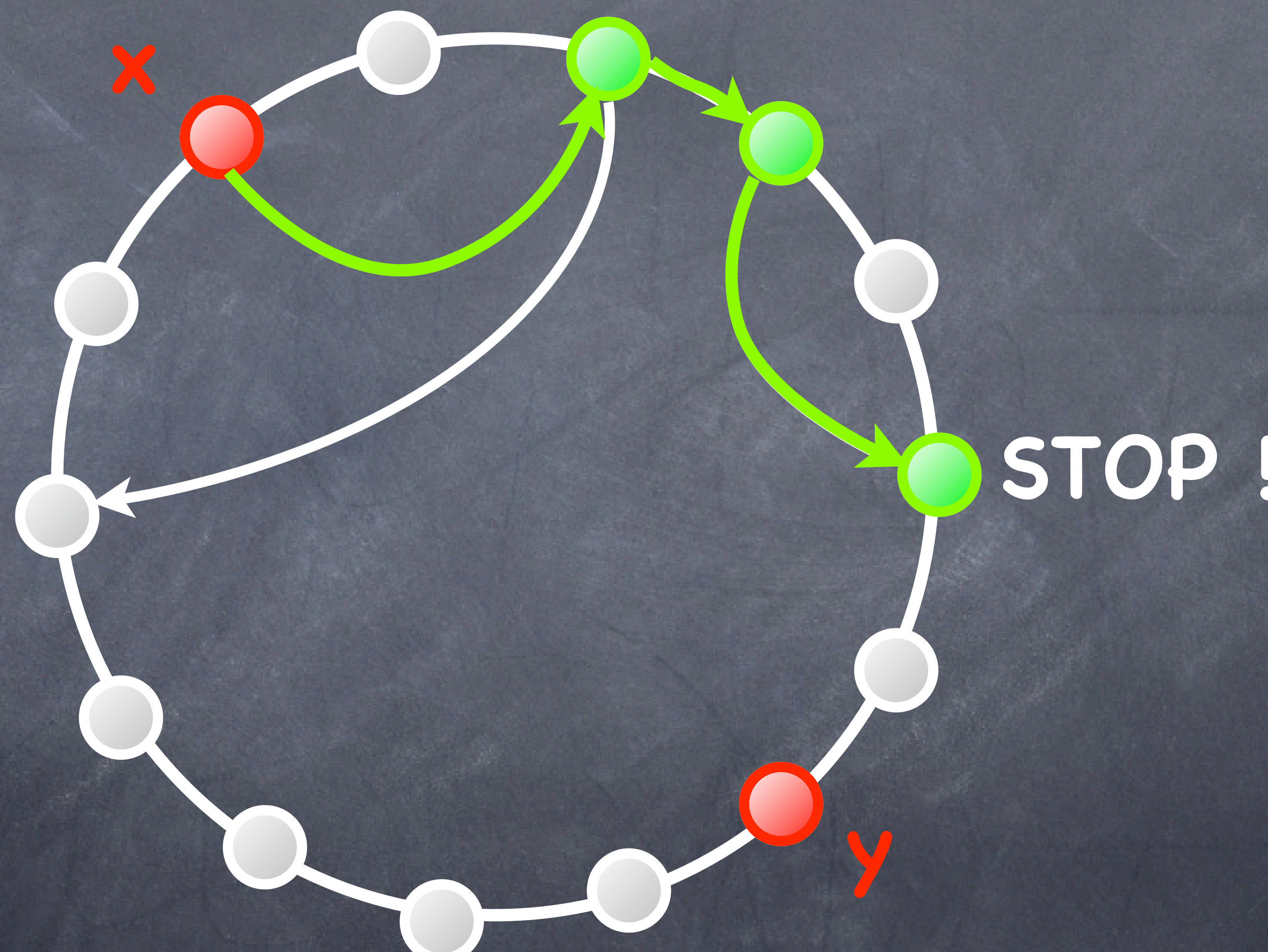


Example

d(x,y) = 7

Pick T = 3

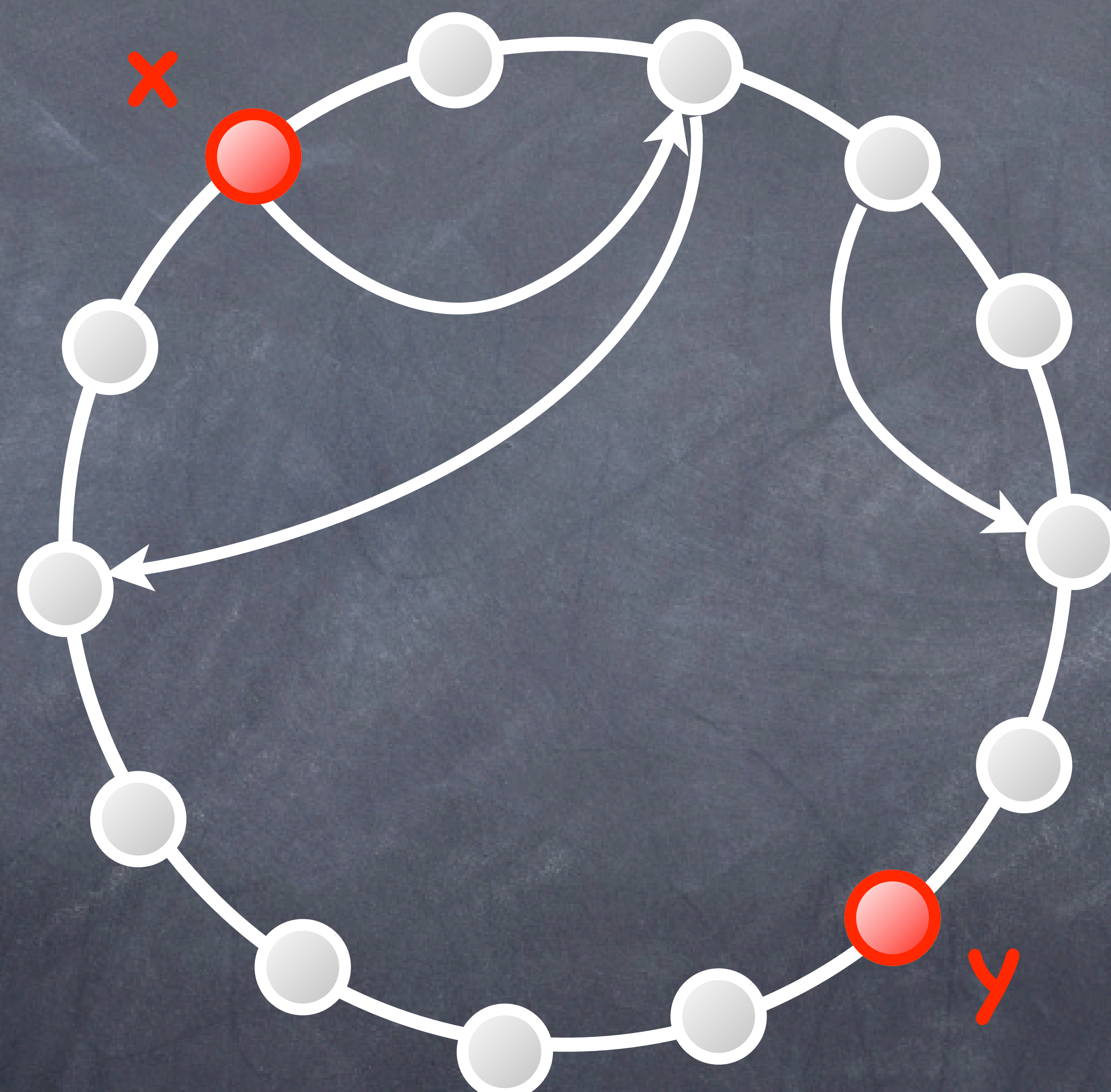
