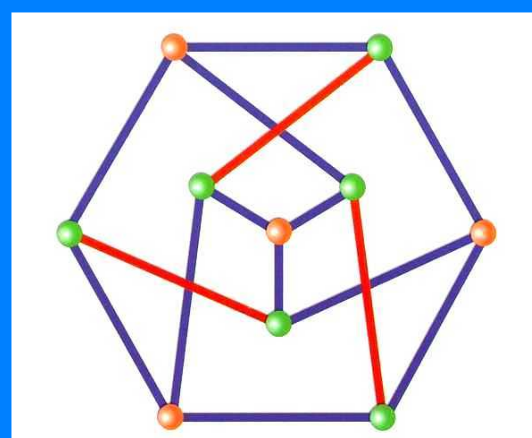
Defending Against H1N1 Virus, Smallpox, and Other Naturally Occurring or Deliberately Caused Diseases: How Can Graph Theory Help?

Fred Roberts, CCICADA



CCICADA
Command, Control, and Interoperability Center for Advanced Data Analysis
A Department of Homeland Security Center of Excellence


x-rated


X


18


Adults Only
X-Rated


X-RATED

Mathematical Models of Disease Spread

Mathematical models of infectious diseases go back to Daniel Bernoulli's mathematical analysis of smallpox in 1760.

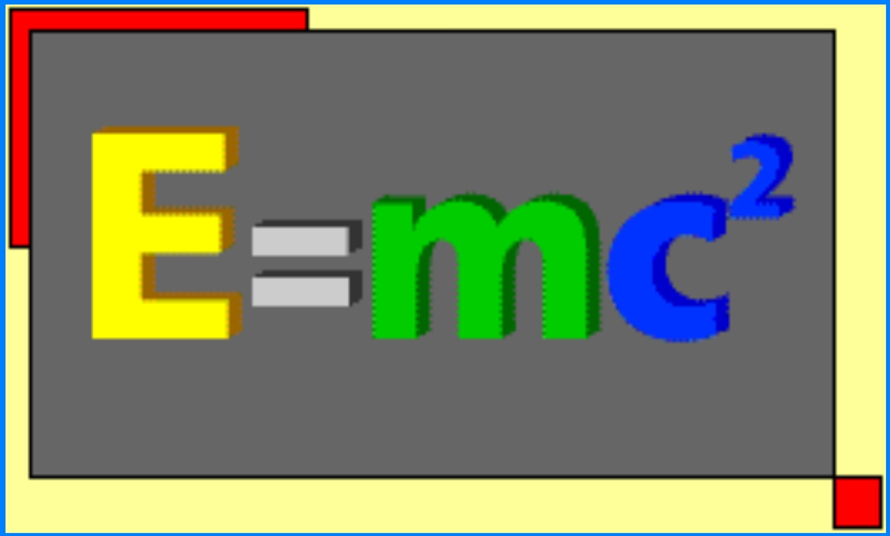
Understanding infectious systems requires being able to reason about highly complex biological systems, with hundreds of demographic and epidemiological variables.

smallpox

Intuition alone is insufficient to fully understand the dynamics of such systems.

Experimentation or field trials are often prohibitively expensive or unethical and do not always lead to fundamental understanding.

Therefore, ***mathematical modeling*** becomes an important experimental and analytical tool.

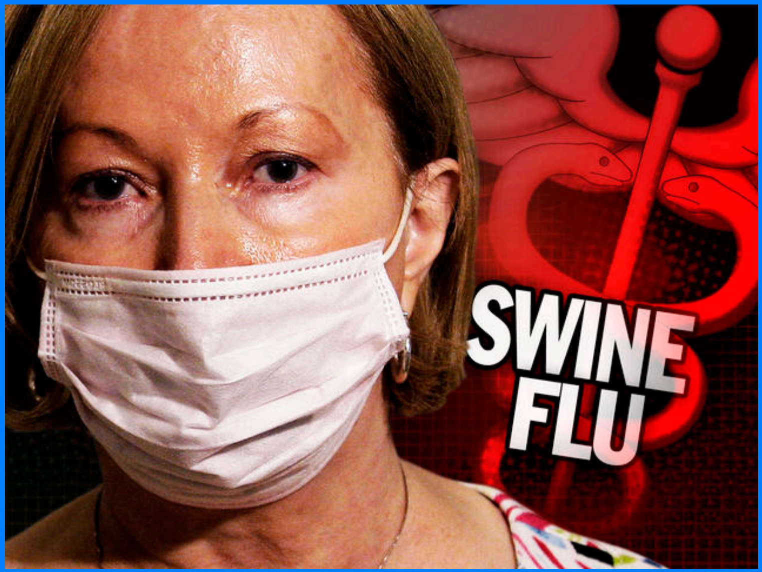


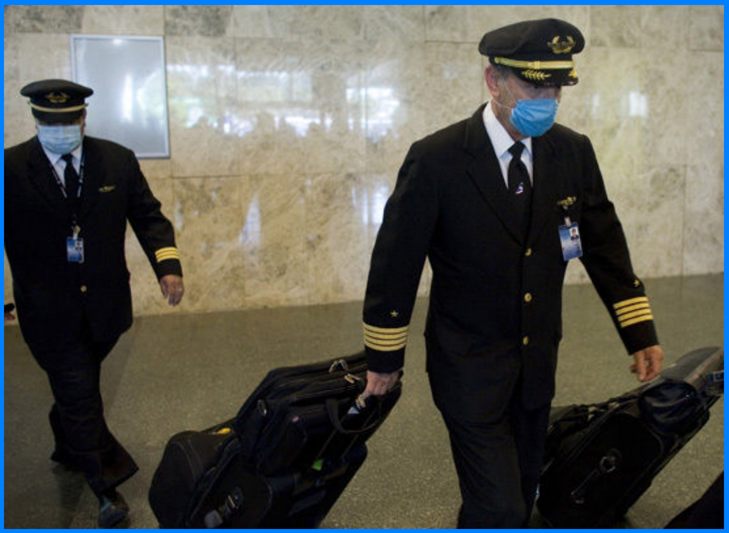
Mathematical models have become important tools in analyzing the spread and control of infectious diseases, especially when combined with powerful, modern computer methods for analyzing and/or simulating the models.

Great concern about the deliberate introduction of diseases by bioterrorists has led to new challenges for mathematical modelers.

anthrax

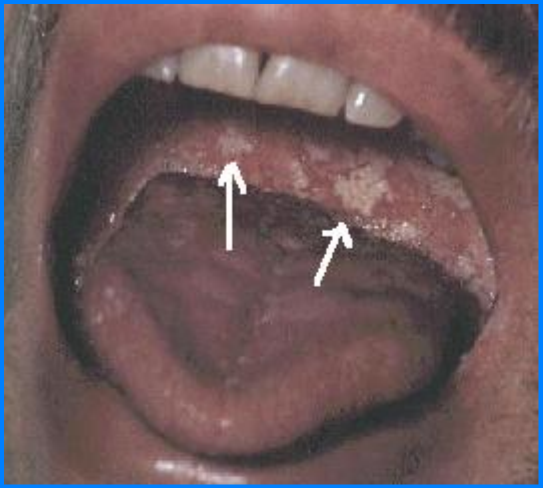
Great concern about possibly devastating new diseases like H1N1 influenza has also led to new challenges for mathematical modelers.





Models of the Spread and Control of Disease through Social Networks

AIDS



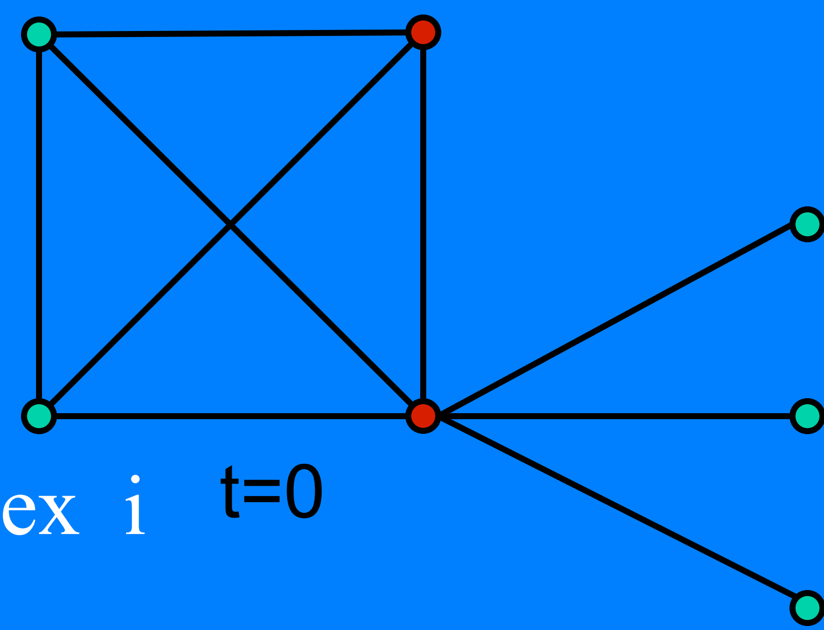
- Diseases are spread through social networks.
- "*Contact tracing*" is an important part of any strategy to combat outbreaks of infectious diseases, whether naturally occurring or resulting from bioterrorist attacks.

The Model: Moving From State to State

Social Network = Graph

Vertices = People

Edges = contact



Let $s_i(t)$ give the state of vertex i at time t.

Simplified Model: Two states:

= susceptible, = infected (SI Model)

Times are discrete: t = 0, 1, 2, …

The Model: Moving From State to State

More complex models: SI, SEI, SEIR, etc.

measles

S = susceptible, E = exposed, I = infected, R = recovered (or removed)

SARS

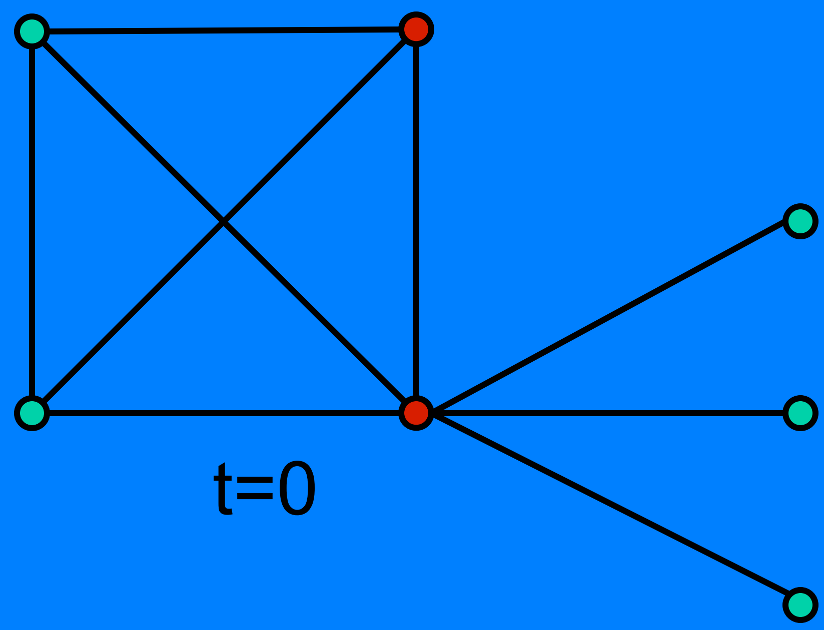
Threshold Processes

Irreversible k-Threshold Process: You change your state from to at time t+1 if at least k of your neighbors have state at time t. You never leave state .

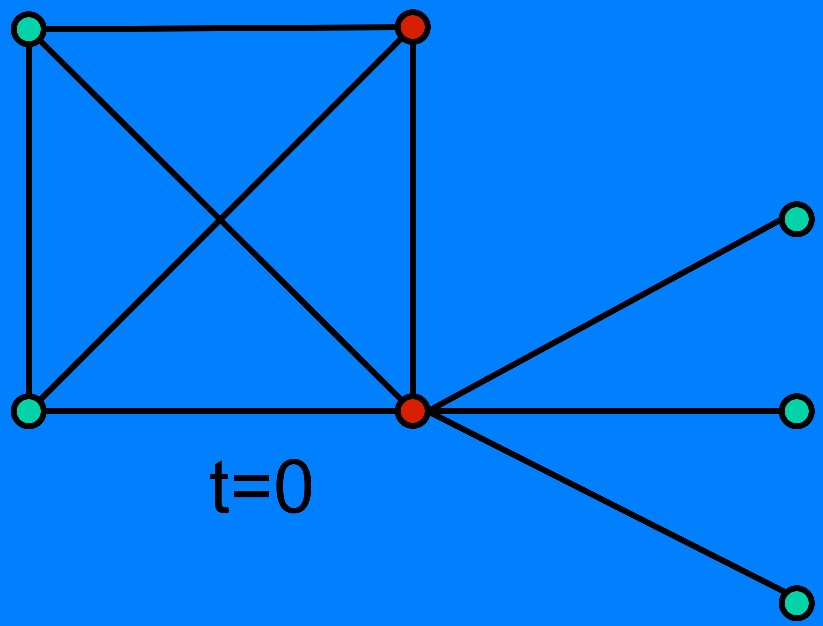
Disease interpretation? Infected if sufficiently many of your neighbors are infected.

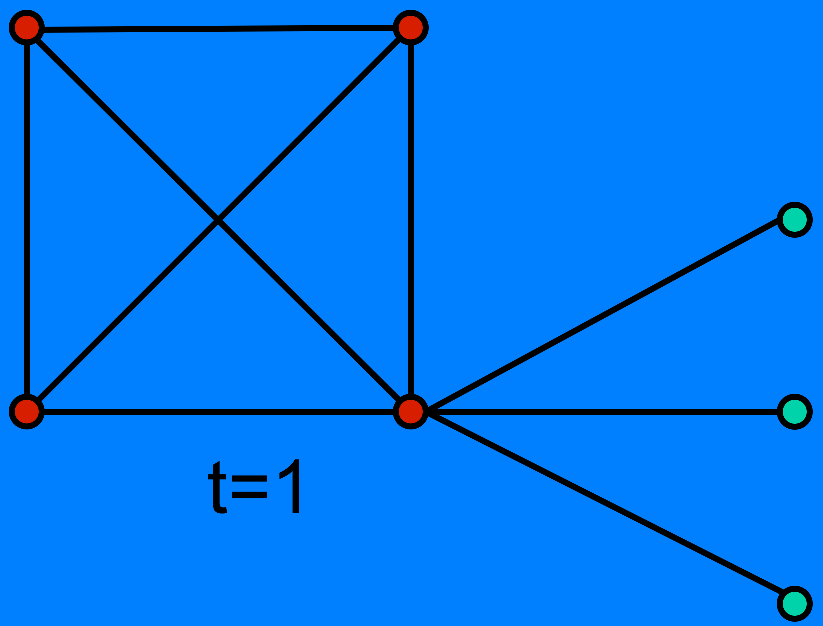
Special Case k = 1: Infected if any of your neighbors is infected.

Irreversible 2-Threshold Process

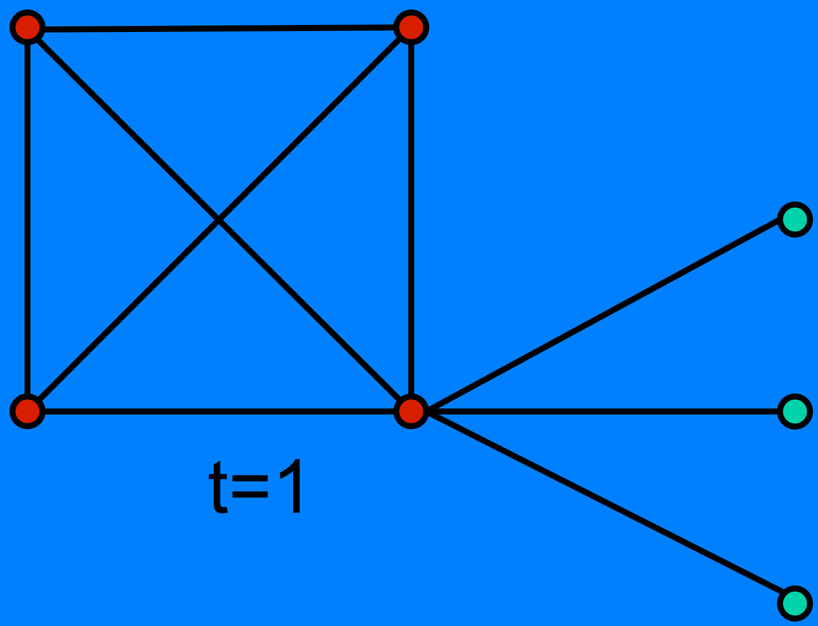


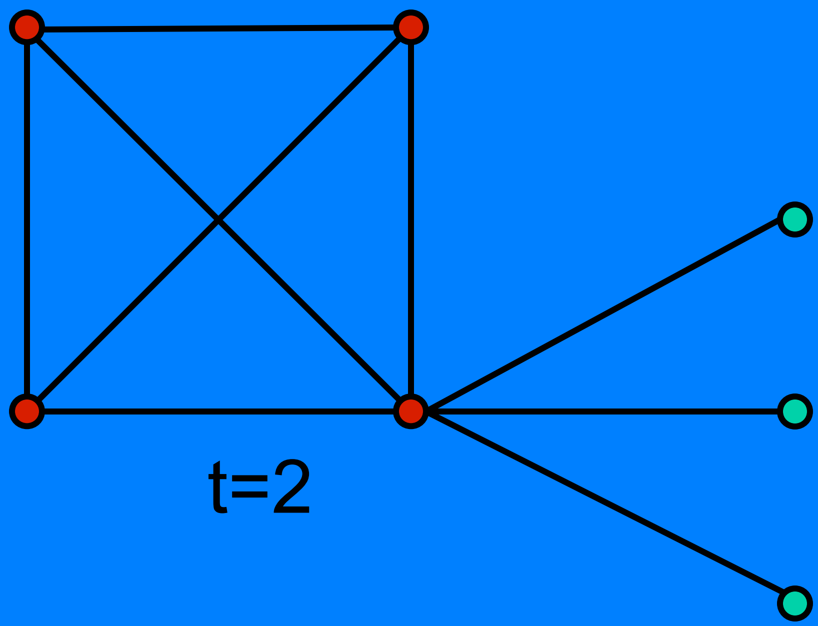
Irreversible 2-Threshold Process



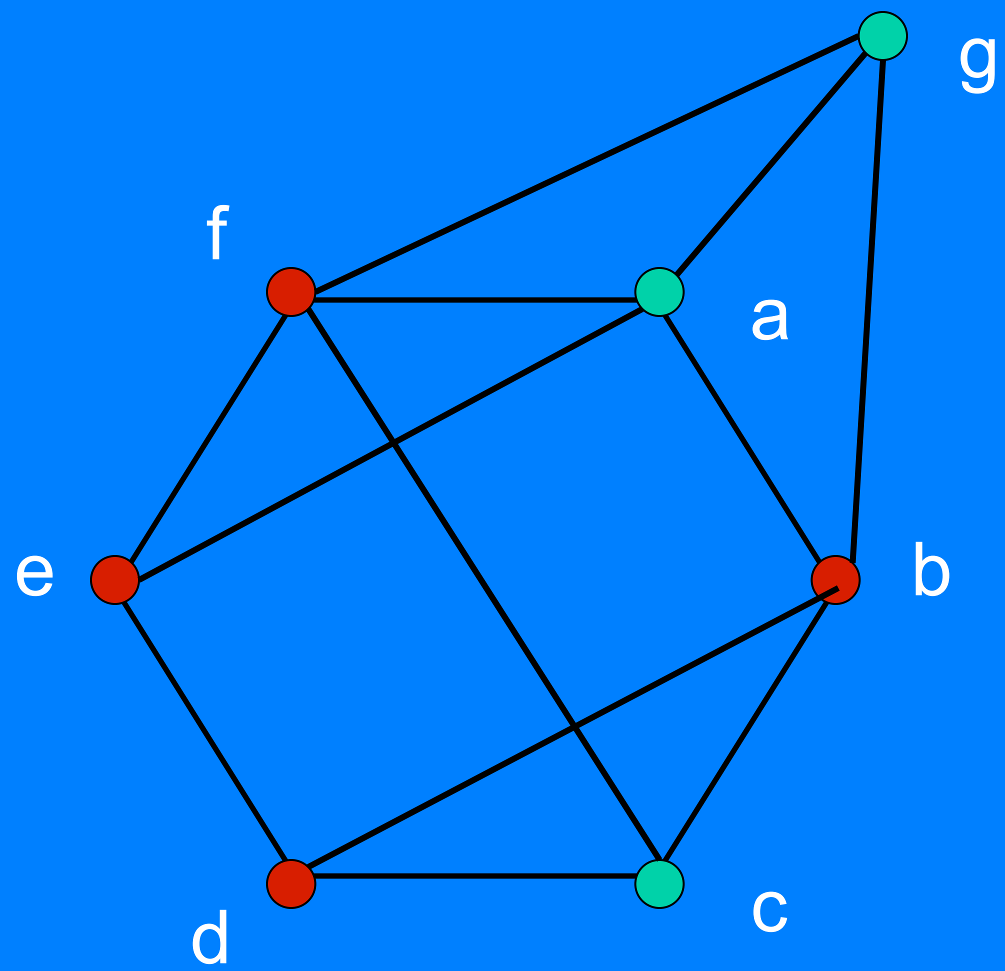


Irreversible 2-Threshold Process



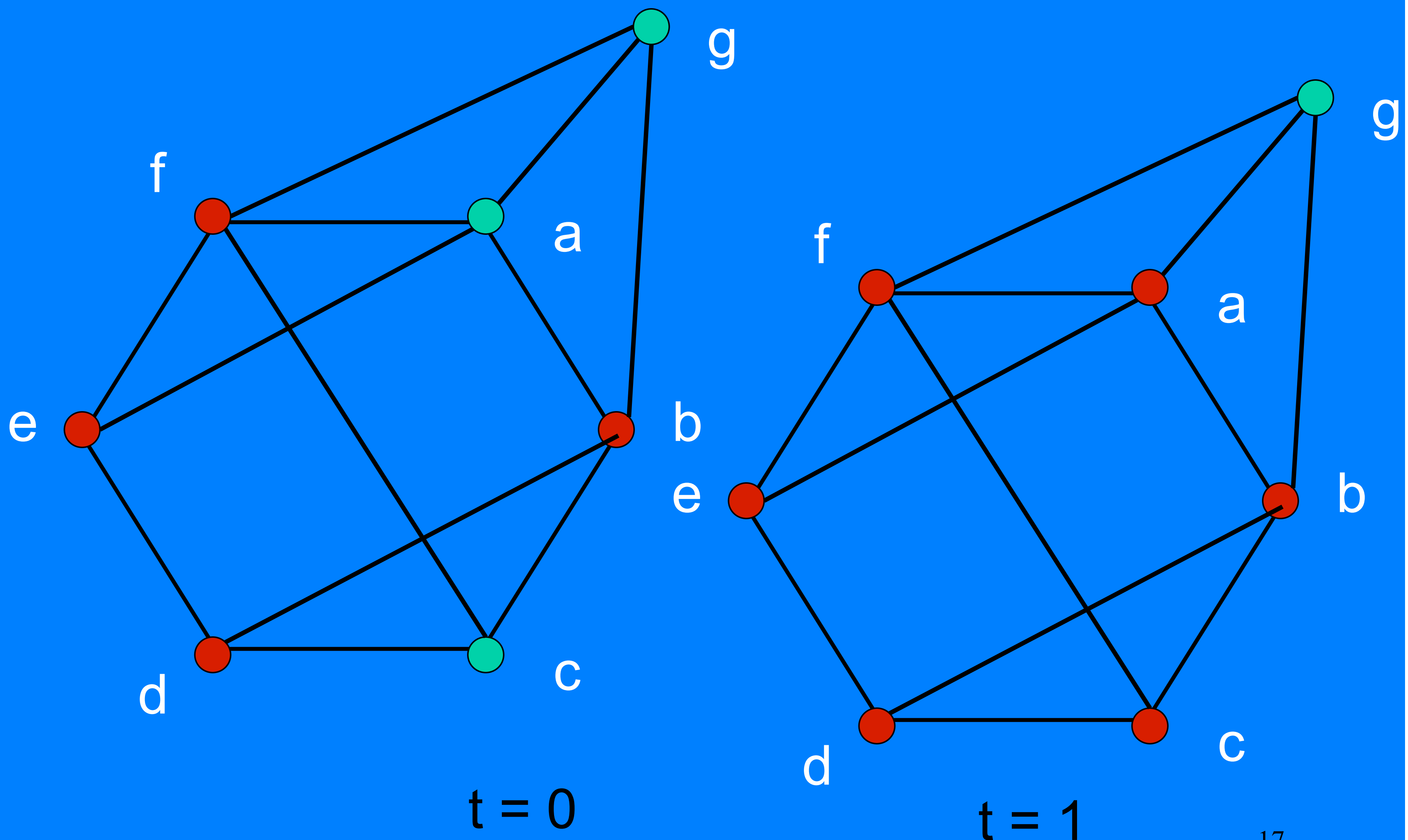


Irreversible 3-Threshold Process

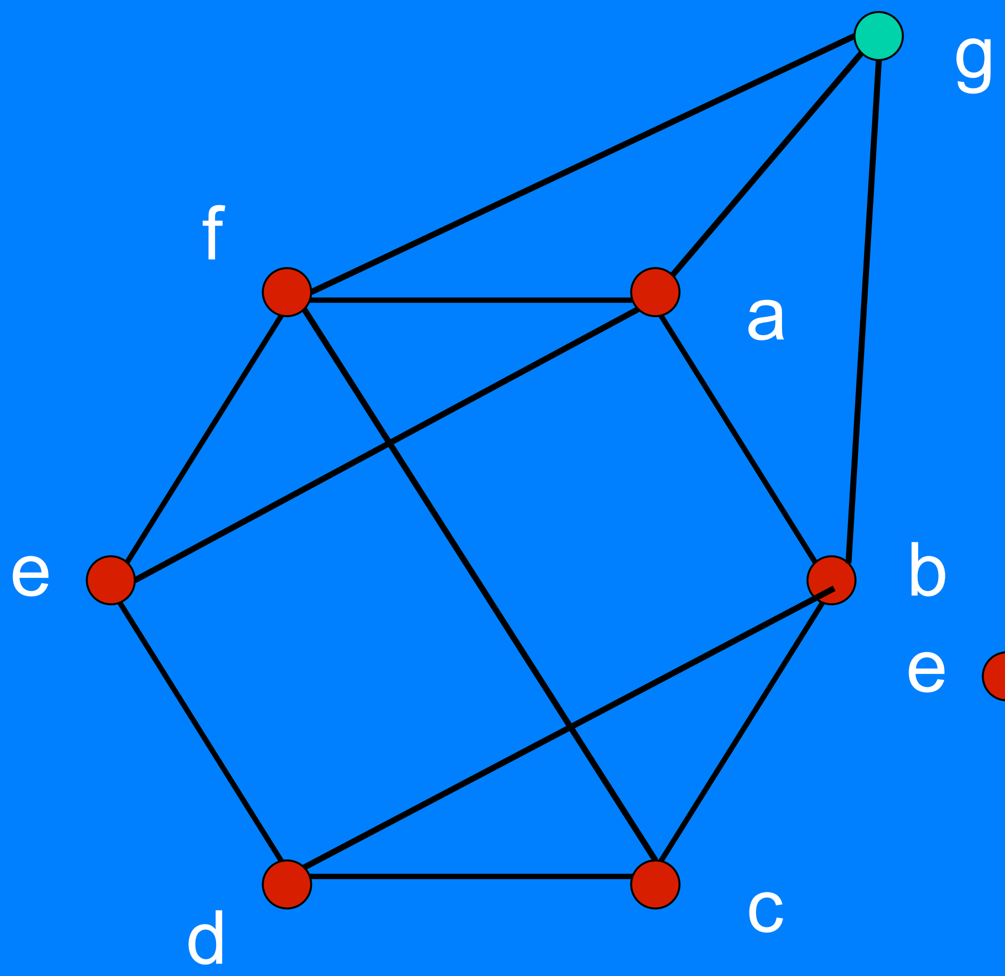


t = 0

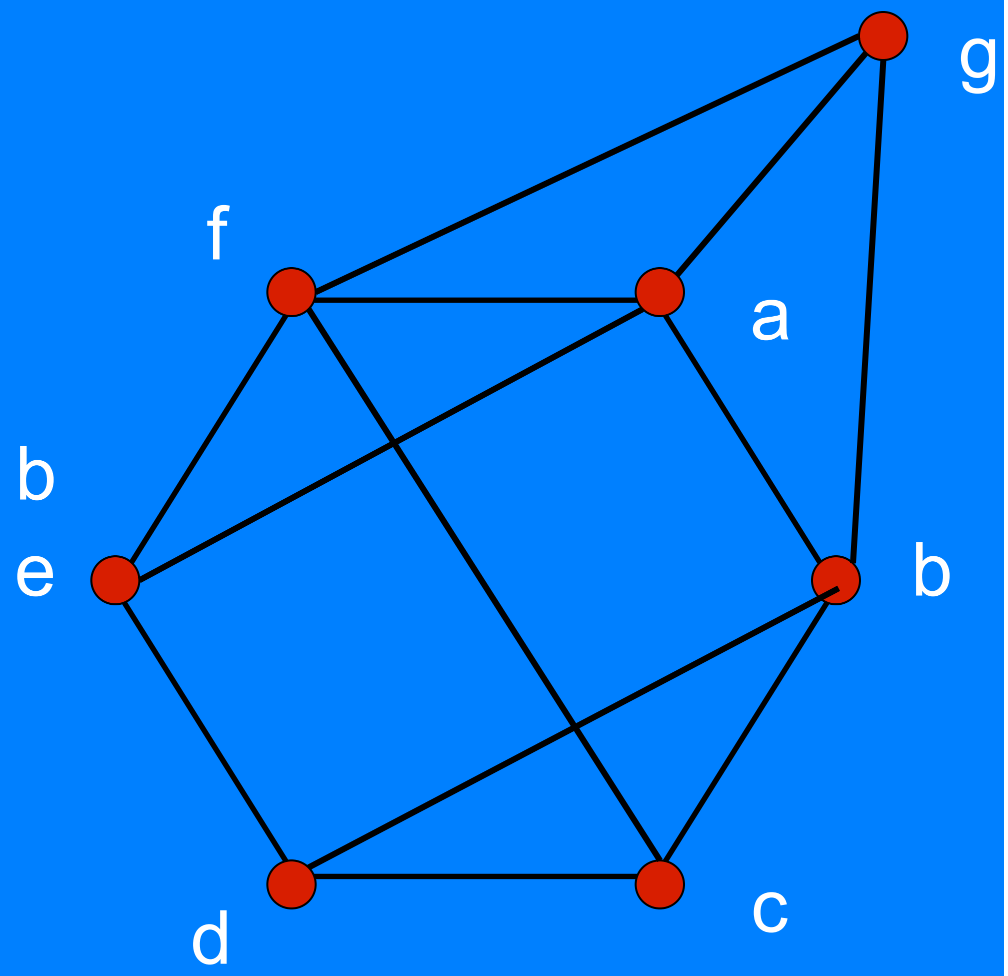
Irreversible 3-Threshold Process



Irreversible 3-Threshold Process



t = 1



t = 2

Complications to Add to Model

•k = 1, but you only get infected with a certain probability.

•You are automatically cured after you are in the infected state for d time periods.

•A public health authority has the ability to "vaccinate" a certain number of vertices, making them immune from infection.

Waiting for smallpox vaccination, NYC, 1947

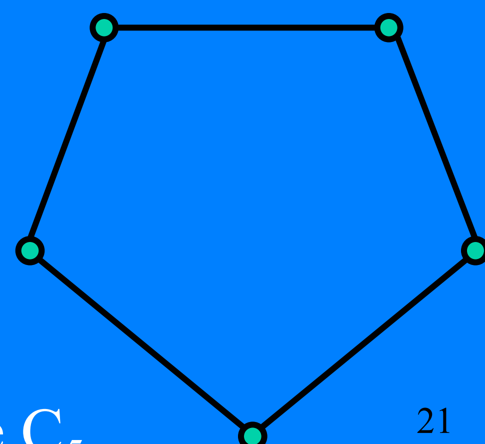
Vaccination Strategies

Mathematical models are very helpful in comparing alternative vaccination strategies. The problem is especially interesting if we think of protecting against deliberate infection by a bioterrorist.

Vaccination Strategies

If you didn't know whom a bioterrorist might infect, what people would you vaccinate to be sure that a disease doesn't spread very much? (Vaccinated vertices stay at state regardless of the state of their neighbors.)

Try odd cycles. Consider an irreversible 2-threshold process. Suppose your adversary has enough supply to infect two individuals.



5-cycle C_5

Vaccination Strategies

One strategy: ***"Mass vaccination"***: Make everyone immune in initial state.

In 5-cycle C_5, mass vaccination means vaccinate 5 vertices. This obviously works.

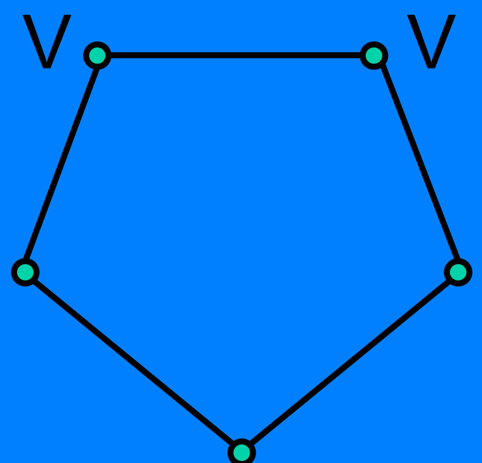
In practice, vaccination is only effective with a certain probability, so results could be different.

Can we do better than mass vaccination?

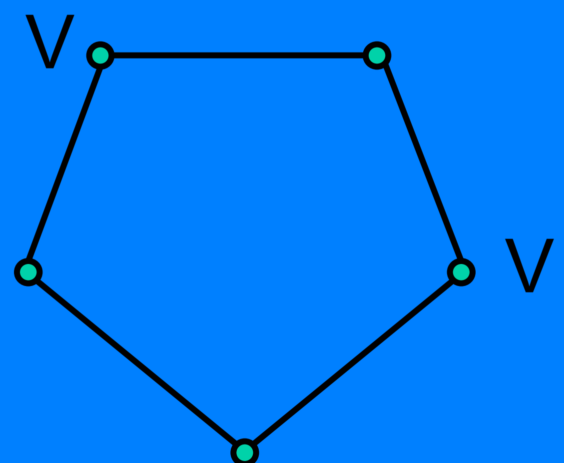
What does better mean? If vaccine has no cost and is unlimited and has no side effects, of course we use mass vaccination.

Vaccination Strategies

What if vaccine is in limited supply? Suppose we only have enough vaccine to vaccinate 2 vertices. Two different vaccination strategies:



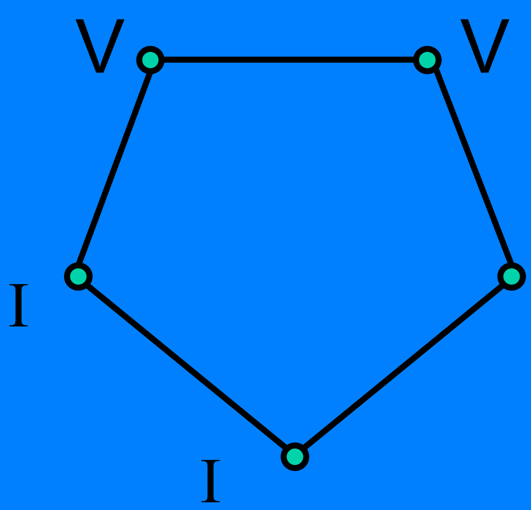
Vaccination Strategy I



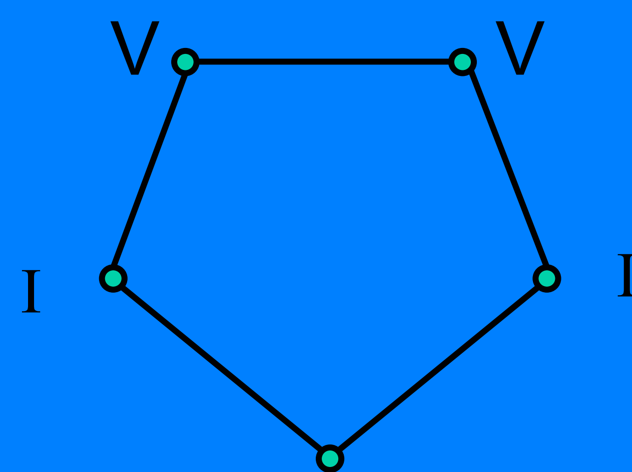
Vaccination Strategy II

Vaccination Strategy I: Worst Case (Adversary Infects Two) Two Strategies for Adversary

This assumes adversary doesn't attack a vaccinated vertex. Problem is interesting if this could happen – or you encourage it to happen.



Adversary Strategy Ia

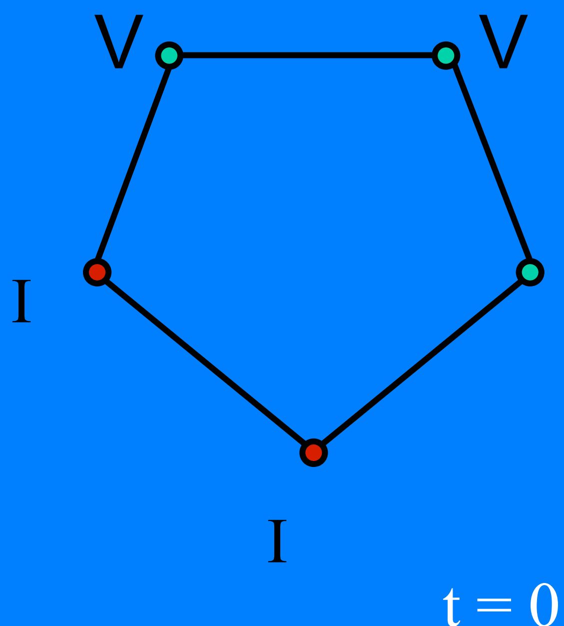


Adversary Strategy Ib

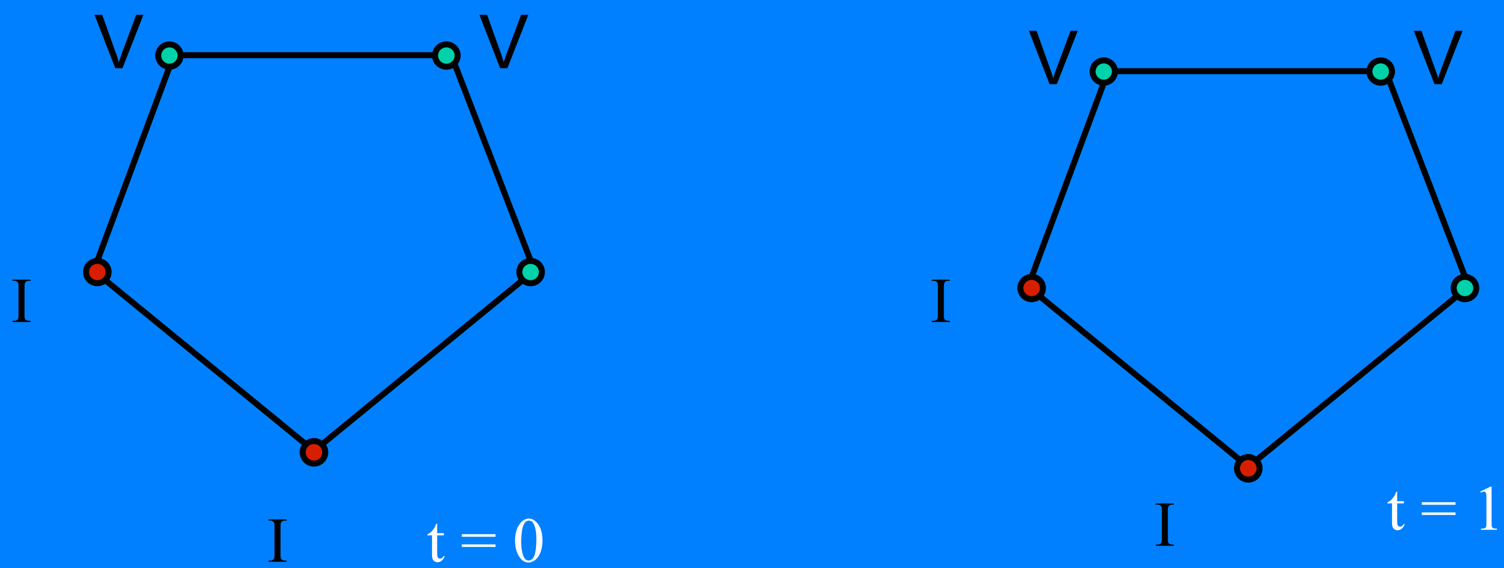
The “alternation” between your choice of a defensive strategy and your adversary’s choice of an offensive strategy suggests we consider the problem from the point of view of game theory.

The Food and Drug Administration is studying the use of game-theoretic models in the defense against bioterrorism.

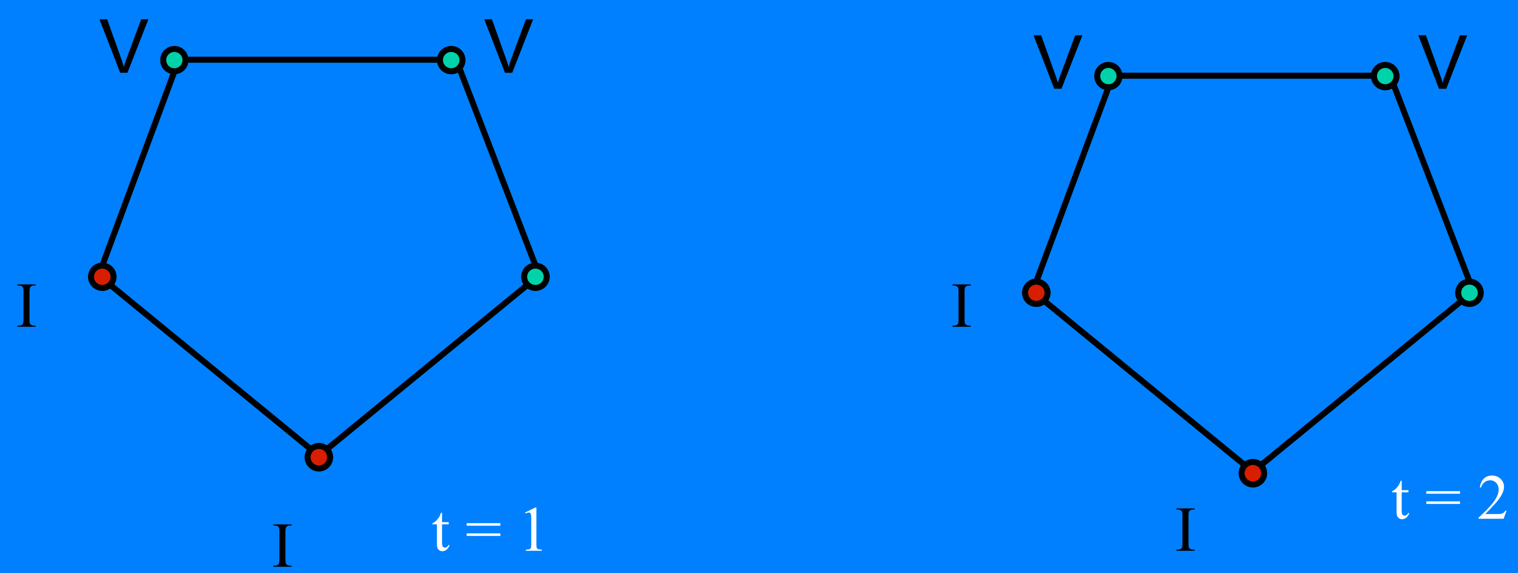
Vaccination Strategy I
Adversary Strategy Ia



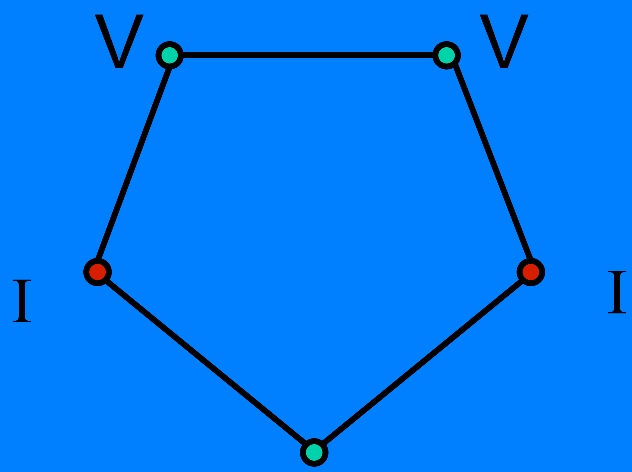
Vaccination Strategy I
Adversary Strategy Ia



Vaccination Strategy I
Adversary Strategy Ia

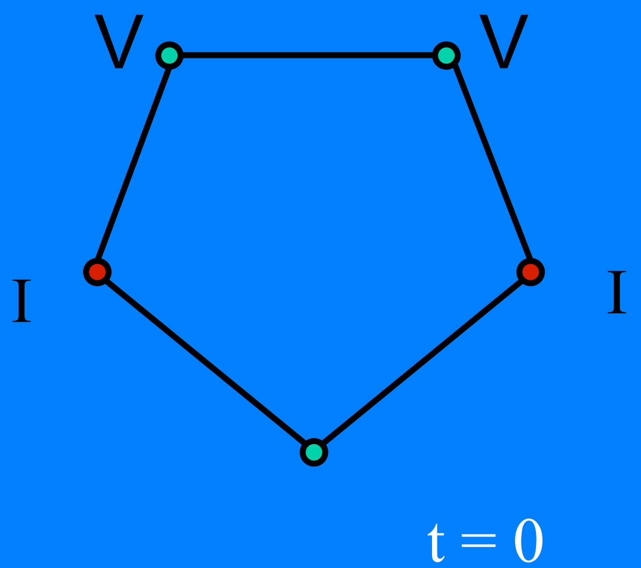


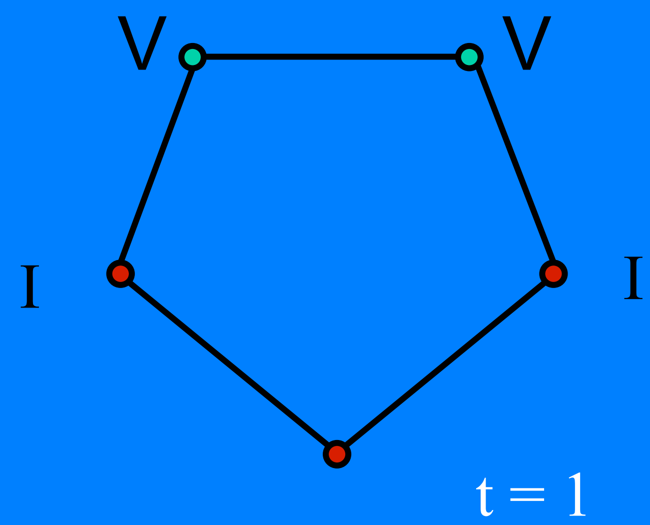
Vaccination Strategy I
Adversary Strategy Ib



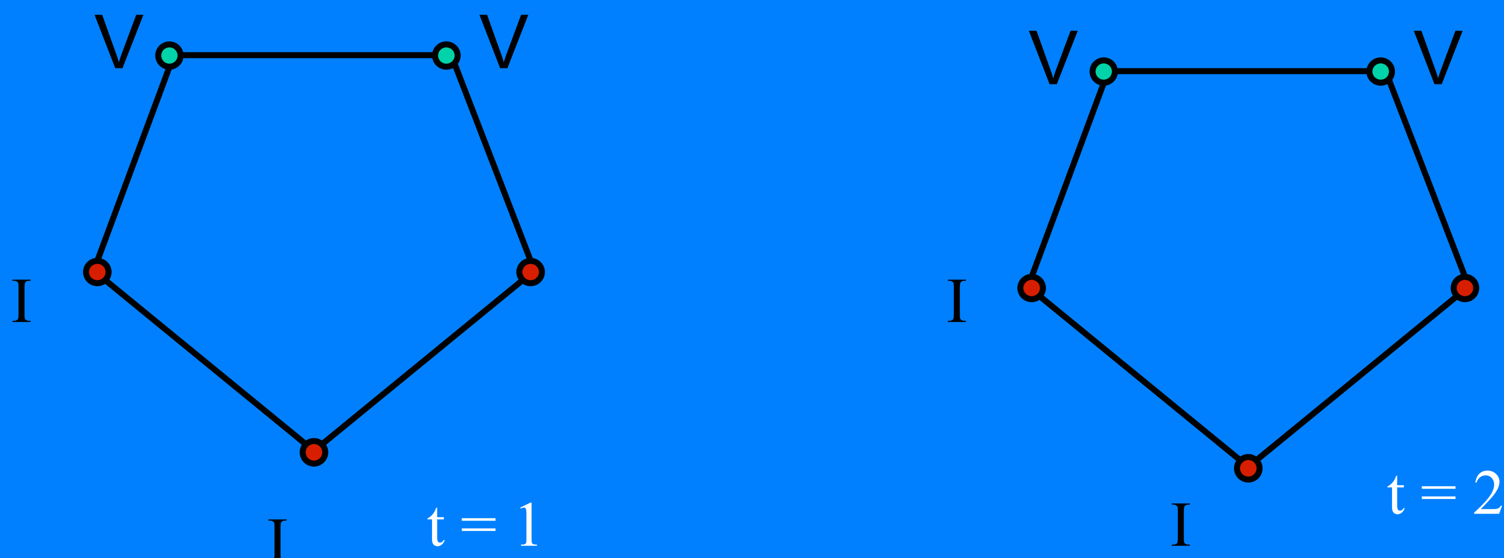
t = 0

Vaccination Strategy I
Adversary Strategy Ib

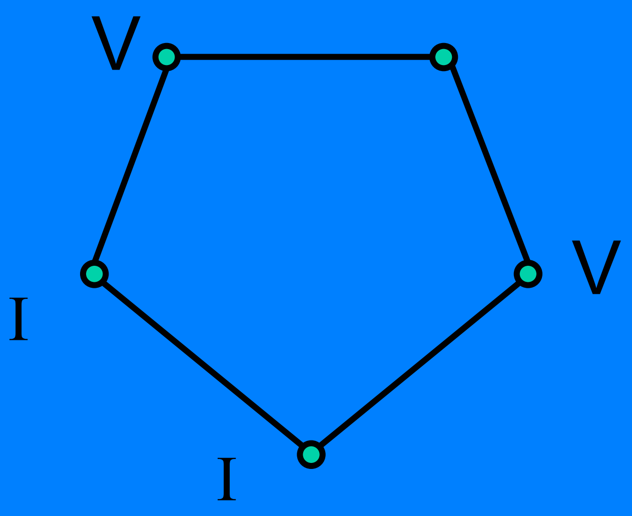




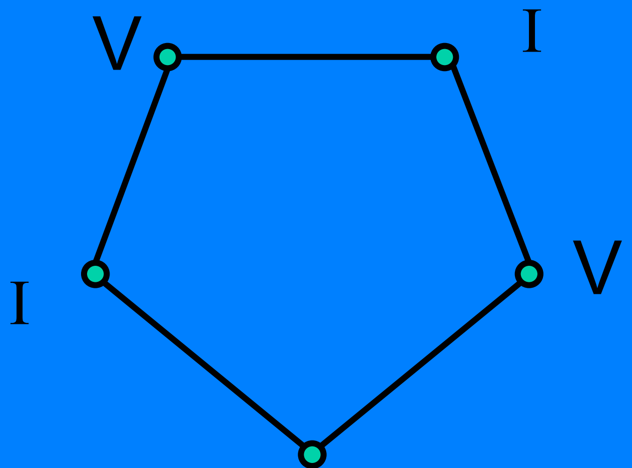
Vaccination Strategy I
Adversary Strategy Ib