u.a.r in {1..7}



Example

d(x,y) = 7

Pick T = 3 u.a.r in {1..7}

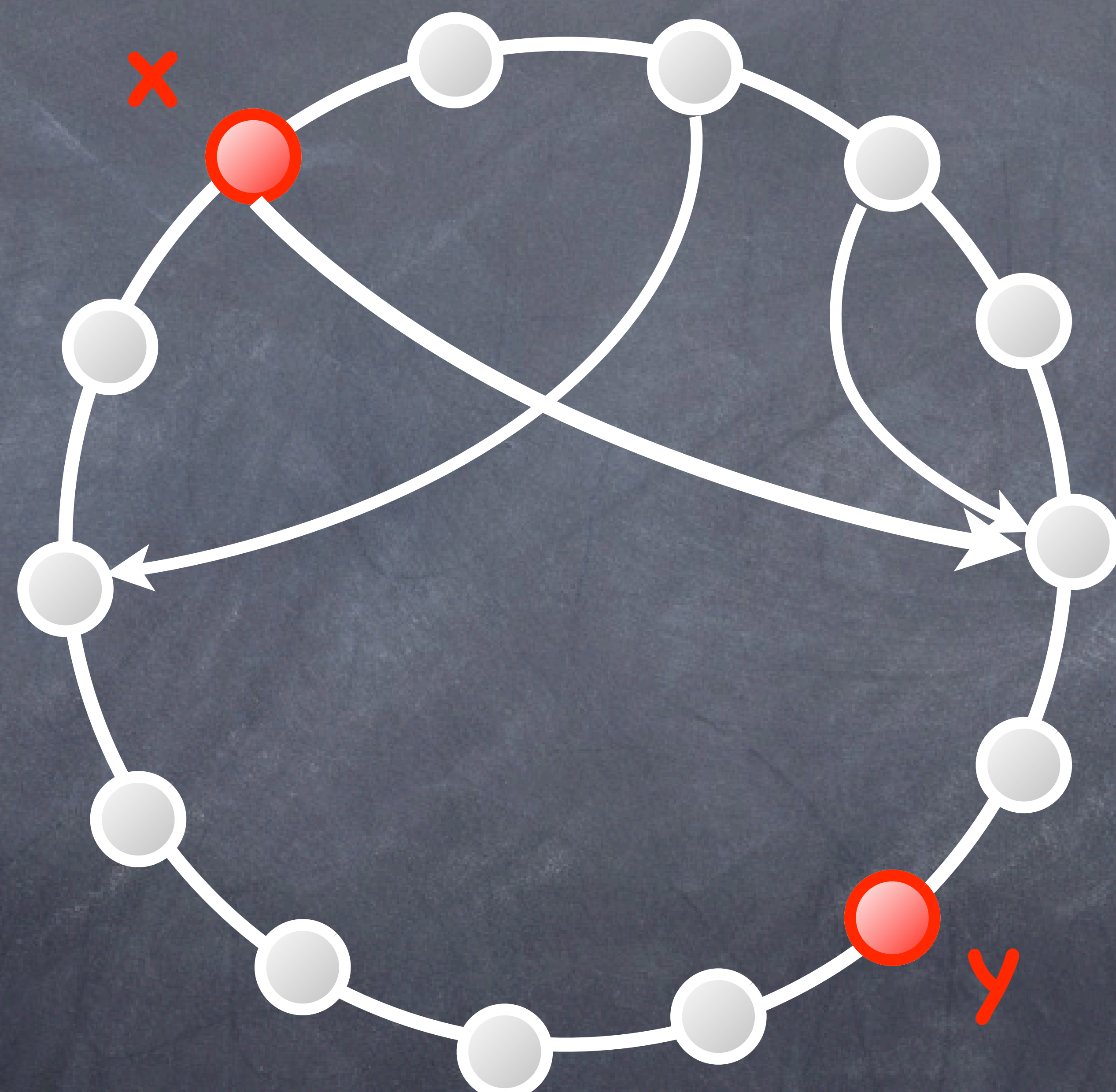


Example

d(x,y) = 7

Pick T = 3 u.a.r in {1..7}

Stop & bookmark !