Vaccination Strategy II: Worst Case (Adversary Infects Two) Two Strategies for Adversary

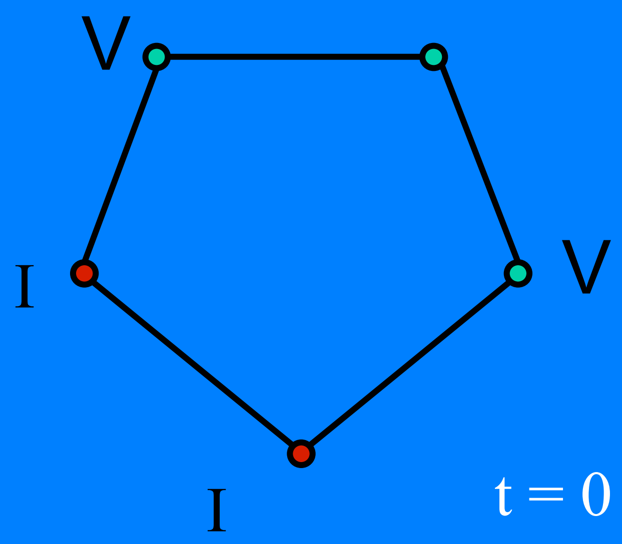


Adversary Strategy IIa

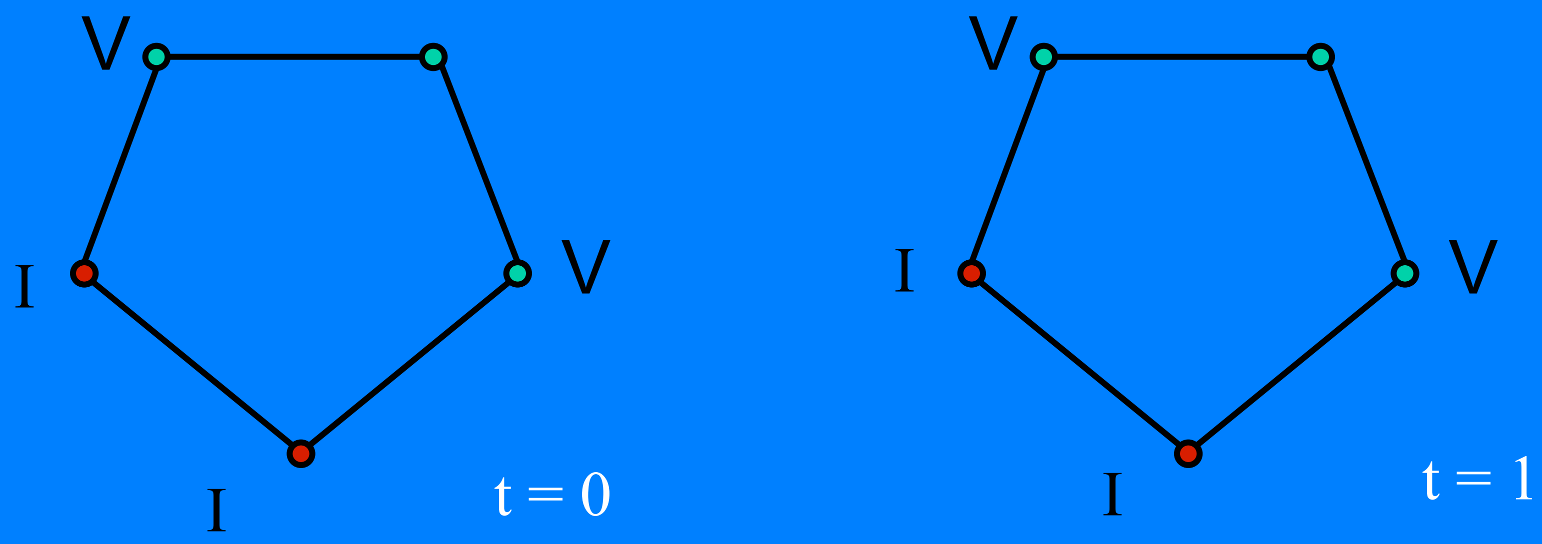


Adversary Strategy IIb

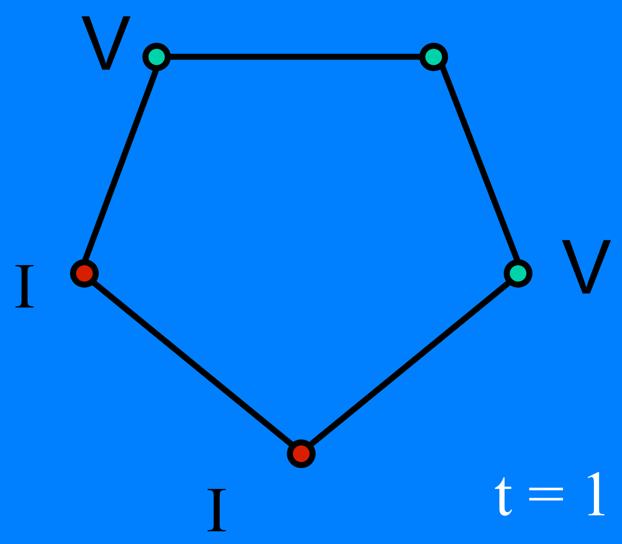
Vaccination Strategy II
Adversary Strategy IIa

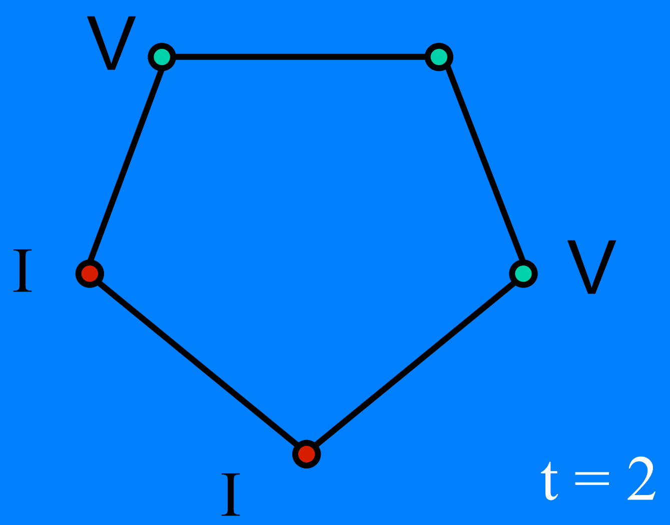


Vaccination Strategy II
Adversary Strategy IIa

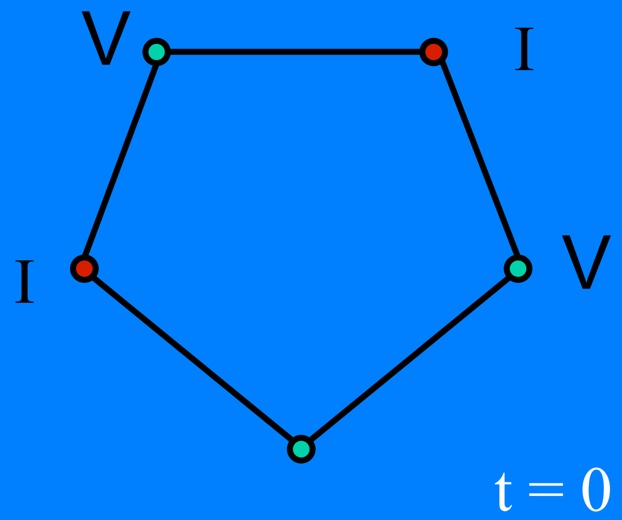


Vaccination Strategy II
Adversary Strategy IIa

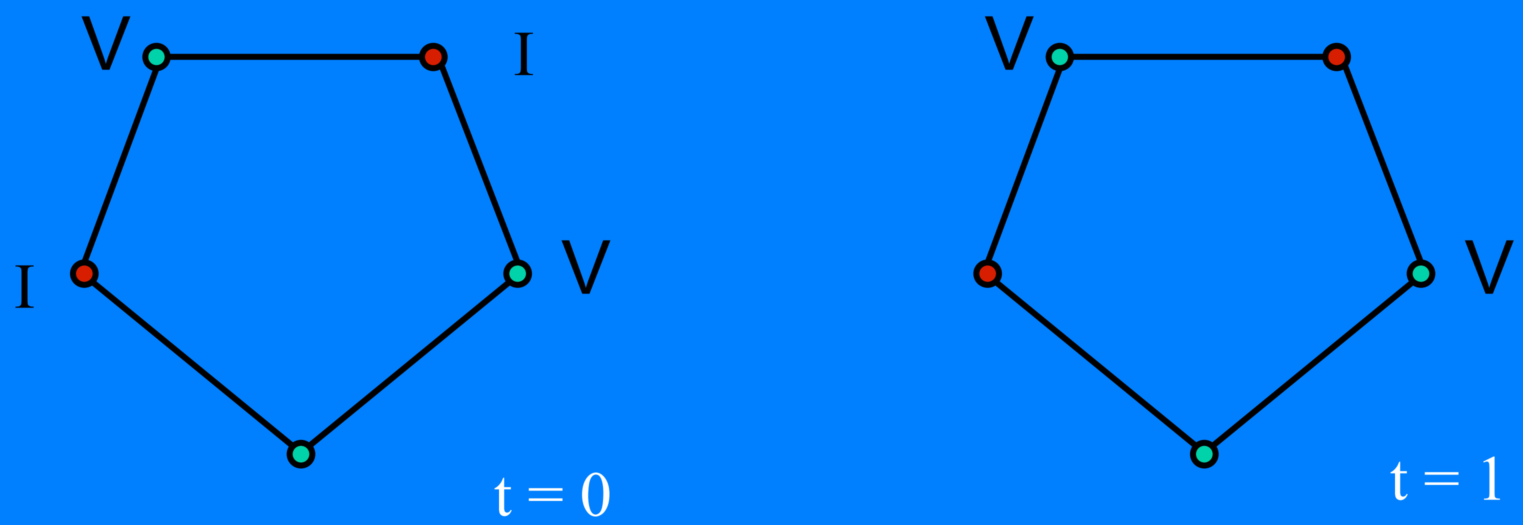




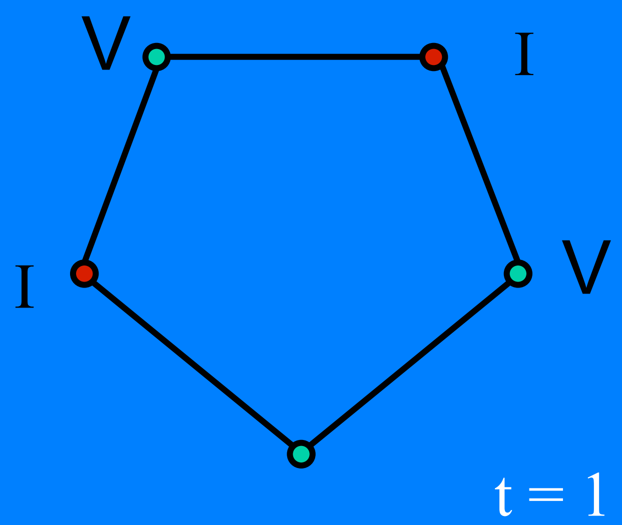
Vaccination Strategy II
Adversary Strategy IIb

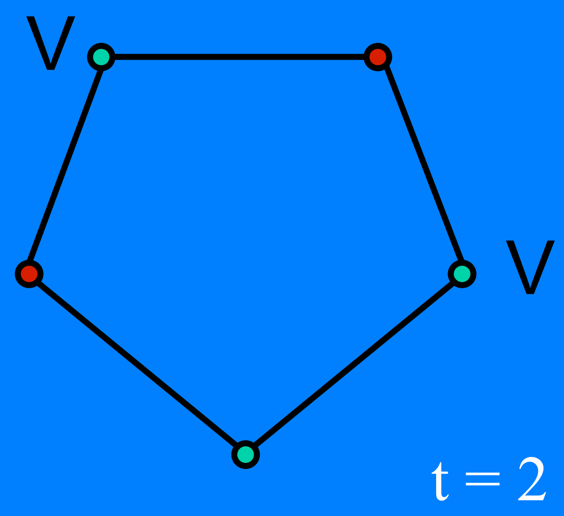


Vaccination Strategy II
Adversary Strategy IIb



Vaccination Strategy II
Adversary Strategy IIb





Conclusions about Strategies I and II

Vaccination Strategy II never leads to more than two infected individuals, while Vaccination Strategy I sometimes leads to three infected individuals (depending upon strategy used by adversary).

Thus, Vaccination Strategy II is better.

More on vaccination strategies later.

The Saturation Problem

Attacker's Problem: Given a graph, what subsets S of the vertices should we plant a disease with so that ultimately the maximum number of people will get it?

Economic interpretation: What set of people do we place a new product with to guarantee "saturation" of the product in the population?

Defender's Problem: Given a graph, what subsets S of the vertices should we vaccinate to guarantee that as few people as possible will be infected?

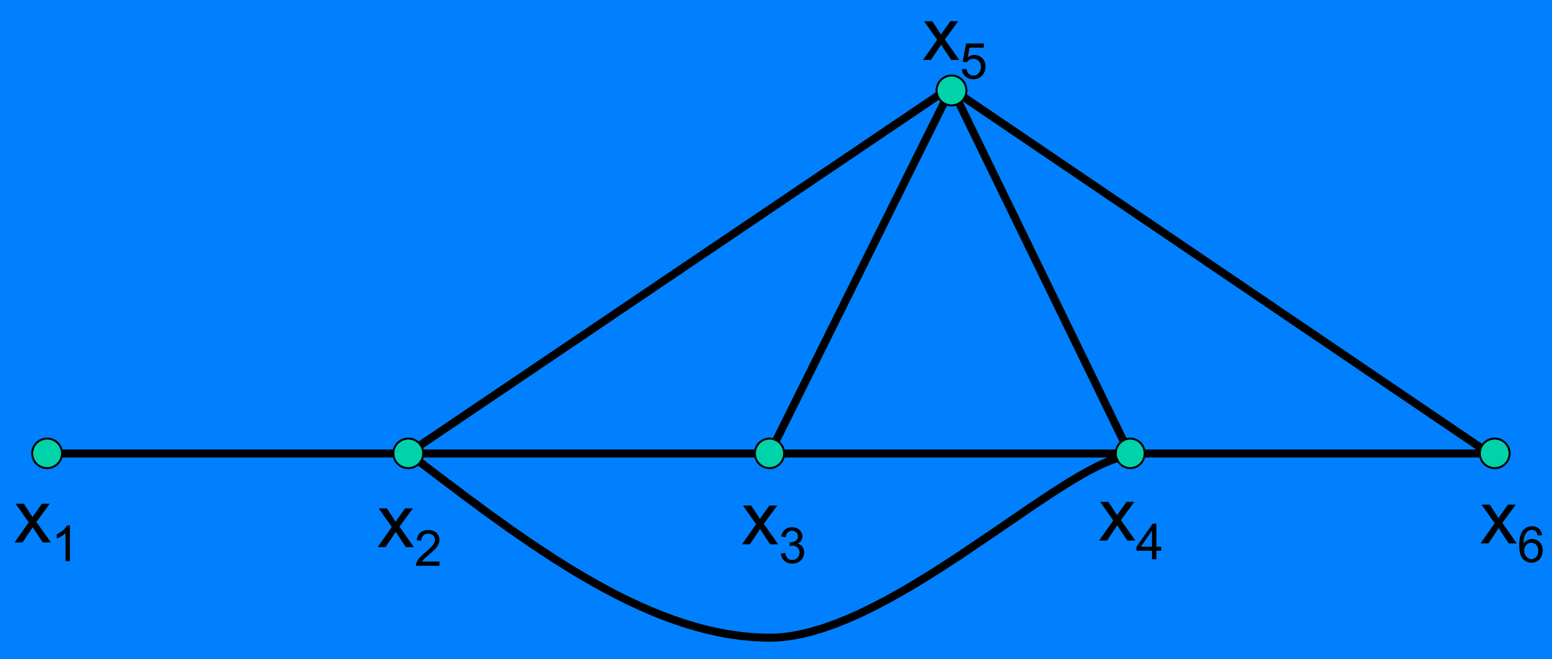
k-Conversion Sets

Attacker's Problem: Can we guarantee that ultimately everyone is infected?

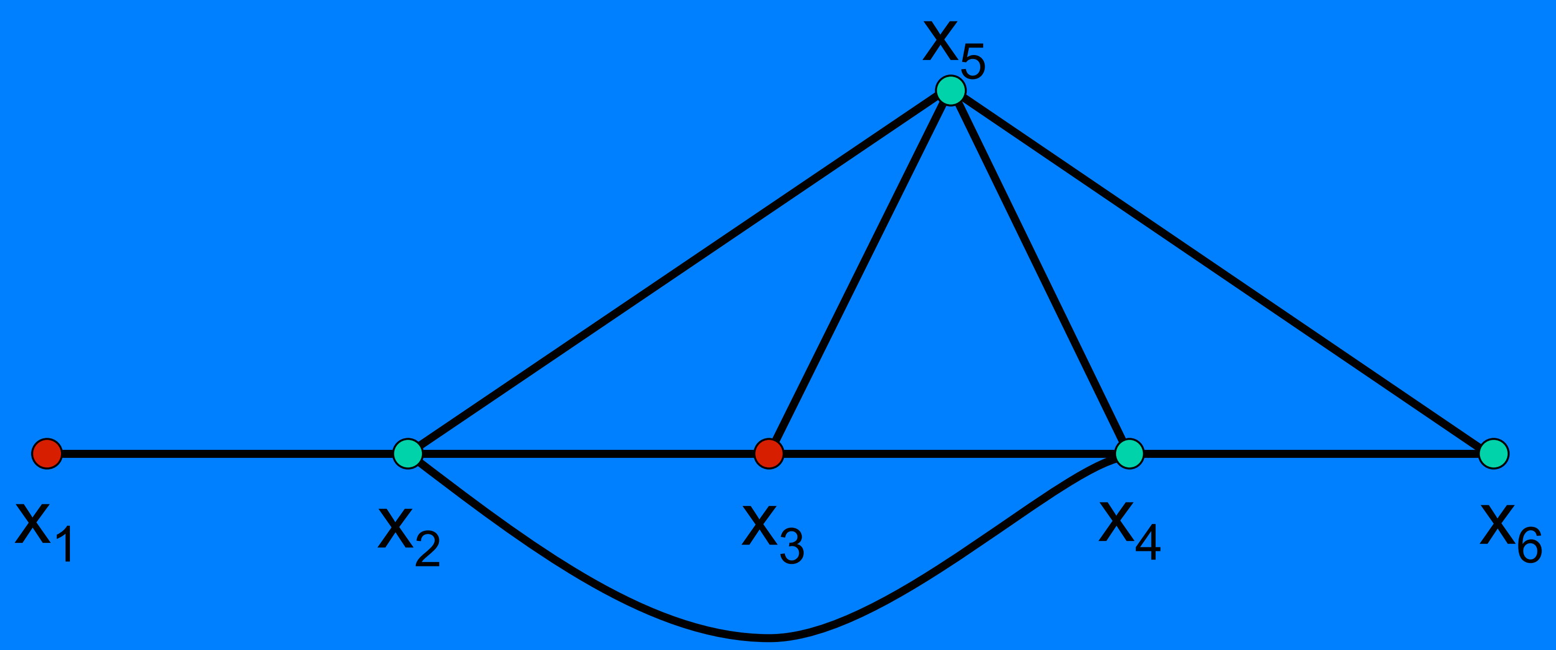
Irreversible k-Conversion Set: Subset S of the vertices that can force an irreversible k-threshold process to the situation where every state $s_i(t) =$ ●

Comment: If we can change back from ● to ○ at least after awhile, we can also consider the Defender's Problem: Can we guarantee that ultimately no one is infected, i.e., all $s_i(t) =$ ○?

What is an irreversible 2-conversion set for the following graph?

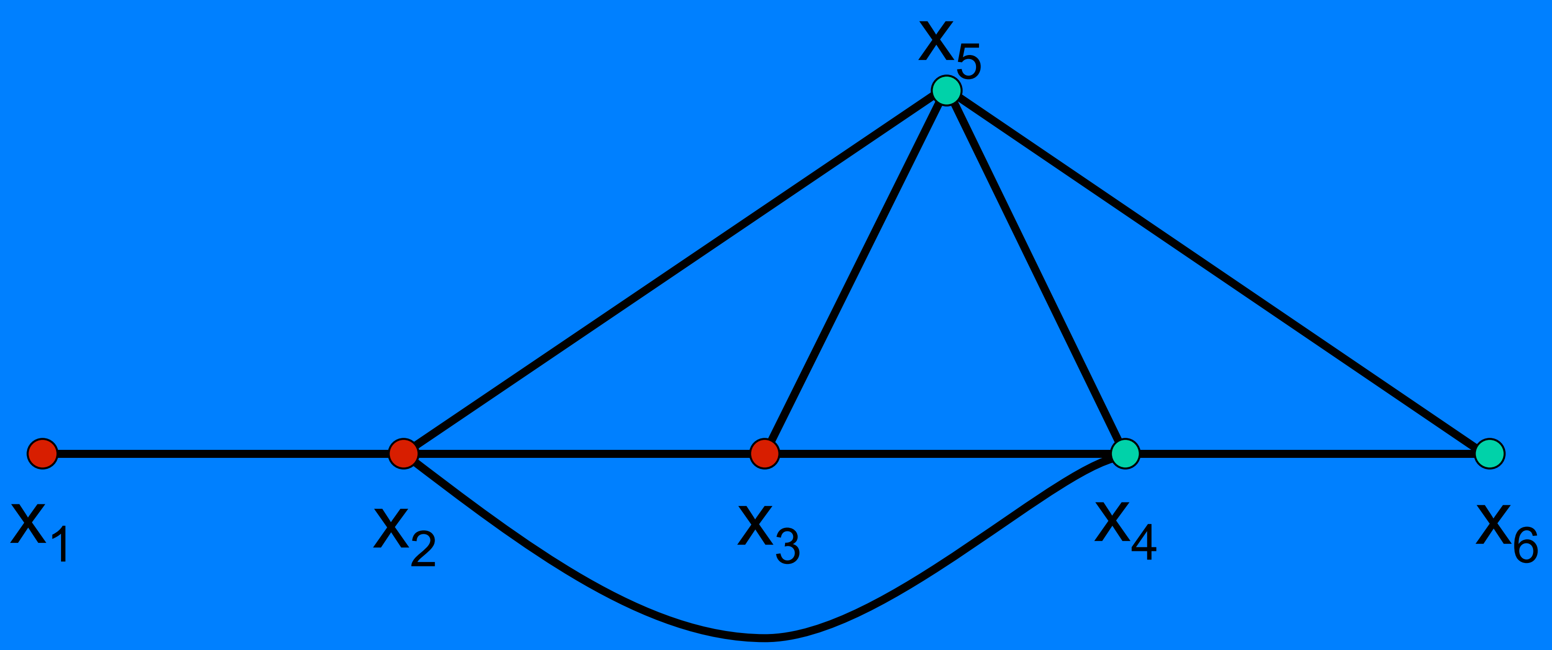


x_1, x_3 is an irreversible 2-conversion set.



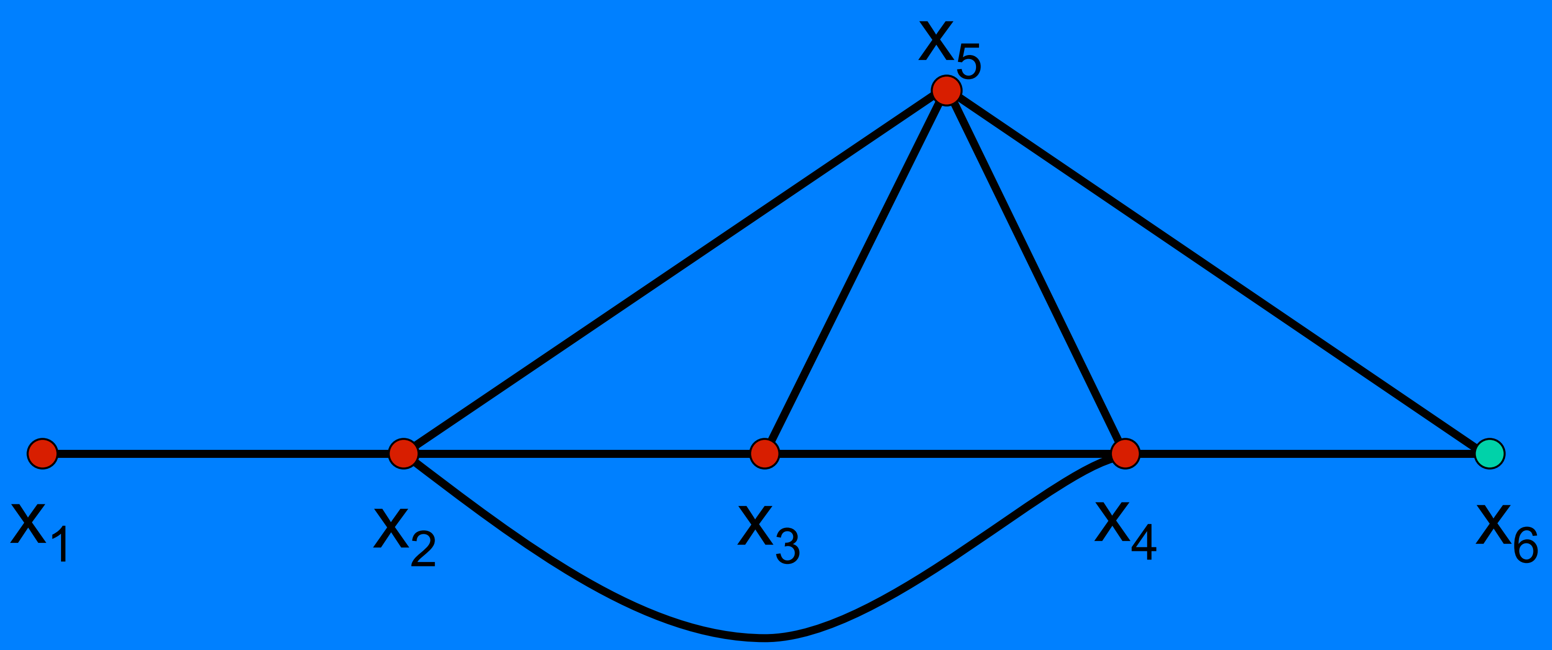
t = 0

x_1, x_3 is an irreversible 2-conversion set.



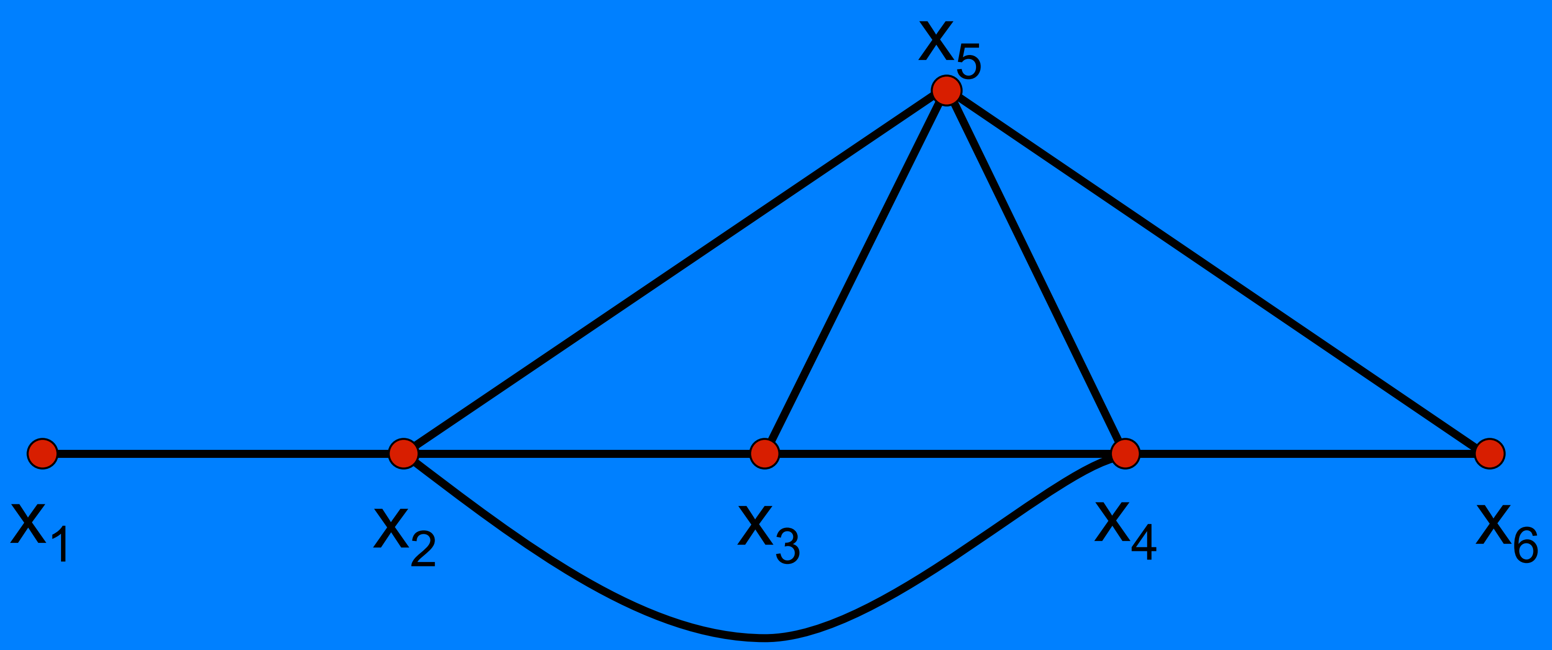
t = 1

x_1, x_3 is an irreversible 2-conversion set.



t = 2

x_1, x_3 is an irreversible 2-conversion set.



t = 3

Irreversible k-Conversion Sets in Regular Graphs

G is ***r-regular*** if every vertex has degree r.
Set of vertices is ***independent*** if there are no edges.

Theorem (Dreyer 2000): Let G = (V,E) be a connected r-regular graph and D be a set of vertices. Then D is an irreversible r-conversion set iff V-D is an independent set.

Note: same r

k-Conversion Sets in Regular Graphs

Corollary (Dreyer 2000):

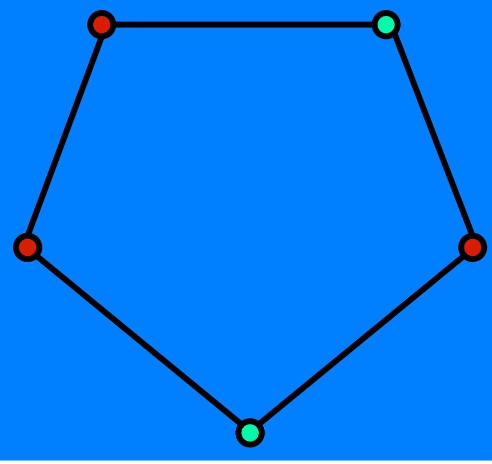
The size of the smallest irreversible 2- conversion set in C_n is ceiling[n/2].

k-Conversion Sets in Regular Graphs

Corollary (Dreyer 2000):

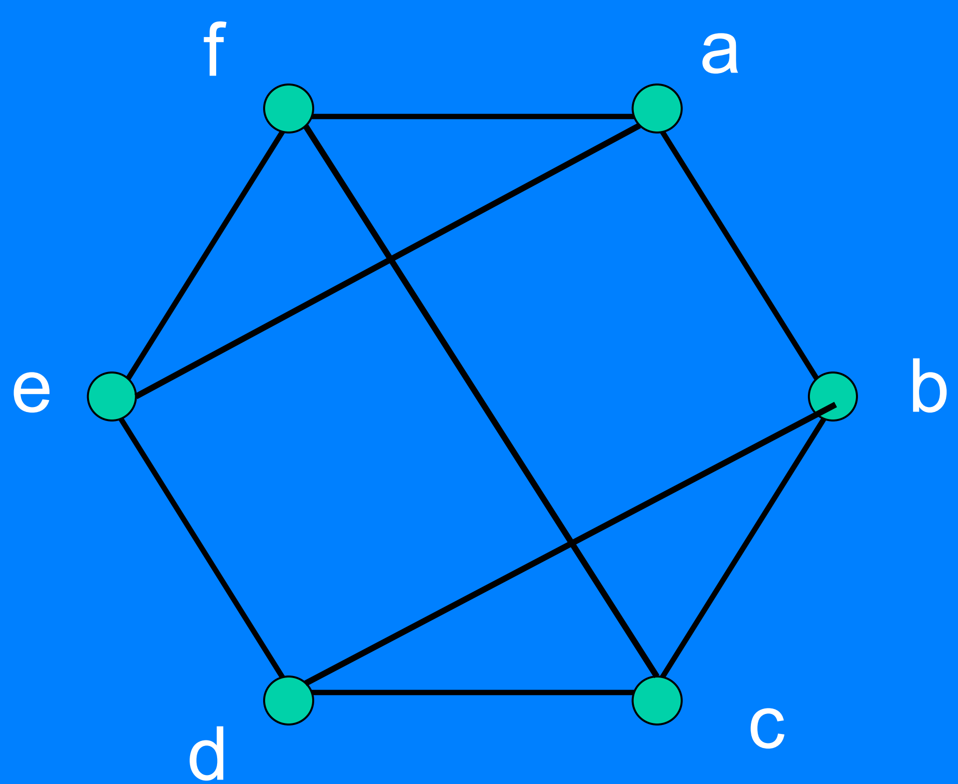
The size of the smallest irreversible 2- conversion set in C_n is ceiling[n/2].

C_5 is 2-regular. The smallest irreversible 2-conversion set has three vertices: the red ones.



k-Conversion Sets in Regular Graphs

Another Example:



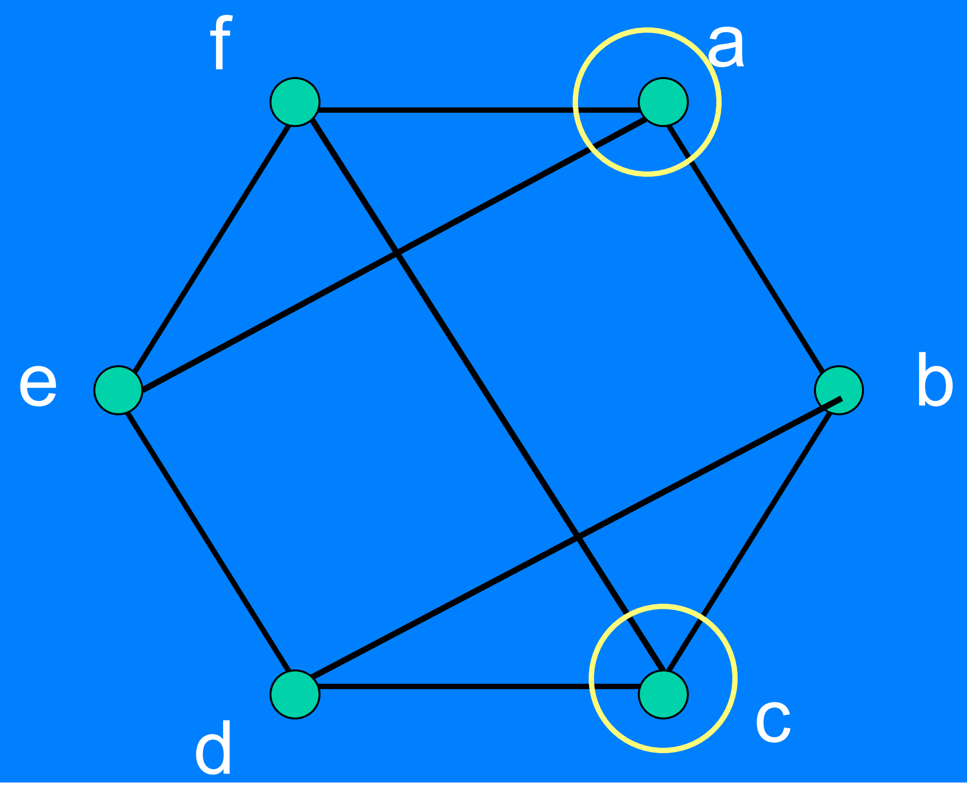
k-Conversion Sets in Regular Graphs

Another Example:

This is 3-regular.

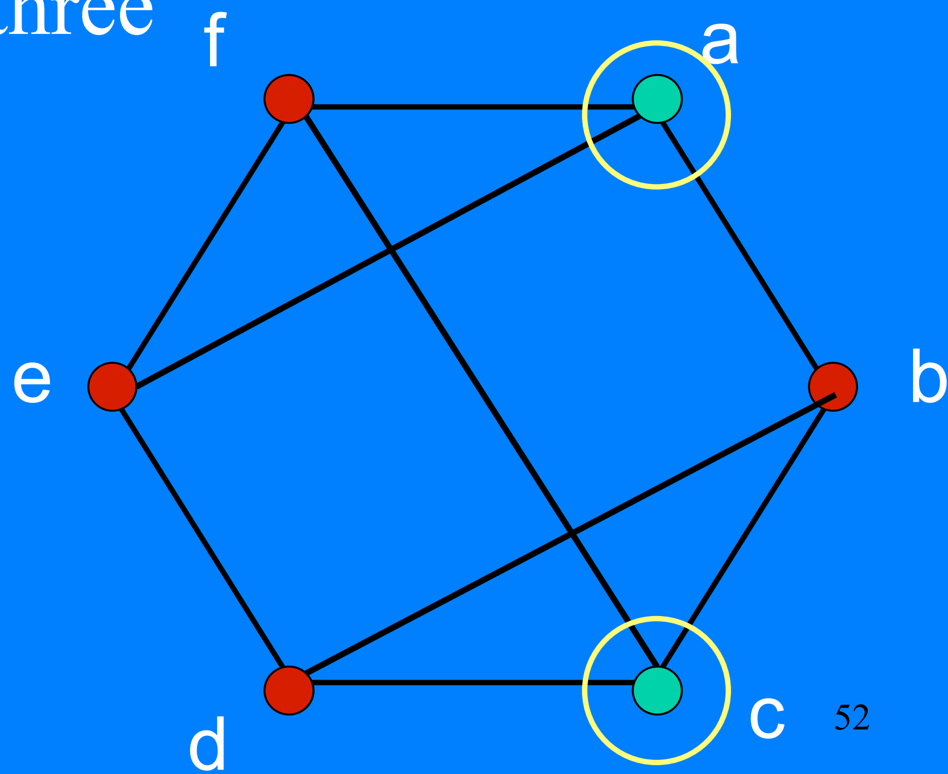
Let k = 3.

The largest independent set has 2 vertices.



k-Conversion Sets in Regular Graphs

- The largest independent set has 2 vertices.
- Thus, the smallest irreversible 3-conversion set has 6-2 = 4 vertices.
- The 4 red vertices form such a set.
- Each other vertex has three red neighbors.



Irreversible k-Conversion Sets in Graphs of Maximum Degree r

Theorem (Dreyer 2000): Let $G = (V,E)$ be a connected graph with maximum degree r and S be the set of all vertices of degree $< r$. If D is a set of vertices, then D is an irreversible r-conversion set iff $S \subseteq D$ and $V-D$ is an independent set.

How Hard is it to Find out if There is an Irreversible k-Conversion Set of Size at Most p?

Problem **IRREVERSIBLE k-CONVERSION SET**: Given a positive integer p and a graph G, does G have an irreversible k-conversion set of size at most p?

How hard is this problem?

Difficulty of Finding Irreversible Conversion Sets

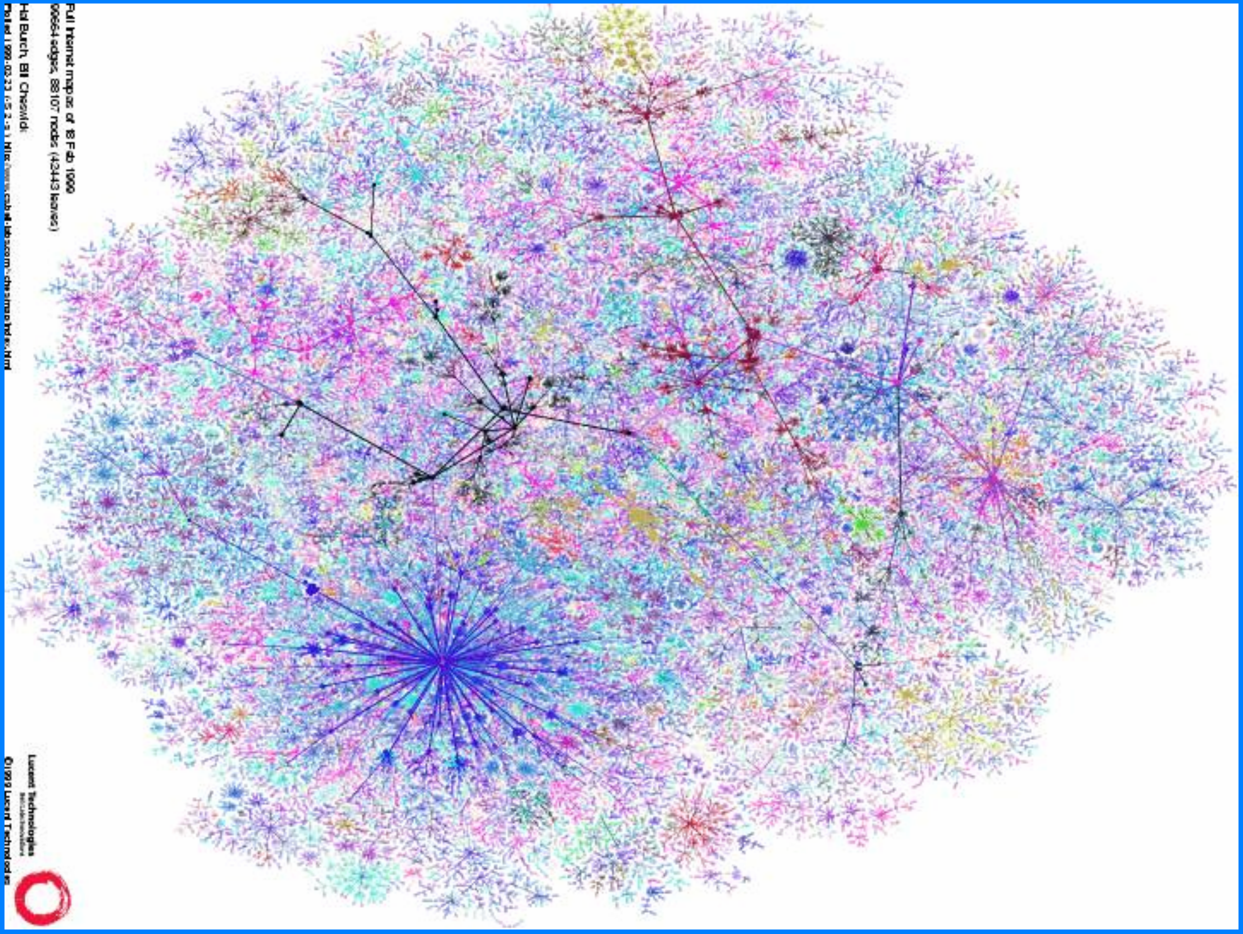
Problem **IRREVERSIBLE k-CONVERSION SET**: Given a positive integer p and a graph G, does G have an irreversible k-conversion set of size at most p?

Theorem (Dreyer 2000): IRREVERSIBLE k-CONVERSION SET is NP-complete for fixed $k > 2$.

(Whether or not it is NP-complete for $k = 2$ remains open.)

Irreversible k-Conversion Sets in Special Graphs

Studied for many special graphs.

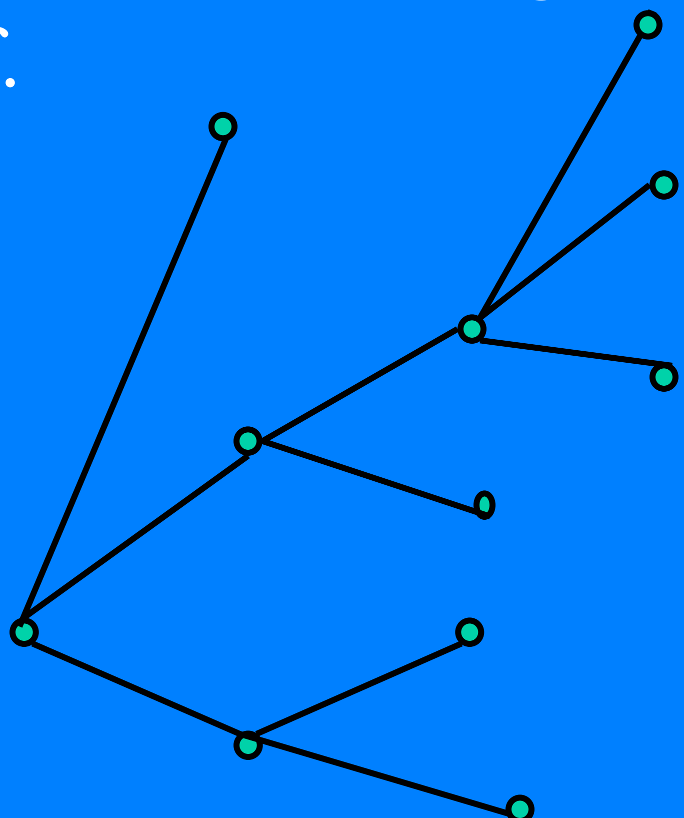


Irreversible k-Conversion Sets in Trees

Irreversible k-Conversion Sets in Trees

The simplest case is when every internal vertex of the tree has degree > k.

Leaf = vertex of degree 1; *internal vertex* = not a leaf.

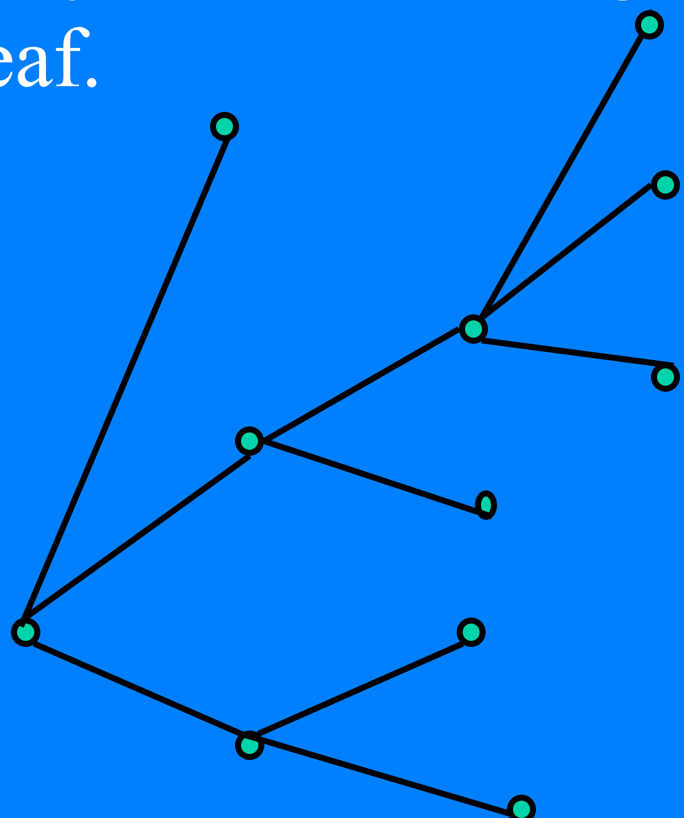


What is an irreversible 2-conversion set here?

Irreversible k-Conversion Sets in Trees

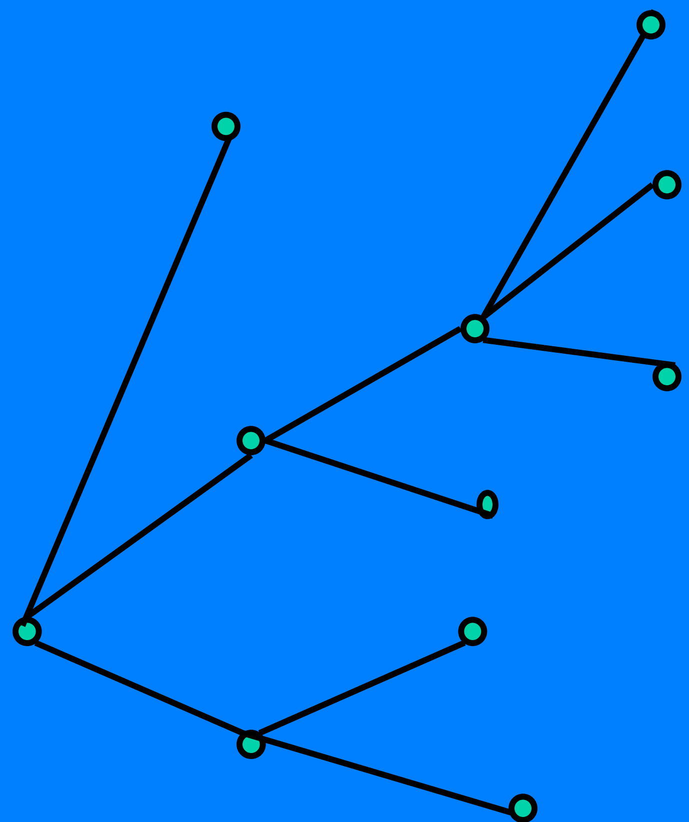
The simplest case is when every internal vertex of the tree has degree > k.