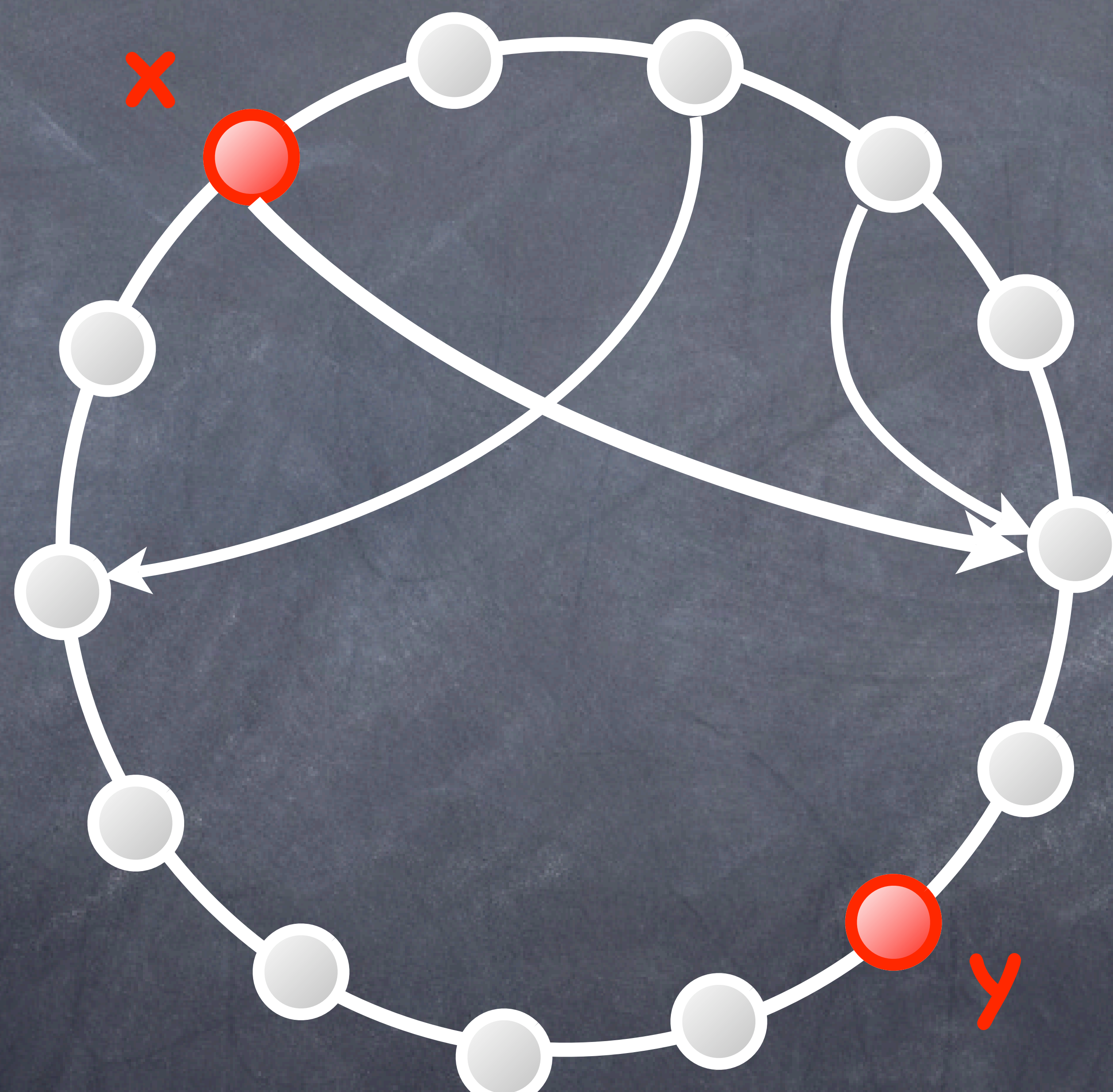


Example

d(x,y) = 7

Pick **T** = 3 u.a.r in {1..7}

Stop & bookmark !



Empirically converges to Kleinberg's law !
BUT NO PROOF YET

Some Goals

- Come up with natural/realistic models & process
- Study relationship "Structure <-> Functionnality"
- Lots of applications in P2P or networks optimization for instance
- Computer scientists may have as important things to say as physicists on the matter
- Example of a question: What is a society anyway? May be its size just depends on the algorithm that running on it (e.g. 1000 users limit of early P2P Gnutella)