Leaf = vertex of degree 1; *internal vertex* = not a leaf.

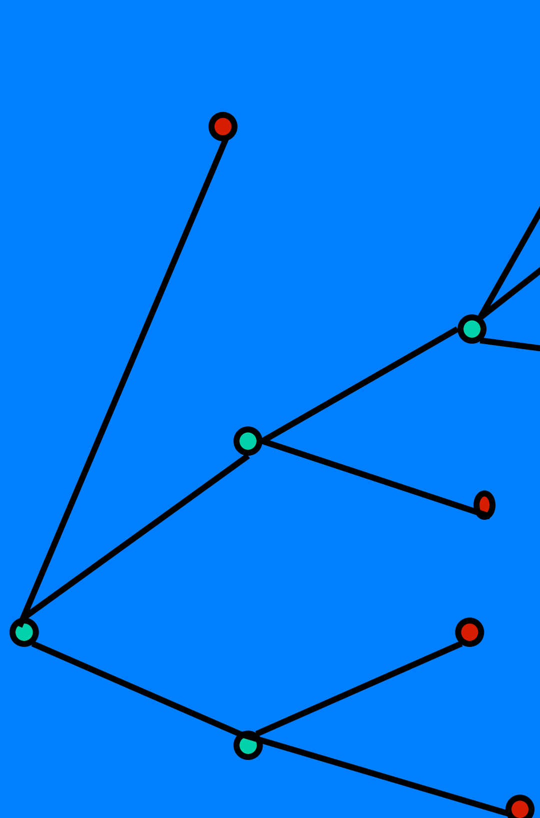


Do you know any vertices that have to be in such a set?

What is an irreversible 2-conversion set here?

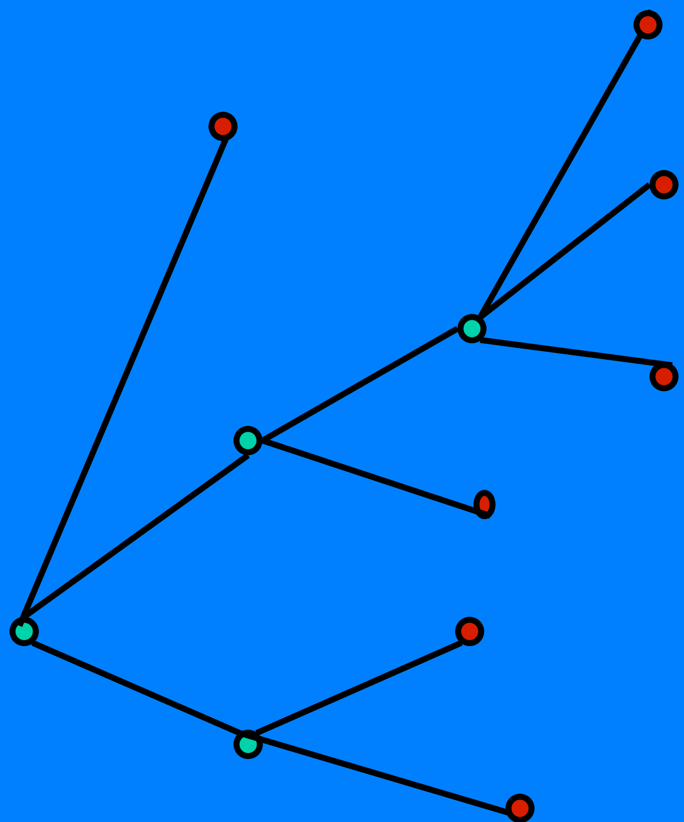


All leaves have to be in it.



All leaves have to be in it.

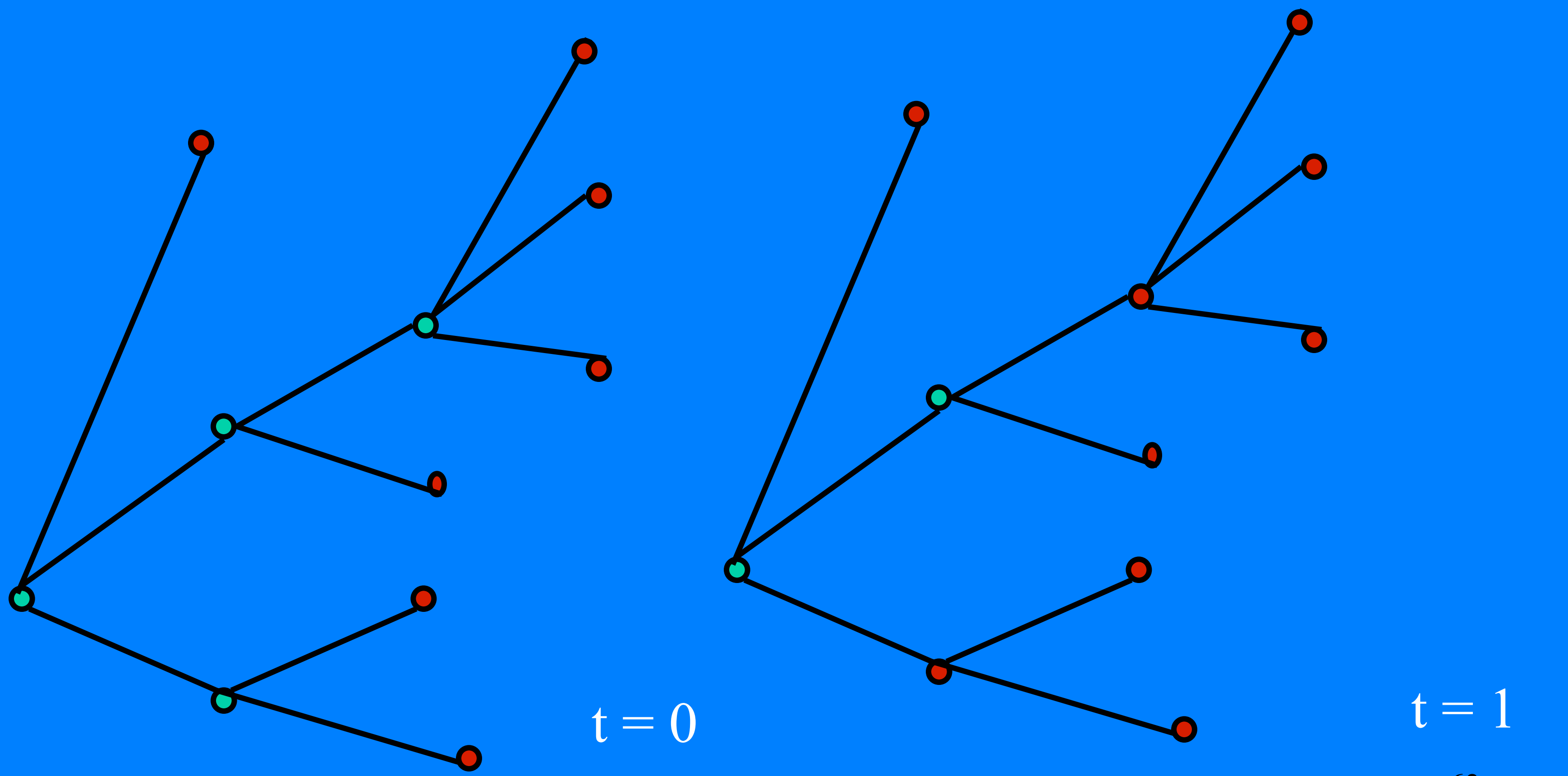
This will suffice.



t = 0

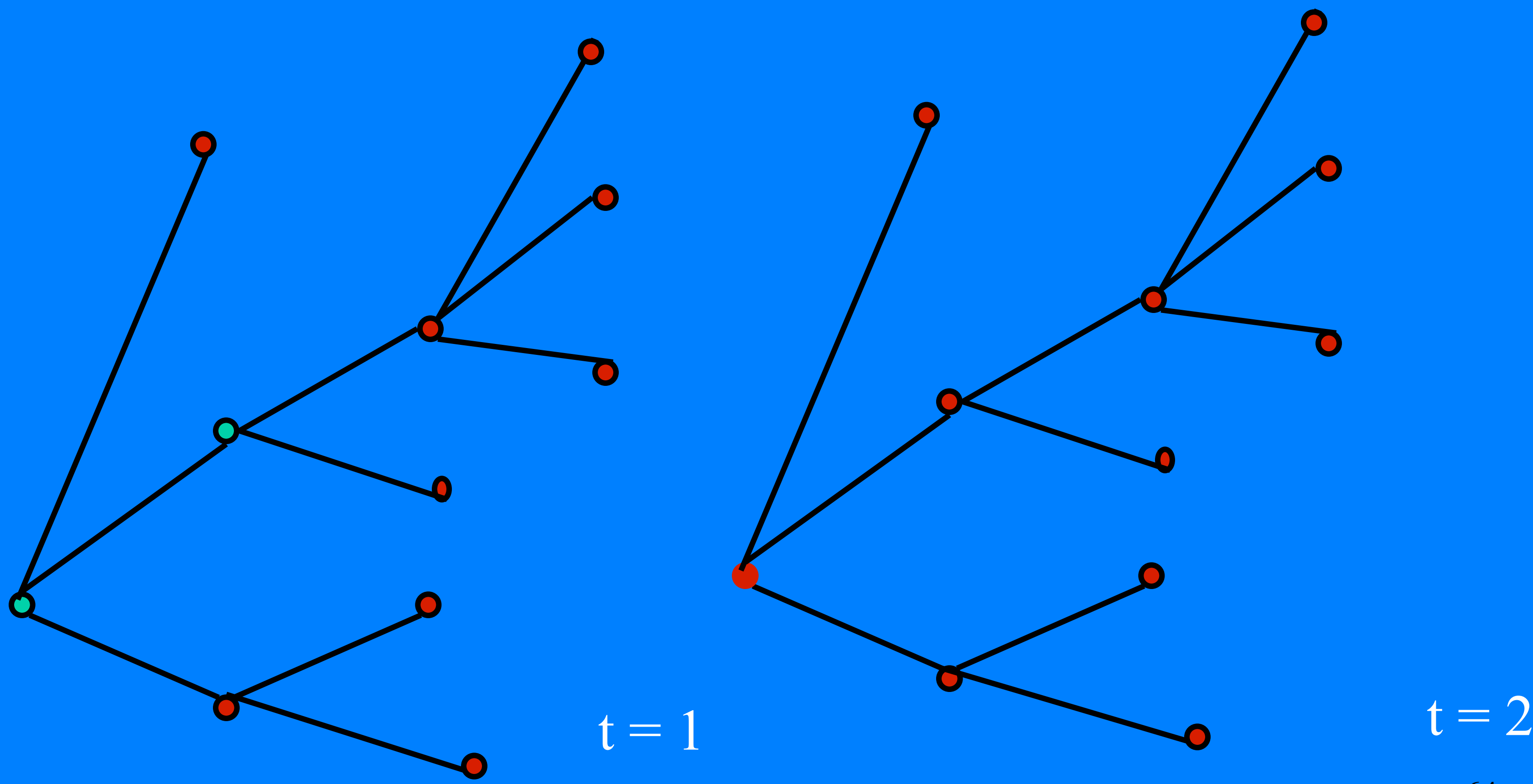
All leaves have to be in it.

This will suffice.



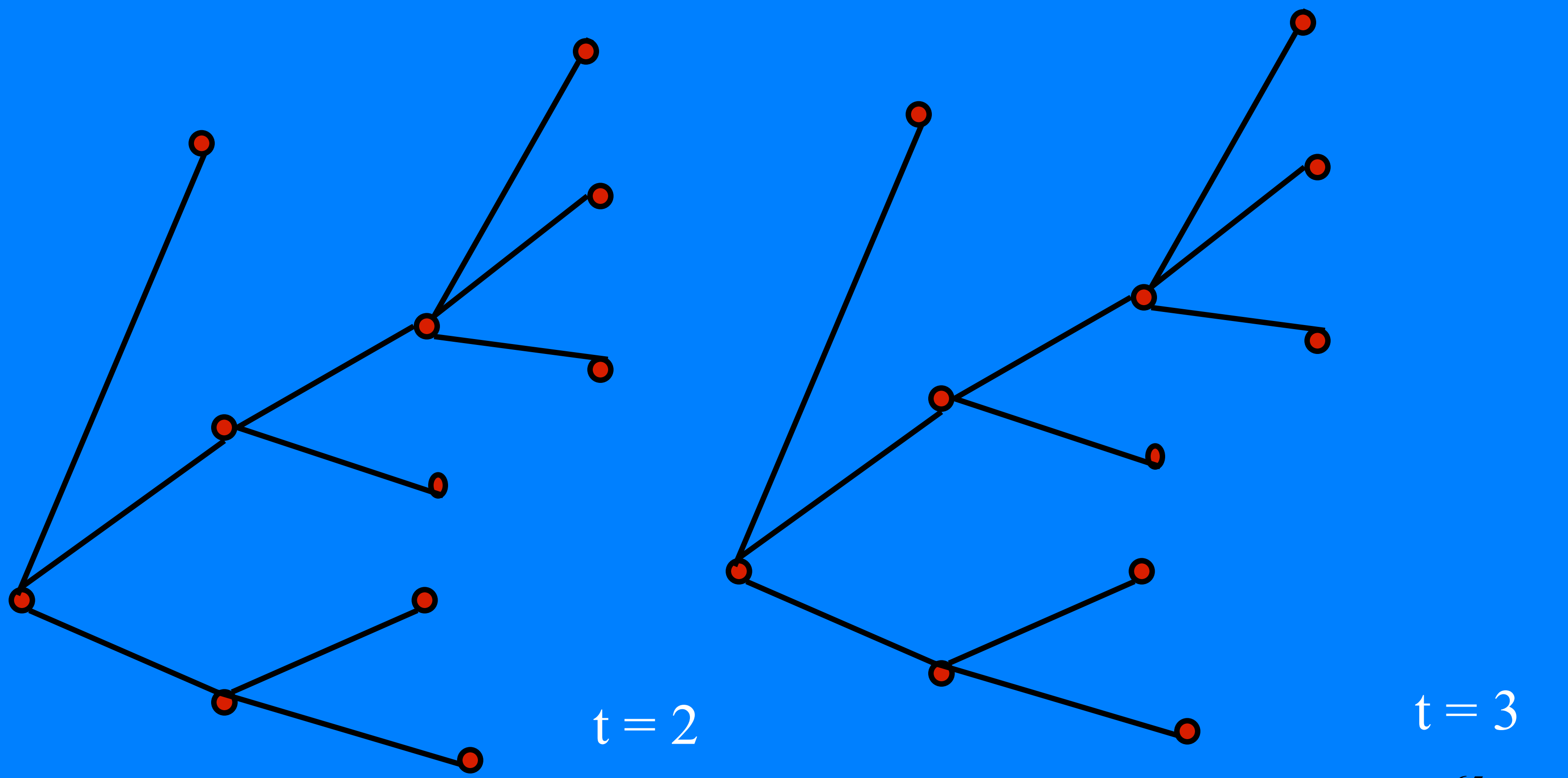
All leaves have to be in it.

This will suffice.



All leaves have to be in it.

This will suffice.



Irreversible k-Conversion Sets in Trees

So k = 2 is easy. What about k > 2? Also easy.

Proposition (Dreyer 2000): Let T be a tree and every internal vertex have degree > k, where k > 1. Then the smallest irreversible k-conversion set has size equal to the number of leaves of the tree.

Irreversible k-Conversion Sets in Trees

What if not every internal vertex has degree > k?

If there is an internal vertex of degree < k, it will have to be in any irreversible k-conversion set and will never change sign.

So, to every neighbor, this vertex v acts like a leaf, and we can break T into deg(v) subtrees with v a leaf in each.

If every internal vertex has degree ≥ k, one can obtain analogous results to those for the > k case by looking at maximal connected subsets of vertices of degree k.

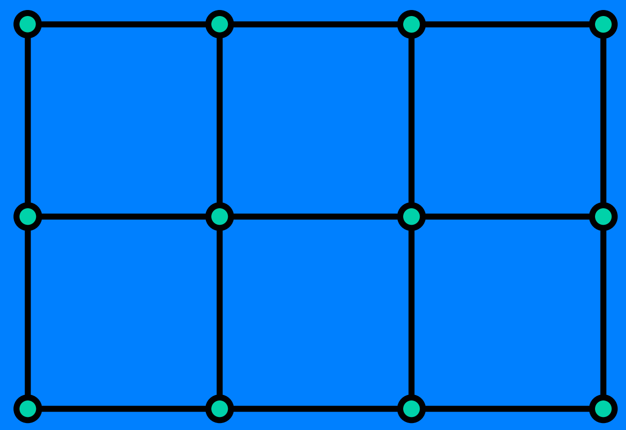
Irreversible k-Conversion Sets in Trees

Dreyer presents an O(n) algorithm for finding the size of the smallest irreversible k-conversion set in a tree of n vertices.

Irreversible k-Conversion Sets in Special Graphs

Studied for many special graphs.

Let G(m,n) be the ***rectangular grid graph*** with m rows and n columns.

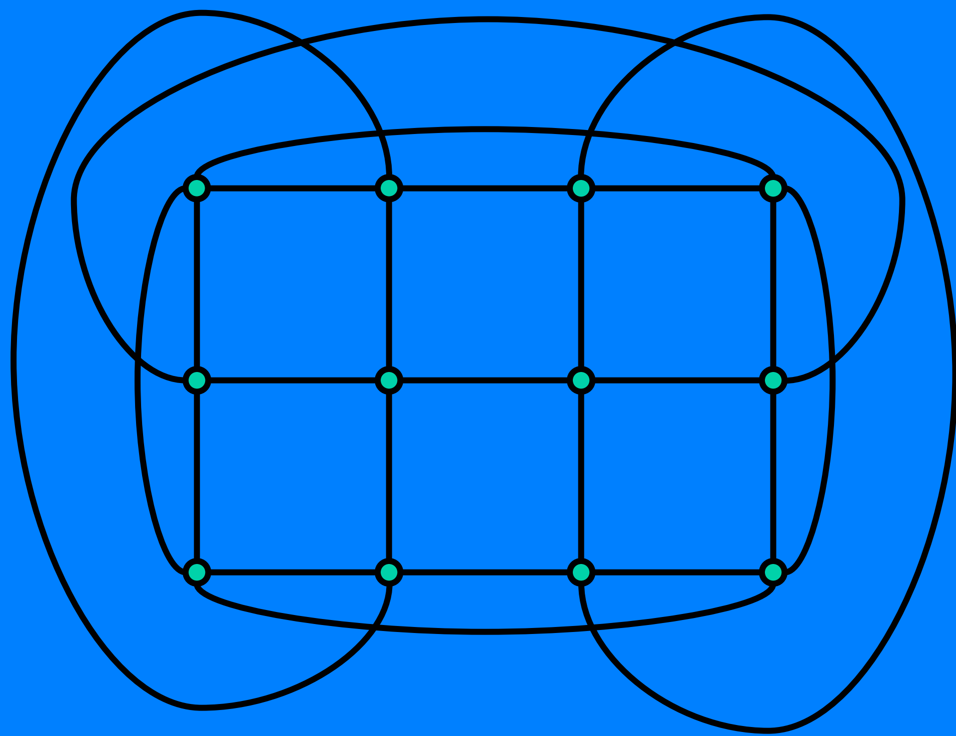


G(3,4)

Toroidal Grids

The ***toroidal grid*** T(m,n) is obtained from the rectangular grid G(m,n) by adding edges from the first vertex in each row to the last and from the first vertex in each column to the last.

Toroidal grids are easier to deal with than rectangular grids because they form regular graphs: Every vertex has degree 4. Thus, we can make use of the results about regular graphs.



T(3,4)

Irreversible4-Conversion Sets in Toroidal Grids

Theorem (Dreyer 2000): In a toroidal grid T(m,n), the size of the smallest irreversible 4-conversion set is

$$\begin{cases} \max\{n(\text{ceiling}[m/2]), m(\text{ceiling}[n/2])\} & \text{m or n odd} \\ mn/2 & \text{m, n even} \end{cases}$$

Part of the Proof: Recall that D is an irreversible 4-conversion set in a 4-regular graph iff V-D is independent.

V-D independent means that every edge {u,v} in G has u or v in D. In particular, the ith row must contain at least ceiling[n/2] vertices in D and the ith column at least ceiling[m/2] vertices in D (alternating starting with the end vertex of the row or column).

We must cover all rows and all columns, and so need at least max{n(ceiling[m/2]), m(ceiling[n/2])} vertices in an irreversible 4-conversion set.

Irreversible k-Conversion Sets for Rectangular Grids

Let $C_k(G)$ be the size of the smallest irreversible k-conversion set in graph G.

Theorem (Dreyer 2000):

$$C_4[G(m,n)] = 2m + 2n - 4 + \text{floor}[(m-2)(n-2)/2]$$

Theorem (Flocchini, Lodi, Luccio, Pagli, and Santoro):

$$C_2[G(m,n)] = \text{ceiling}([m+n]/2)$$

Irreversible 3-Conversion Sets for Rectangular Grids

For 3-conversion sets, the best we have are bounds:

Theorem (Flocchini, Lodi, Luccio, Pagli, and Santoro):

$$[(m-1)(n-1)+1]/3 \leq C_3[G(m,n)] \leq [(m-1)(n-1)+1]/3 + [3m+2n-3]/4 + 5$$

Finding the exact value is an open problem.

Vaccination Strategies

Stephen Hartke and others worked on a different problem:

Defender: can vaccinate v people ***per time period***.

Attacker: can only infect people at the beginning.
Irreversible k-threshold model.

What vaccination strategy minimizes number of people infected?

Sometimes called the ***firefighter problem***:

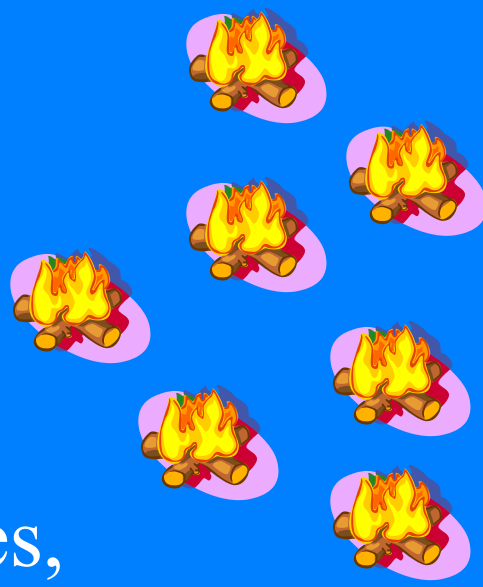
alternate fire spread and firefighter placement.

Usual assumption: k = 1. (We will assume this.)

Variation: The vaccinator and infector alternate turns, having v vaccinations per period and i doses of pathogen per period. What is a good strategy for the vaccinator?

Problem goes back to Bert Hartnell 1995

A Survey of Some Results on the Firefighter Problem



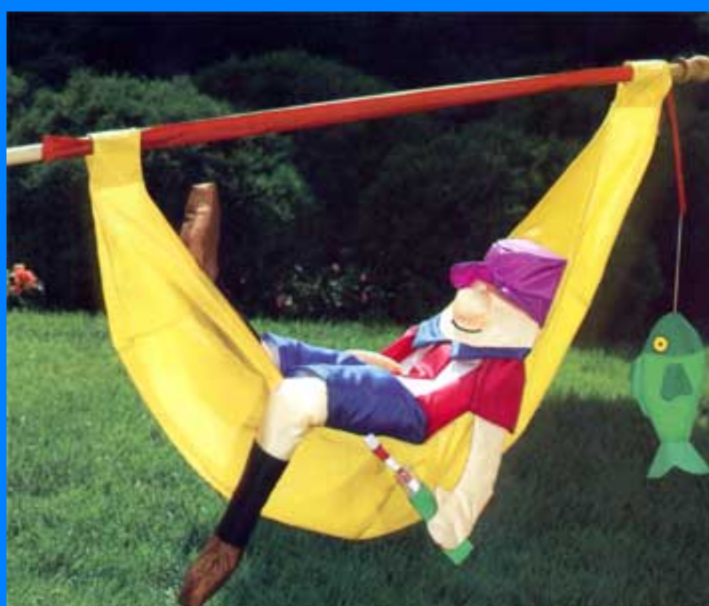
Thanks to
Kah Loon Ng
DIMACS
For the animated slides,
slightly modified by me

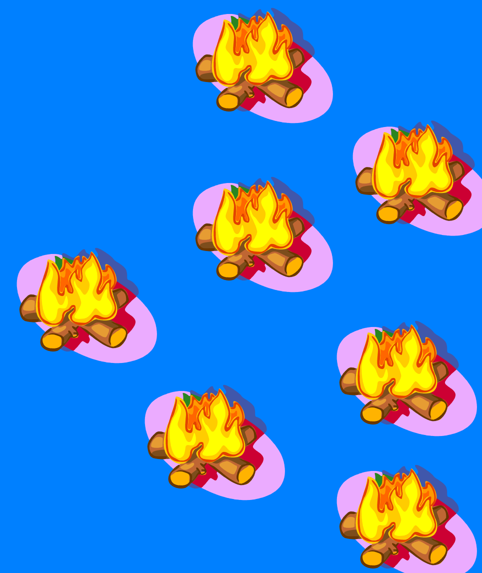
Mathematicians can be Lazy

Mathematicians can be Lazy

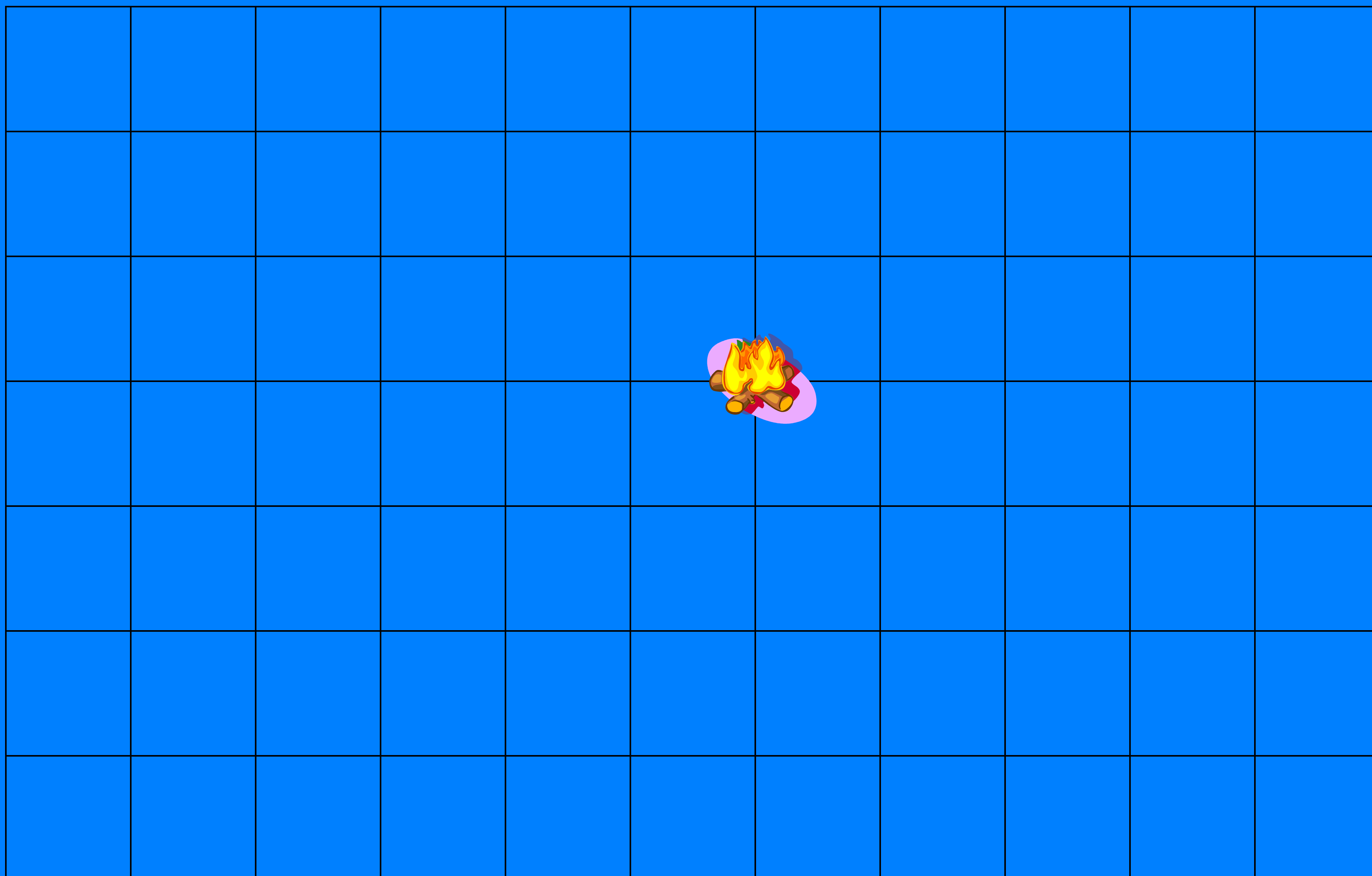
- Different application.
- Different terminology
- Same mathematical model.

measles

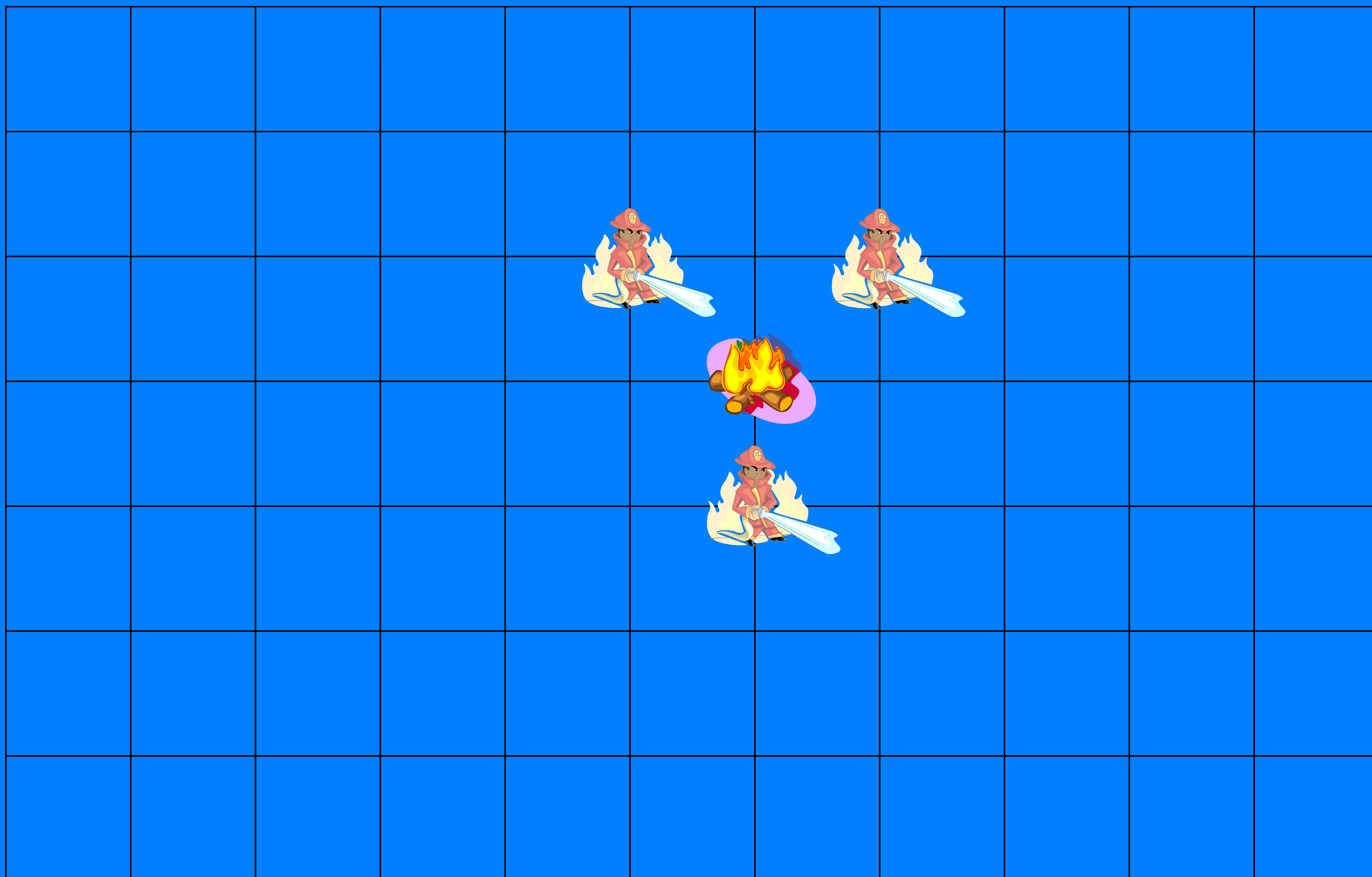




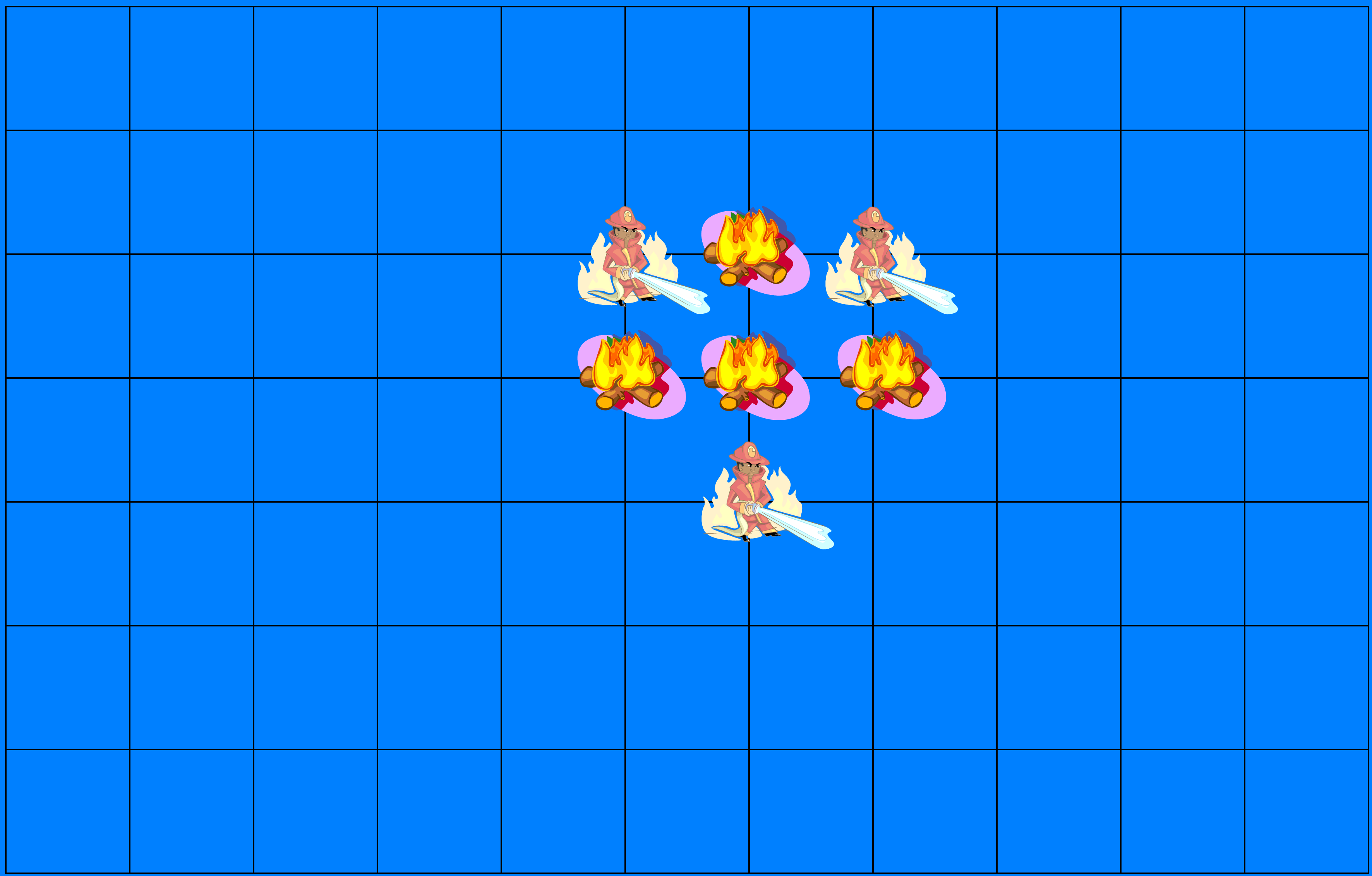
A Simple Model (k = 1) (v = 3)



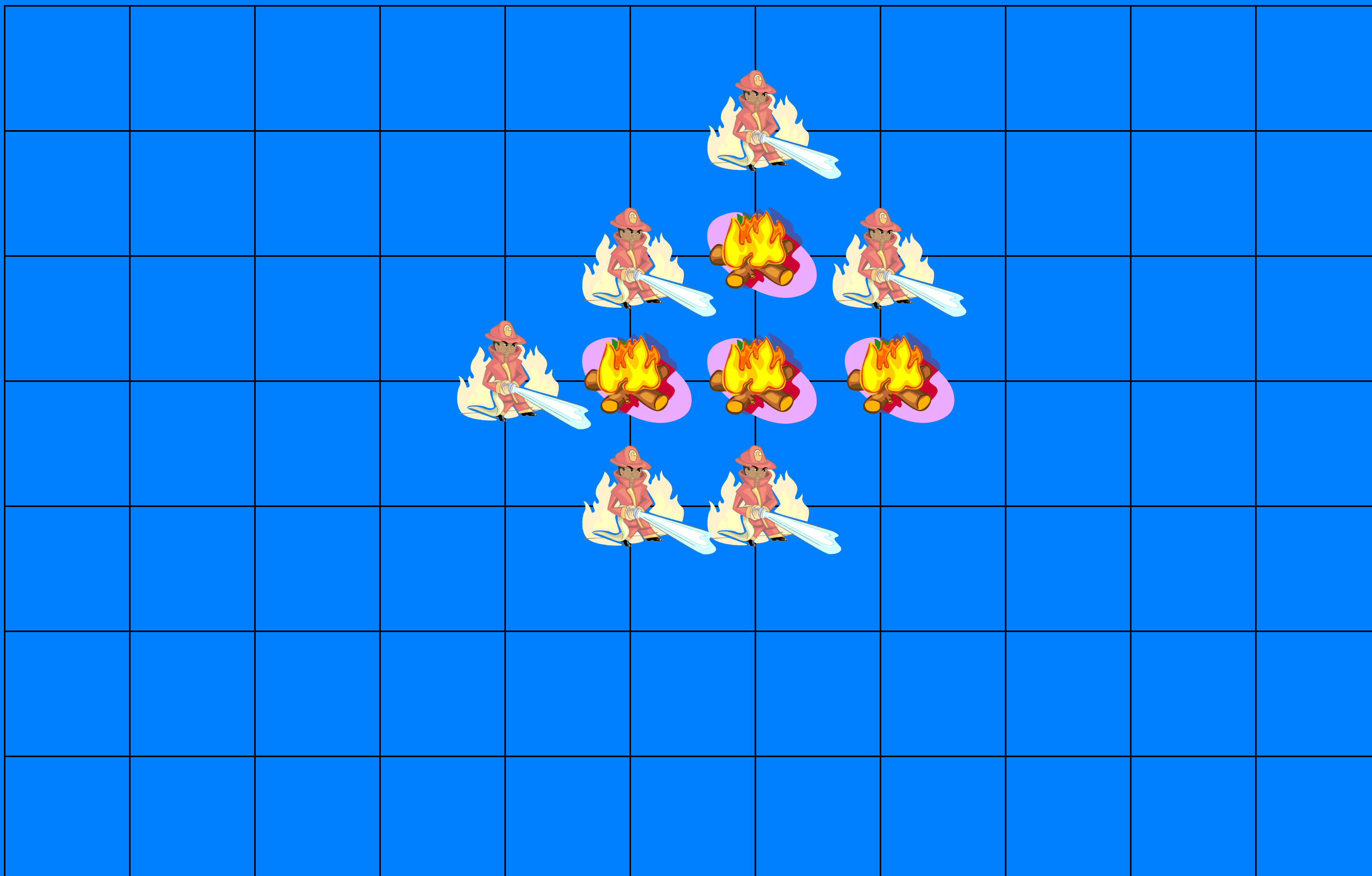
A Simple Model



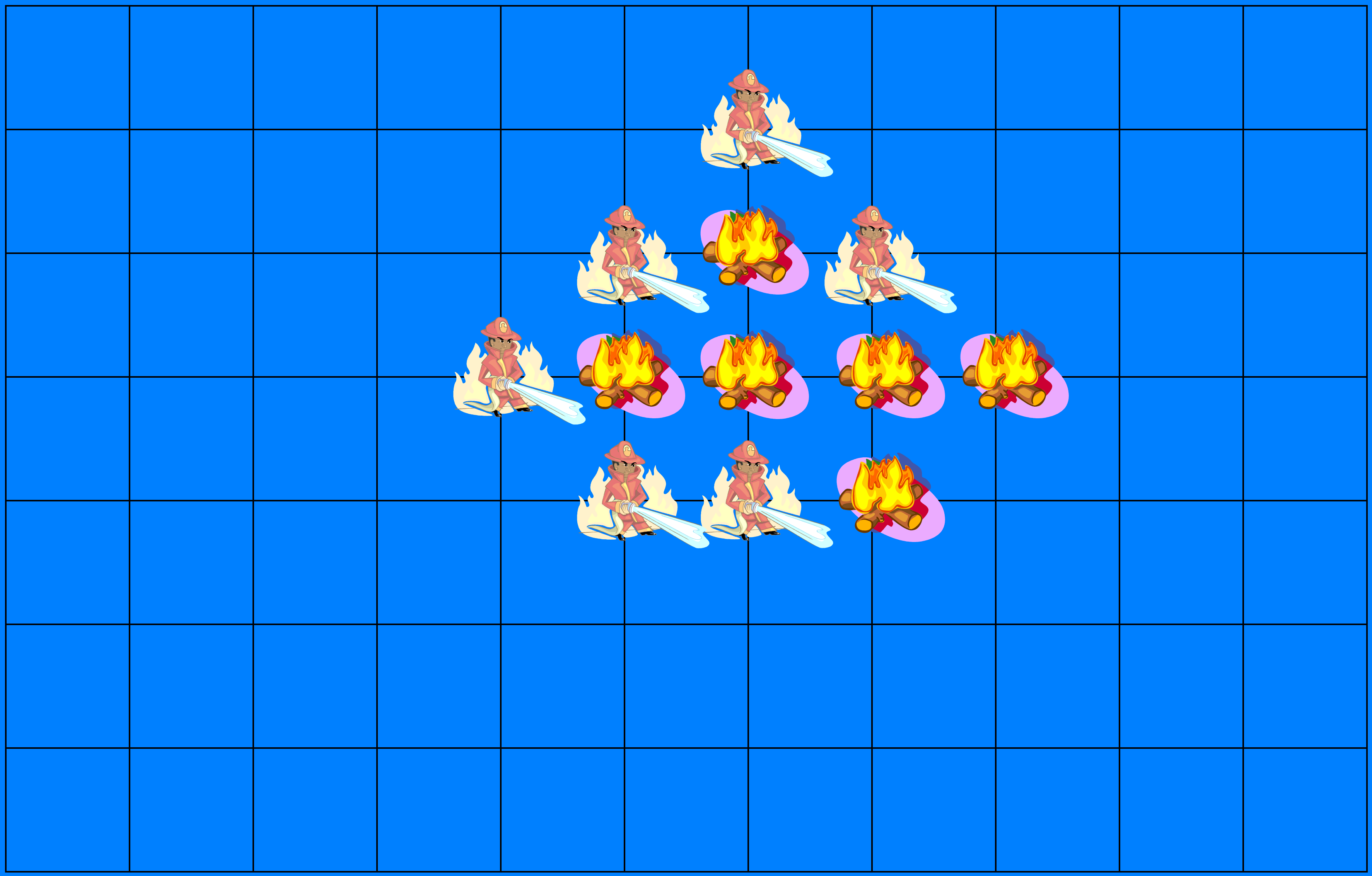
A Simple Model



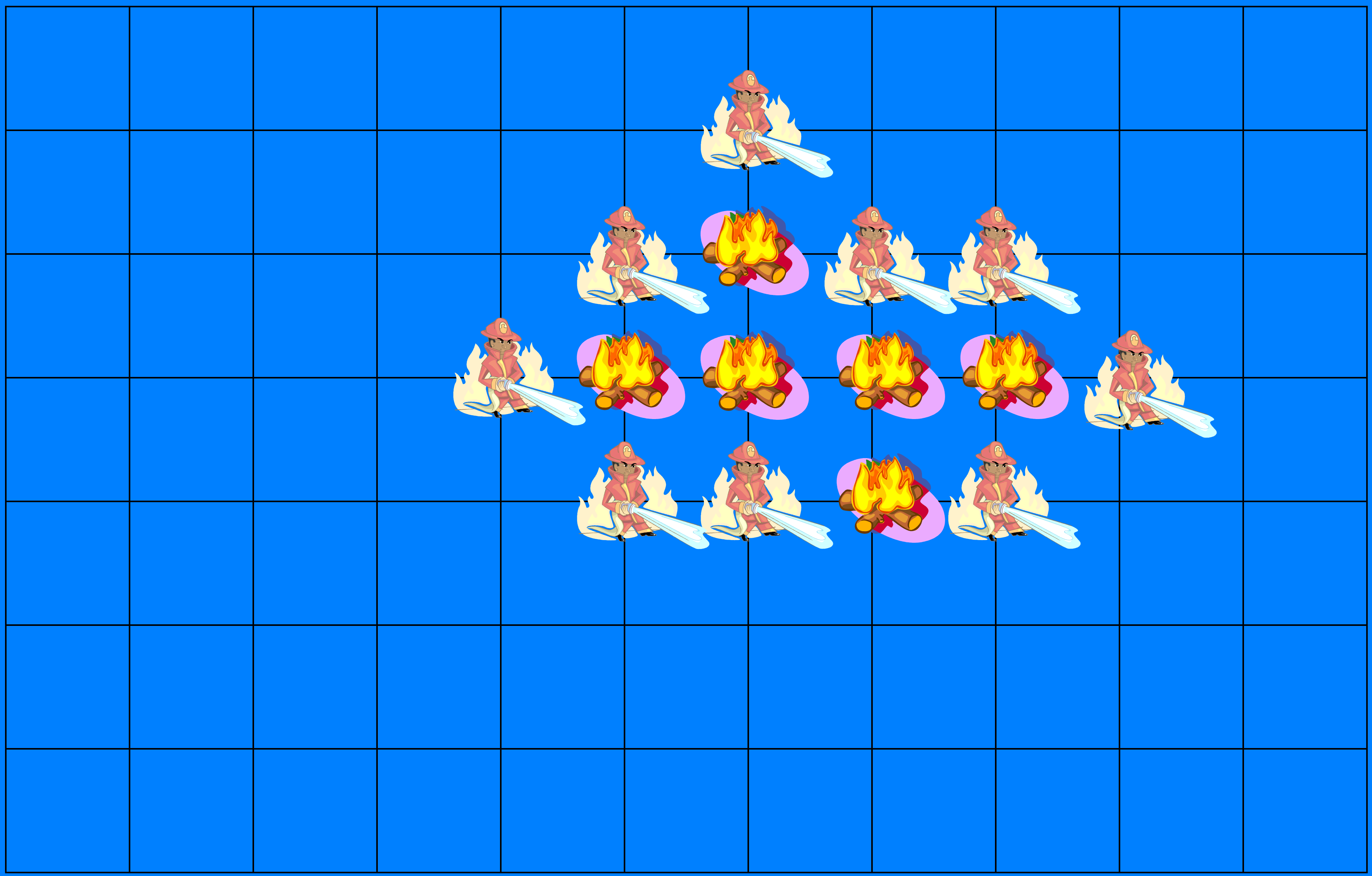
A Simple Model



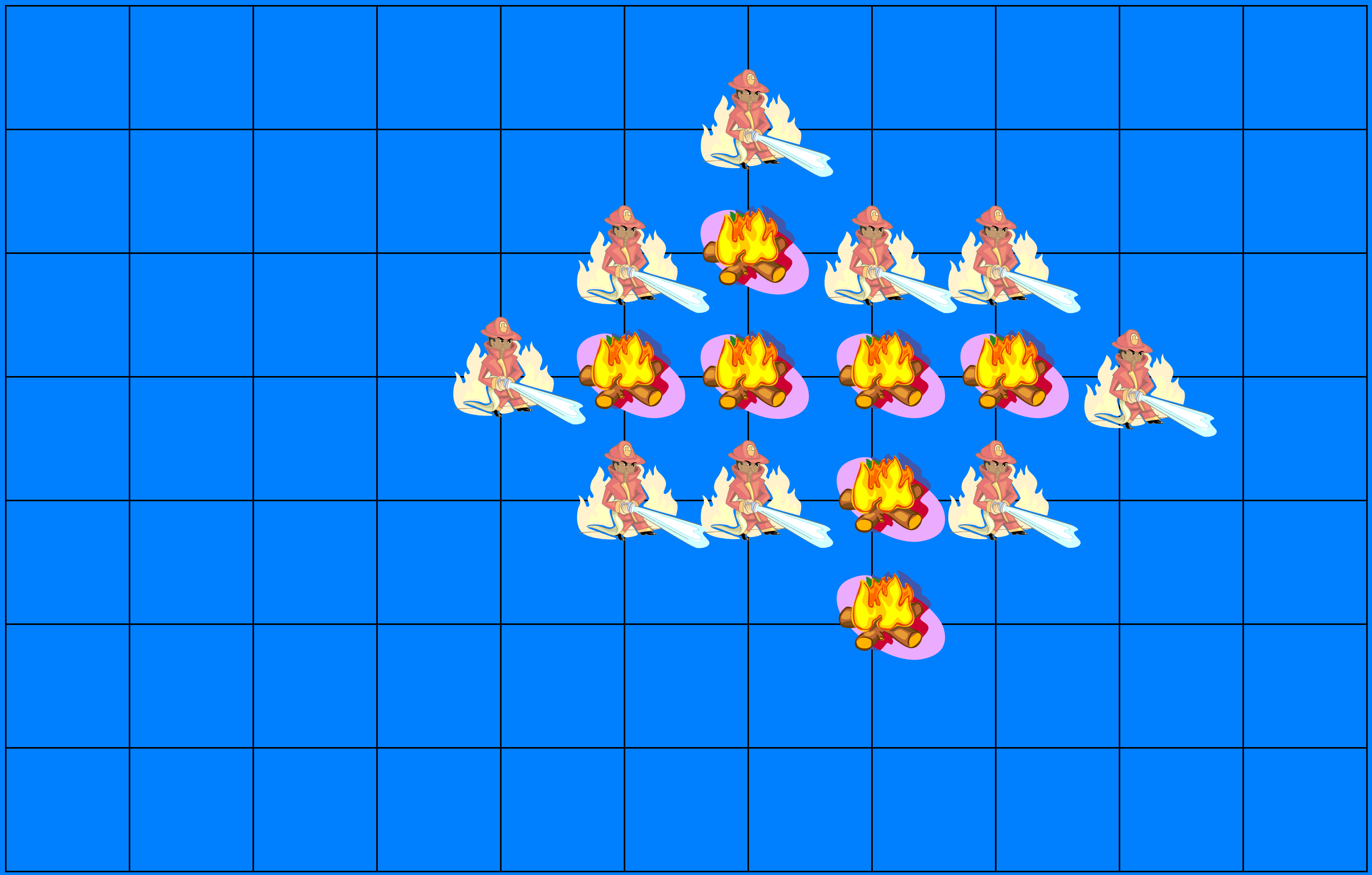
A Simple Model



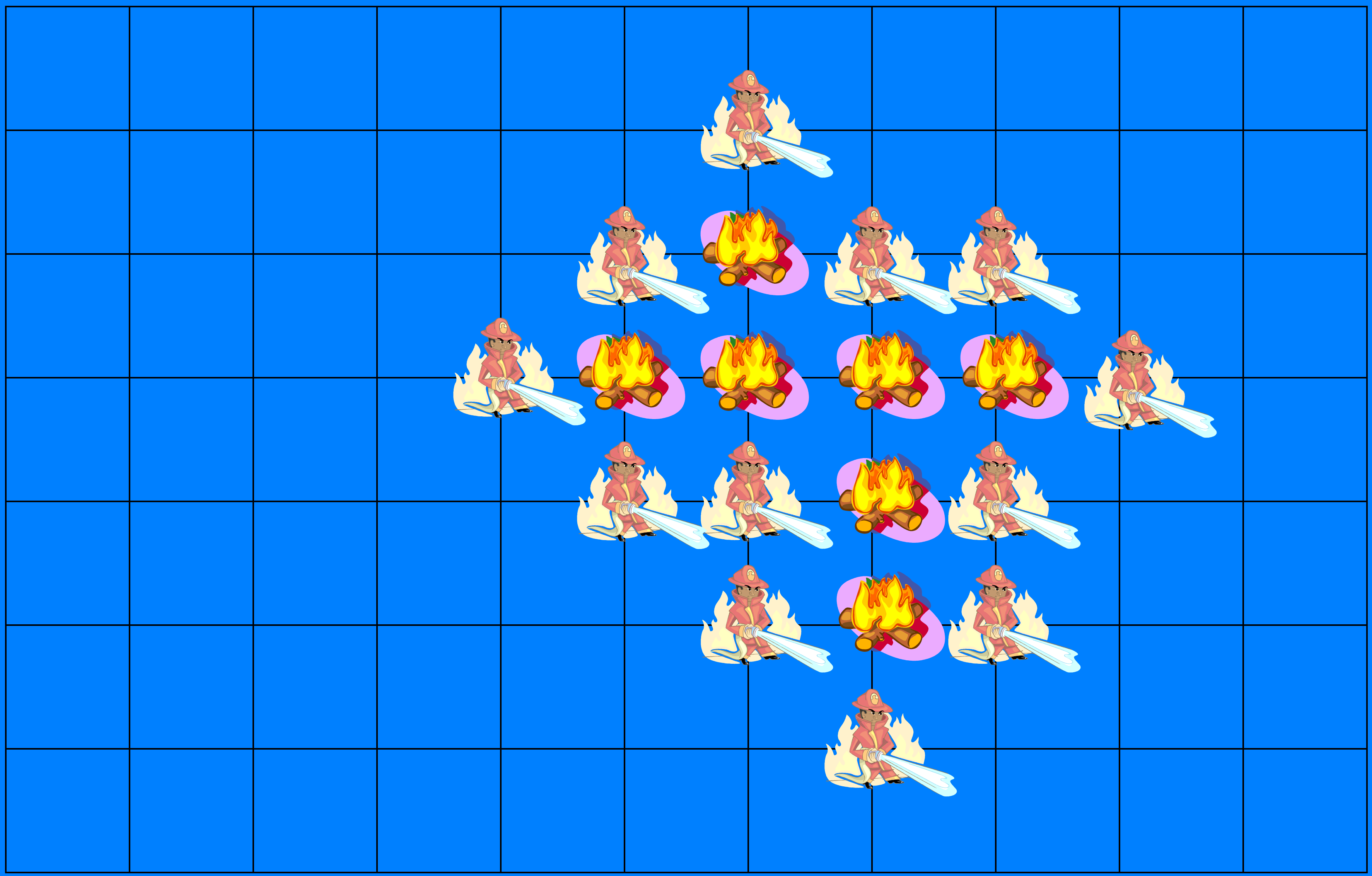
A Simple Model



A Simple Model



A Simple Model



Some questions that can be asked (but not necessarily answered!)

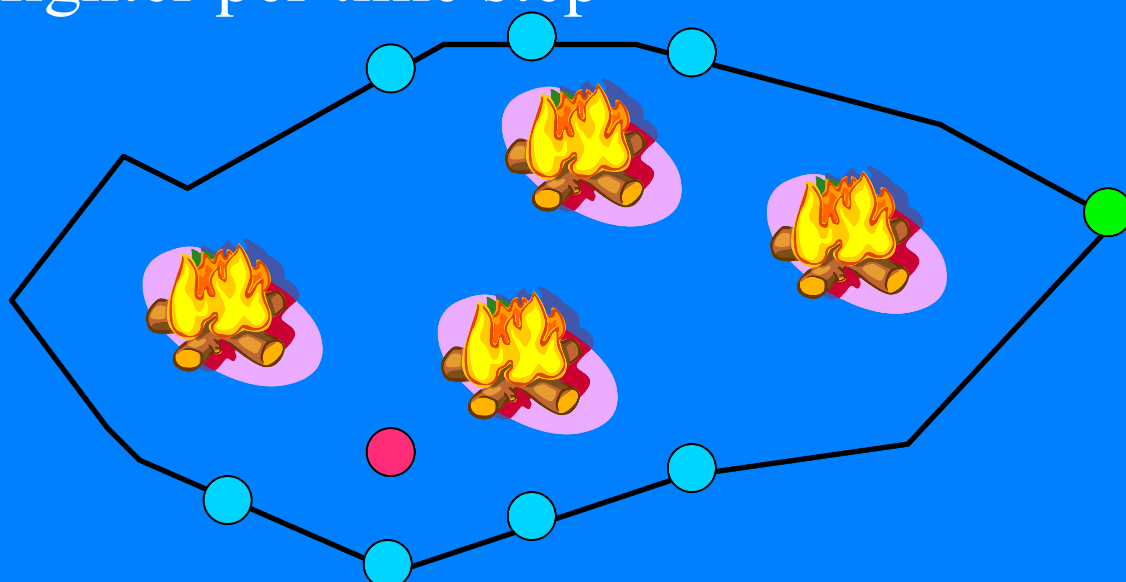
- Can the fire be contained?
- How many time steps are required before fire is contained?
- How many firefighters per time step are necessary?
- What fraction of all vertices will be saved (burnt)?
- Does where the fire breaks out matter?
- Fire starting at more than 1 vertex?
- Consider different graphs. Construction of (connected) graphs to minimize damage.
- Complexity/Algorithmic issues

Containing Fires in Infinite Grids L_d

Fire starts at only one vertex:

d =1: Trivial.

d = 2: Impossible to contain the fire with 1 firefighter per time step

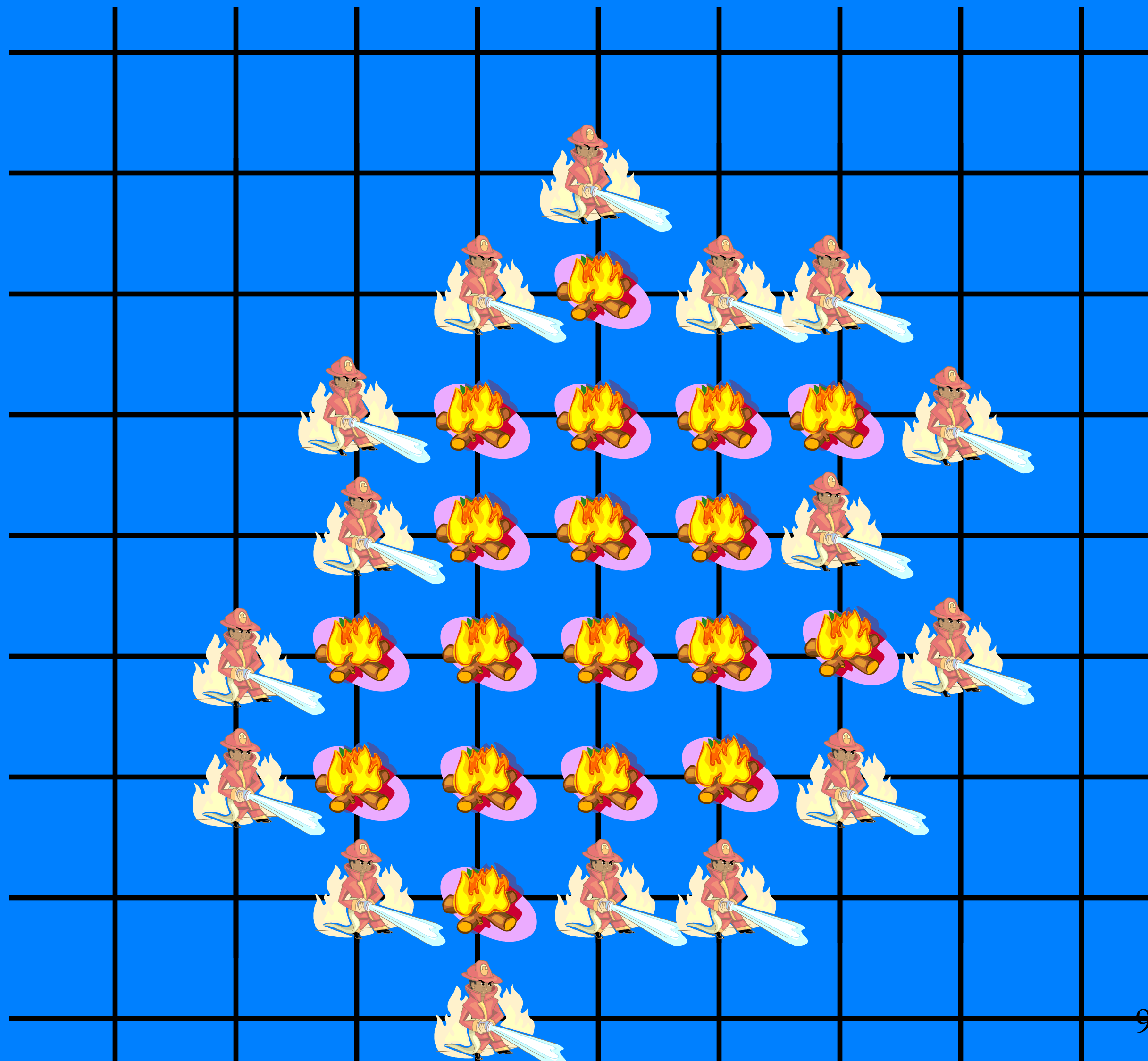


Containing Fires in Infinite Grids L_d

$d = 2$: Two firefighters per time step needed to contain the fire.

8 time steps

18 burnt vertices



Containing Fires in Infinite Grids L_d

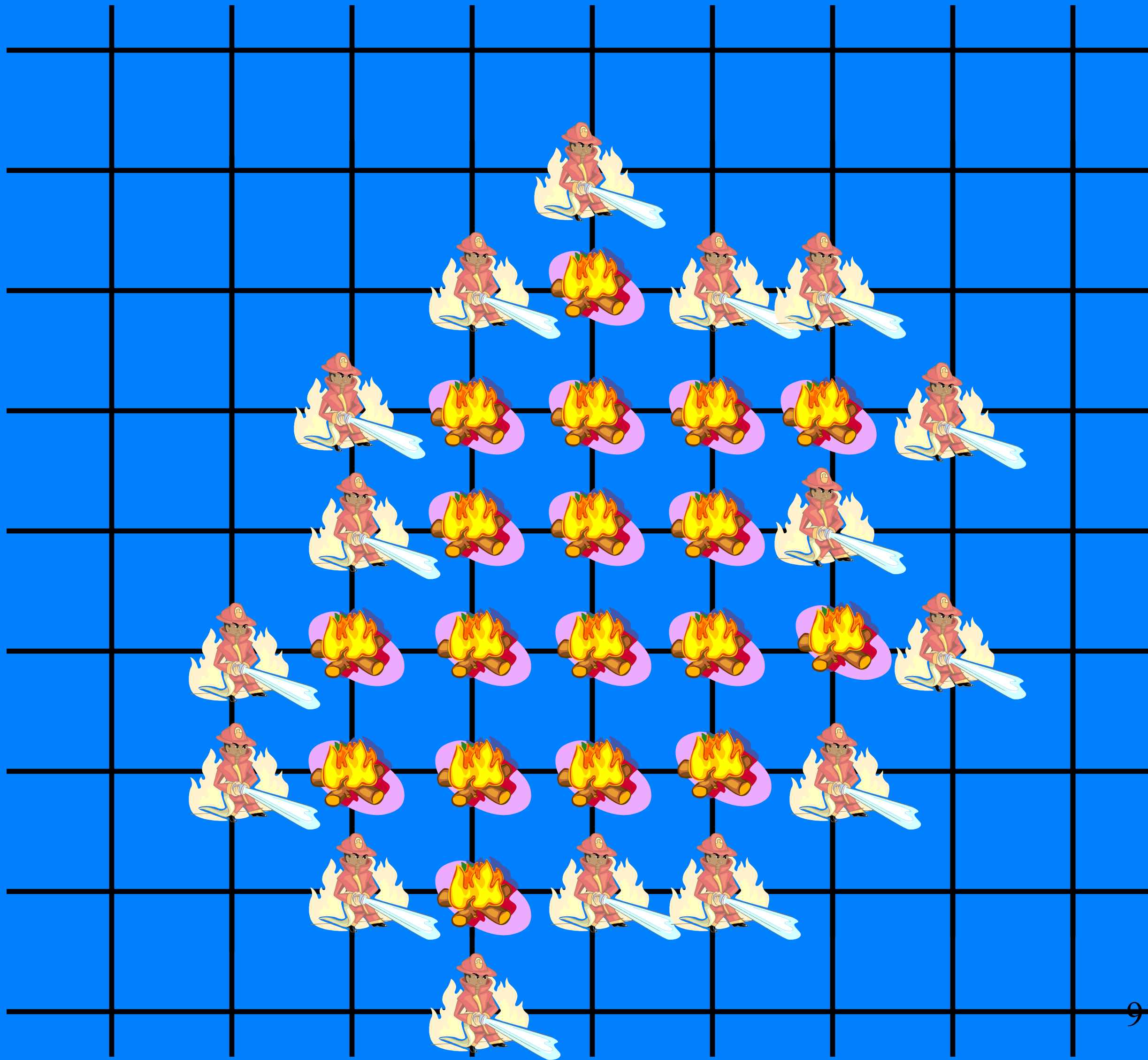
$d = 2$: Two firefighters per time step needed to contain the fire.

8 time steps

18 burnt vertices

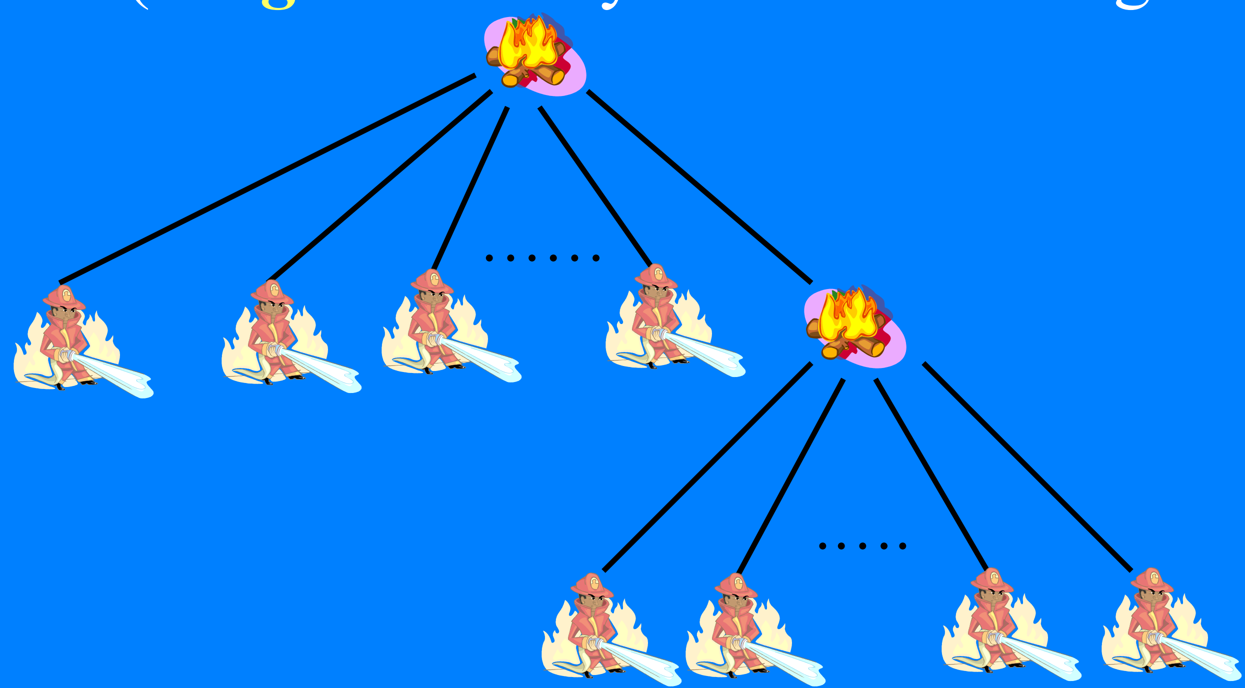
Develin & Hartke (2007): cannot do better than 18

Wang & Moeller (2002): Cannot contain fire in < 8 steps



Containing Fires in Infinite Grids L_d

$d \geq 3$: Wang and Moeller (2002): If G is an r-regular graph, $r - 1$ firefighters per time step is always sufficient to contain any fire outbreak (at a single vertex) in G. (***r-regular***: every vertex has r neighbors.)



Containing Fires in Infinite Grids L_d

$d \geq 3$: In L_d, every vertex has degree $2d$.

Thus: $2d$-1 firefighters per time step are sufficient to contain any outbreak starting at a single vertex.

<u>Theorem (Hartke 2004):</u> If $d \geq 3$, $2d - 2$ firefighters per time step are not enough to contain an outbreak in L_d.

Thus, $2d - 1$ firefighters per time step is the minimum number required to contain an outbreak in L_d and containment can be attained in 2 time steps.

Containing Fires in Infinite Grids L_d

Fire can start at more than one vertex.

$d = 2$: Fogarty (2003): Two firefighters per time step are sufficient to contain any outbreak at a finite number of vertices.

$d \geq 3$: Hartke (2004): For any $d \geq 3$ and any positive integer f, f firefighters per time step is not sufficient to contain all finite outbreaks in L_d. In other words, for $d \geq 3$ and any positive integer f, there is an outbreak such that f firefighters per time step cannot contain the outbreak.

Containing Fires in Infinite Grids L_d

The case of a different number of firefighters per time step.

Let f(t) = number firefighters available at time t.

Assume f(t) is periodic with period p_f.

Possible motivations for periodicity:

•Firefighters arrive in batches.

•Firefighters need to stay at a vertex for several time periods before redeployment.

Containing Fires in Infinite Grids L_d

The case of a different number of firefighters per time step.

$N_f = f(1) + f(2) + \ldots + f(p_f)$

$R_f = N_f/p_f$

(average number firefighters available per time period)

Theorem (Ng and Raff 2006): If $d = 2$ and f is periodic with period $p_f \geq 1$ and $R_f > 1.5$, then an outbreak at any number of vertices can be contained at a finite number of vertices.

Containing Fires in Infinite Grids L_d

The case of a different number of firefighters per time step.

Conjecture (Develin and Hartke 2007): Suppose that $f(t)/t^{d-2}$ goes to 0 as t gets large. Then there is some fire on L_d that cannot be contained by deploying f(t) firefighters at time t.

Containing Fires in Infinite Grids

Other work has been done on infinite triangular grids and infinite hexagonal grids

Saving Vertices in Finite Grids *G*

Assumptions:

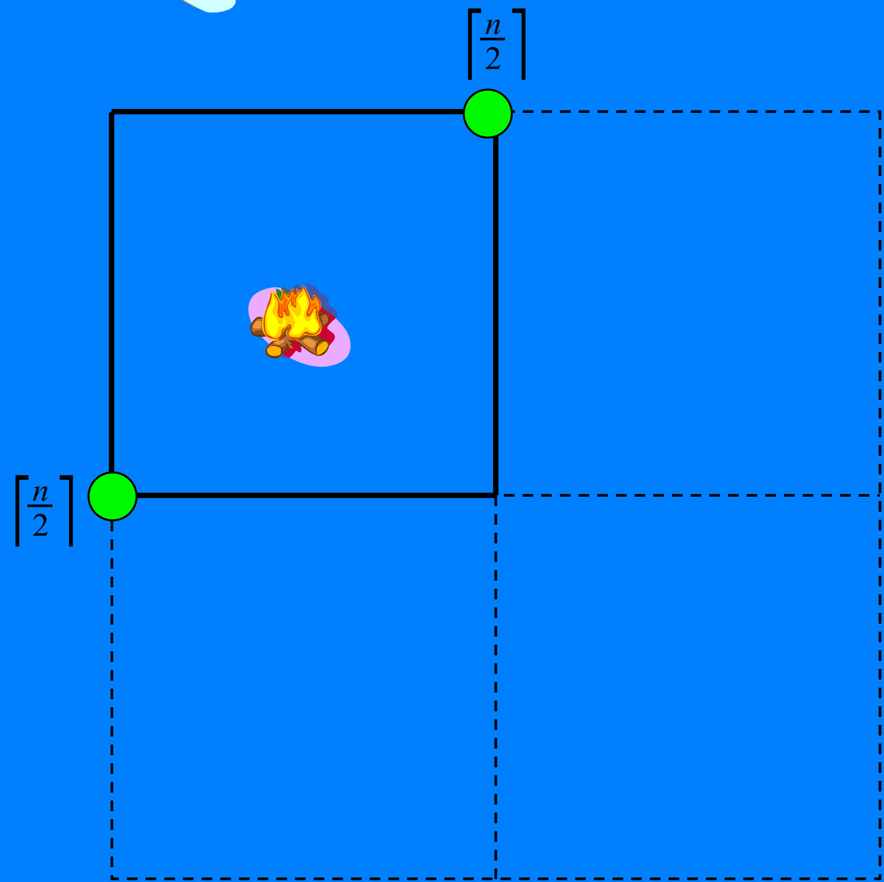
1. 1 firefighter is deployed per time step
2. Fire starts at one vertex

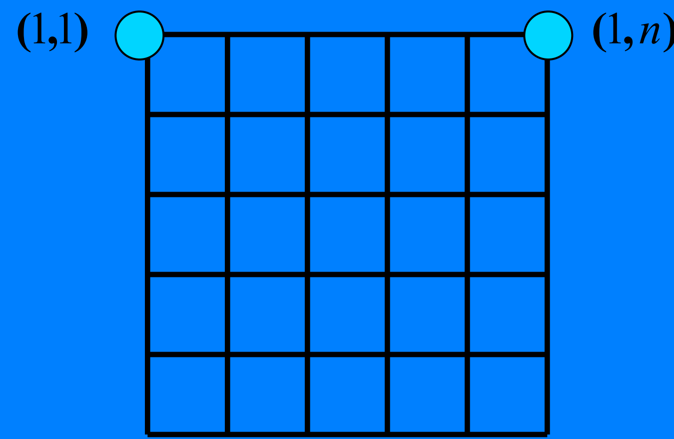
Let

***MVS*(*G*, *v*)** = maximum number of vertices that can be saved in *G* if fire starts at *v*.

Saving Vertices in Finite Grids G

$G = P_n \times P_n \qquad V(G) = \{(a,b) \mid 1 \le a, b \le n\}$

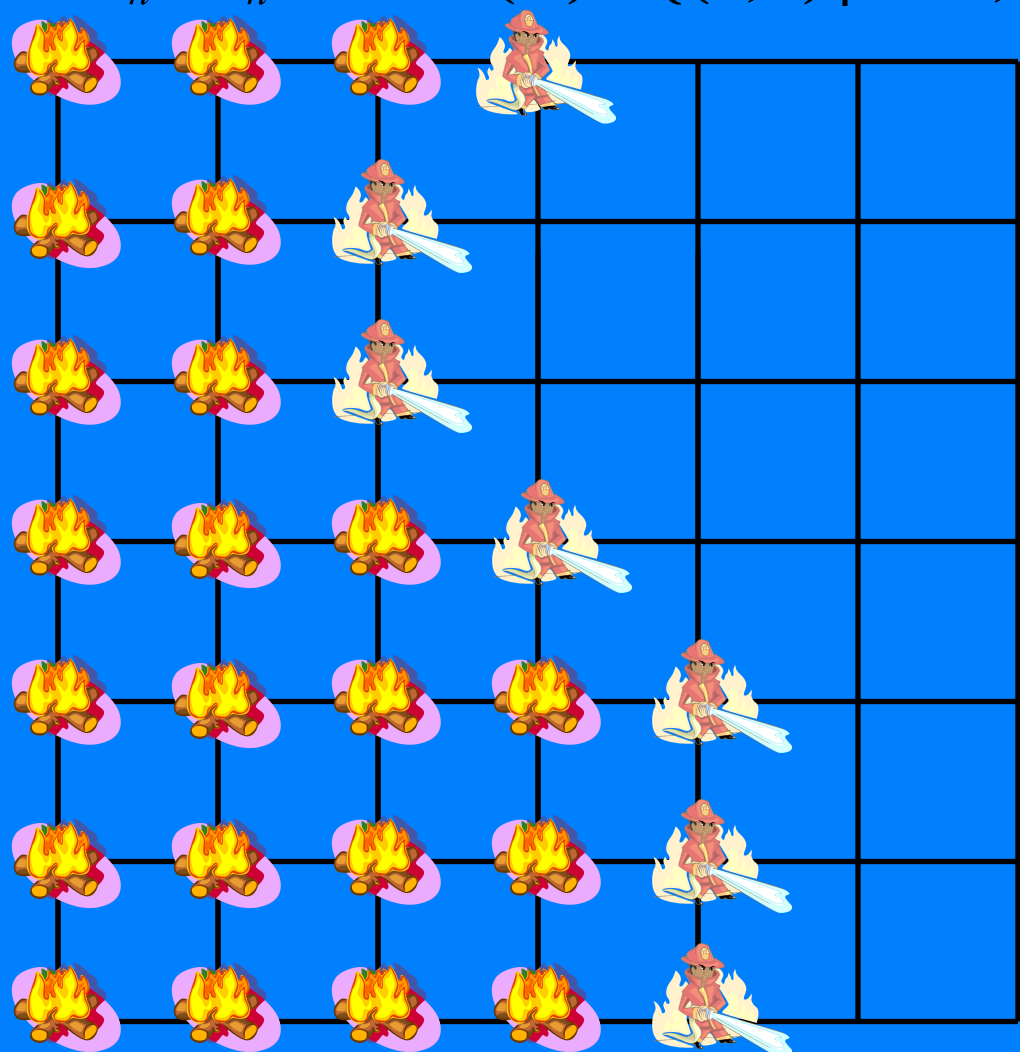




$$MVS(P_n \times P_n, (a,b)) \ge n(n-b) - (a-1)(n-a) \qquad 1 \le b \le a \le \left\lceil \frac{n}{2} \right\rceil$$

Saving Vertices in Finite Grids G

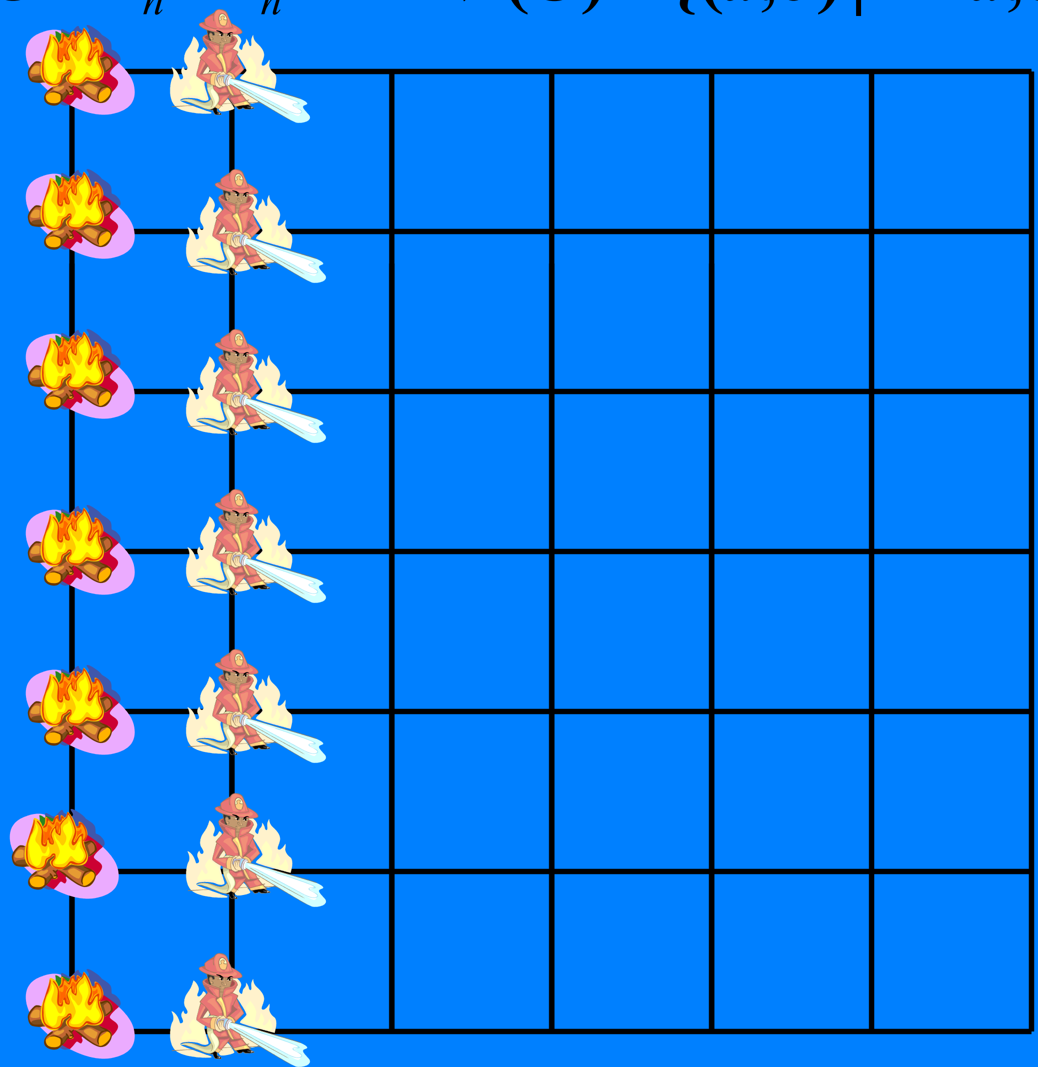
$G = P_n \times P_n$ $\quad V(G) = \{(a,b) \mid 1 \le a,b \le n\}$



$MVS(P_n \times P_n,(a,b)) \ge n(n-b)-(a-1)(n-a)$ $\quad 1 \le b \le a \le \left\lceil \frac{n}{2} \right\rceil$

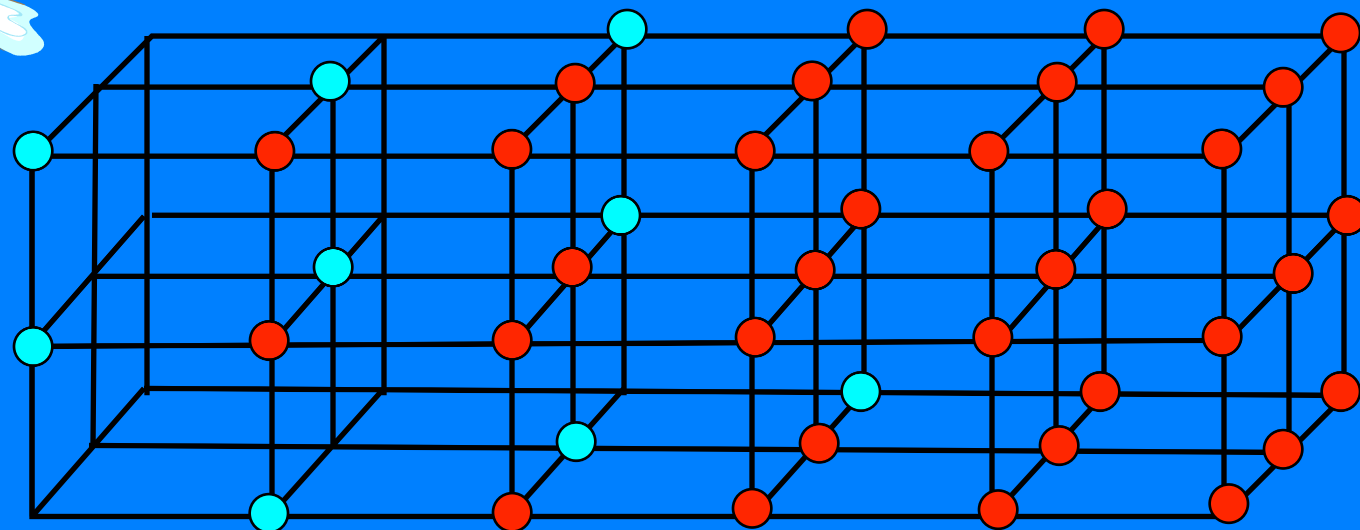
Saving Vertices in Finite Grids G

$G = P_n \times P_n \qquad V(G) = \{(a,b) \mid 1 \le a,b \le n\}$



$$MVS(P_n \times P_n, (1,1)) = n(n-1) = n^2 - n$$

Saving Vertices in $P_l \times P_m \times P_n$



$$MVS(P_3 \times P_3 \times P_6, (1,1,1)) = 21$$

$$MVS(P_3 \times P_3 \times P_n, (1,1,1)) = 9n - 33, \quad n \geq 6$$

Saving Vertices in P_n x P_n x P_n

Conjecture (Moeller and Wang):

$\lim_{n \to \infty} MVS(P_n x P_n x P_n, v)/n^3 = 0$ for all v

Algorithmic and Complexity Matters

FIREFIGHTER:

Instance: A rooted graph (G,v) and an integer $p \geq 1$.

Question: Is MVS(G,v) $\geq$ p? That is, is there a finite sequence $d_1, d_2, \ldots, d_t$ of vertices of G such that if the fire breaks out at v, then,

1. vertex d_i is neither burning nor defended at time i

2. at time t, no undefended vertex is next to a burning vertex

3. at least p vertices are saved at the end of time t.

Algorithmic and Complexity Matters

Theorem (MacGillivray and Wang, 2003): FIREFIGHTER is NP-complete.

Theorem (Finbow, Kind, MacGillivray, Wang (2007): FIREFIGHTER is NP-complete even if restricted to trees with maximum degree 3.

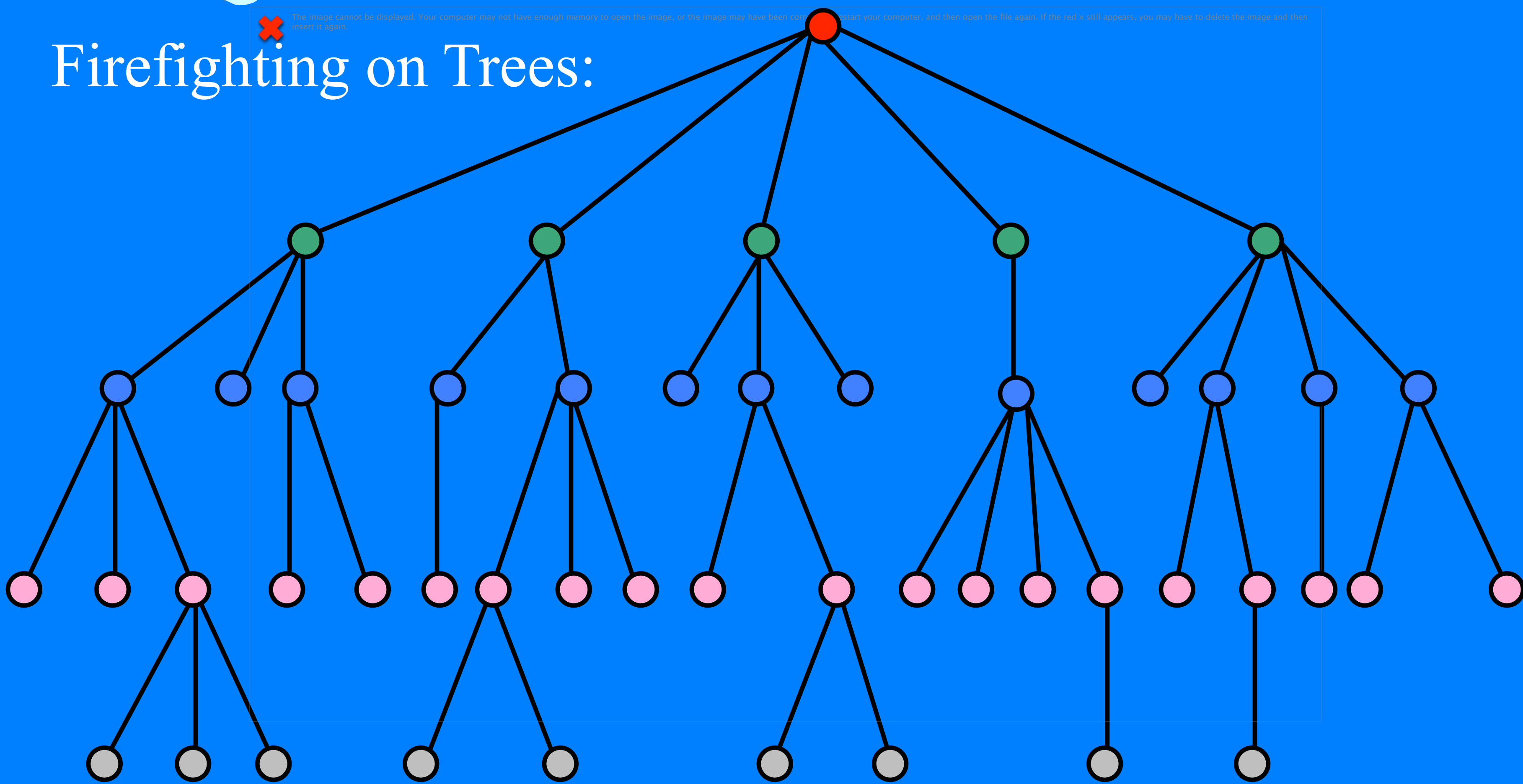
Theorem (Finbow, Kind, MacGillivray, Wang (2007): The problem is solvable in polynomial time for graphs of maximum degree 3 if the fire starts at a vertex of degree 2.

Algorithmic and Complexity Matters

Theorem (King and MacGivillray, 2010):
FIREFIGHTER is NP-complete for cubic graphs.

Algorithmic and Complexity Matters

Firefighting on Trees:



Algorithmic and Complexity Matters

Greedy algorithm:

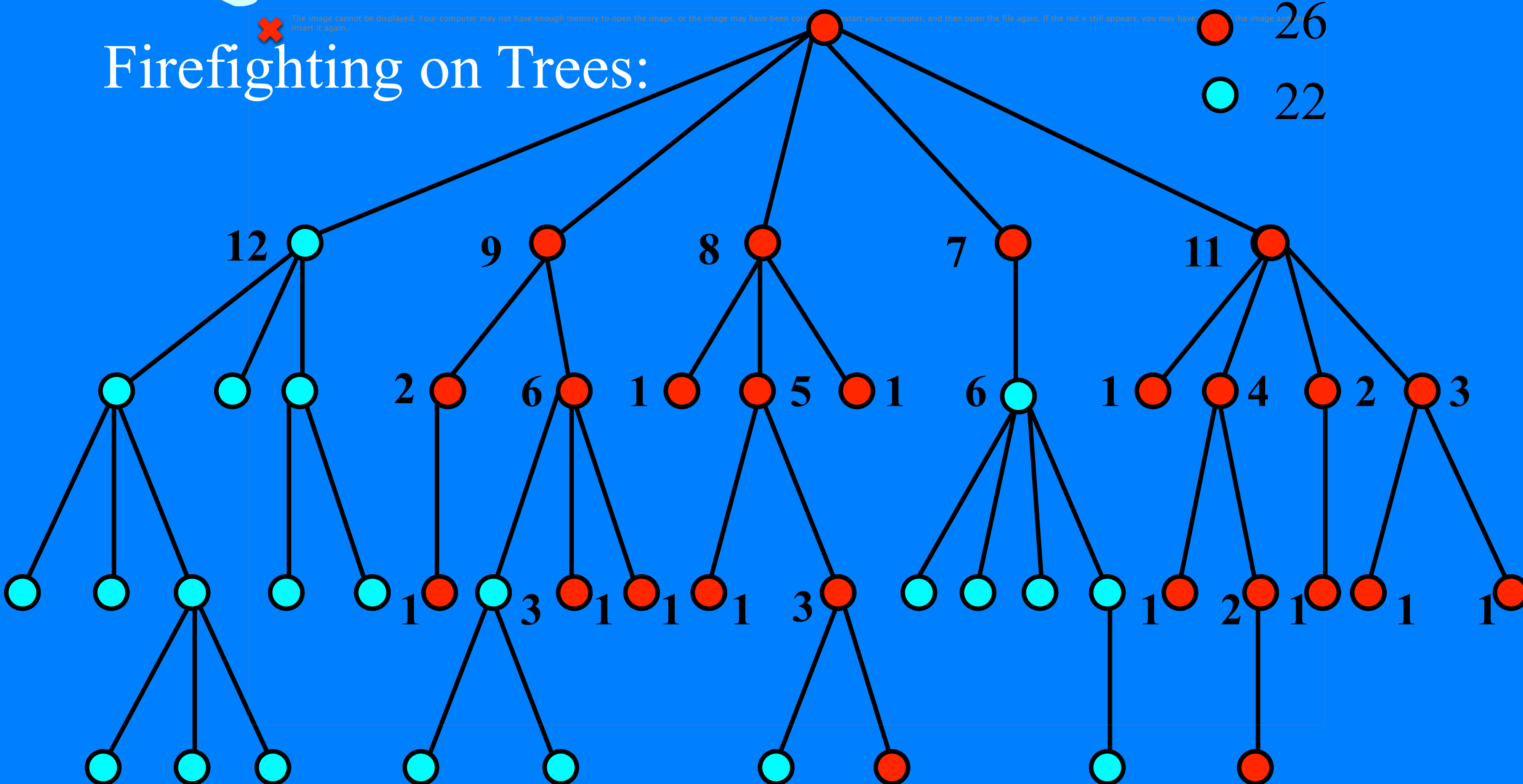
For each v in V(T), define

weight (v) = number descendants of v + 1

Algorithm: At each time step, place firefighter at vertex that has not been saved such that weight (*v*) is maximized.

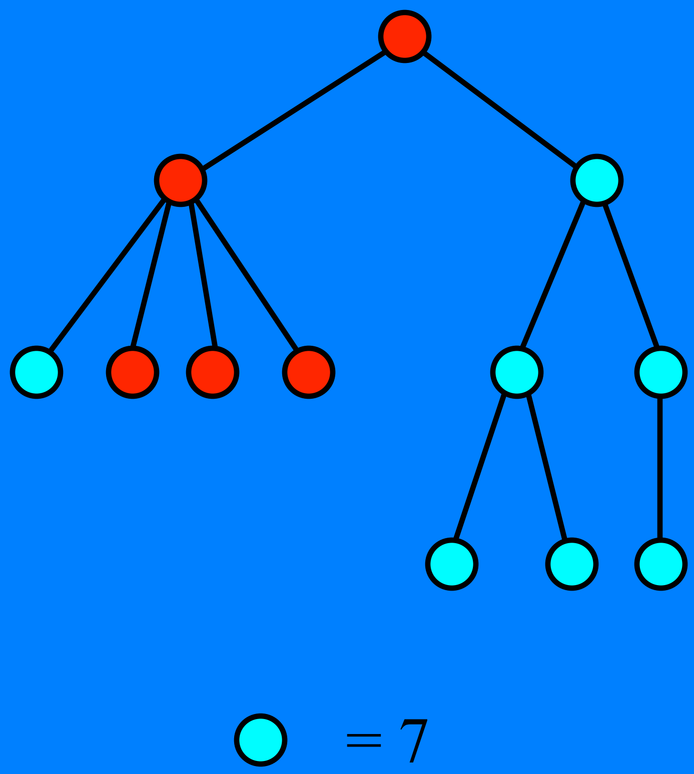
Algorithmic and Complexity Matters

Firefighting on Trees:

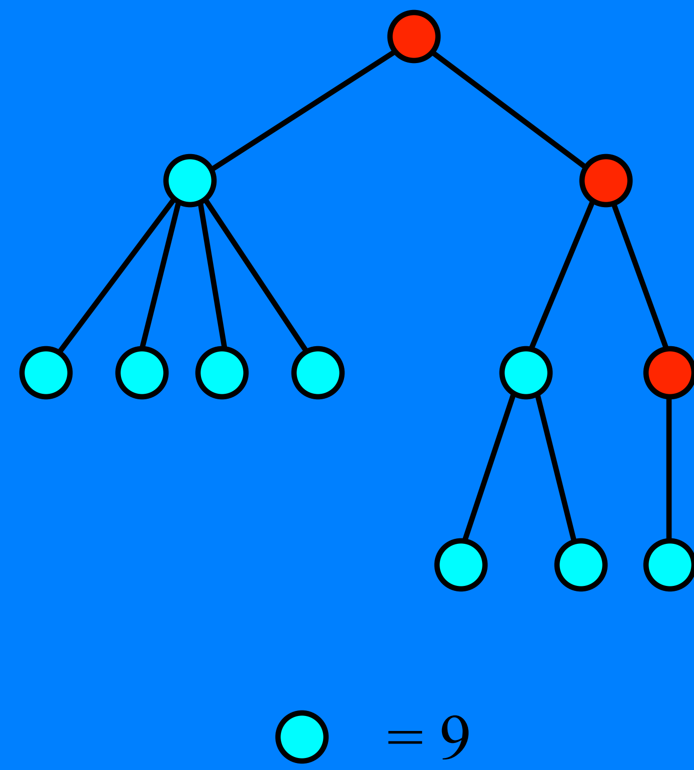


Algorithmic and Complexity Matters

Greedy



Optimal



Algorithmic and Complexity Matters

Theorem (Hartnell and Li, 2000): For any tree with one fire starting at the root and one firefighter to be deployed per time step, the greedy algorithm always saves more than ½ of the vertices that any algorithm saves.

Algorithmic and Complexity Matters

<u>Theorem (Finbow and MacGillivray 2009):</u>
The FireFighter problem is solvable in polynomial time for caterpillars and for trees of maximum degree 3 where the root has degree 2. (This includes binary trees.)

Would Graph Theory help with a deliberate outbreak of Anthrax?

What about a deliberate release of smallpox?

Similar approaches using mathematical models have proven useful in public health and many other fields, to:

- make policy
- plan operations
- analyze risk
- compare interventions
- identify the cause of observed events

More Realistic Models

Many oversimplifications in both of our models.
For instance:

measles

- What if you stay infected (burning) only a certain number of days?
- What if you are not necessarily infective for the first few days you are sick?
- What if your threshold k for changes from ○ to ● (uninfected to infected) changes depending upon how long you have been uninfected?

More Realistic Models

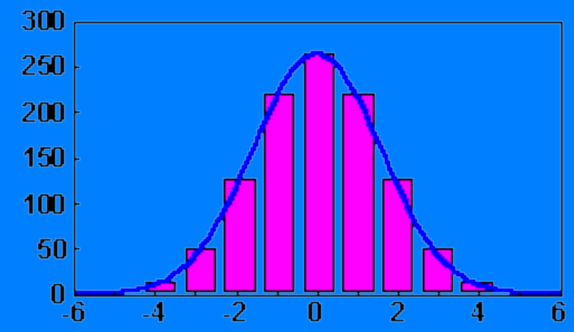
Consider an irreversible process in which you stay in the infected state (state ●) for d time periods after entering it and then go back to the uninfected state (state ●).

Consider an irreversible k-threshold process in which we vaccinate a person in state ● once k-1 neighbors are infected (in state ●).

Etc. – experiment with a variety of assumptions

More Realistic Models

Our models are *deterministic*. How do *probabilities* enter?



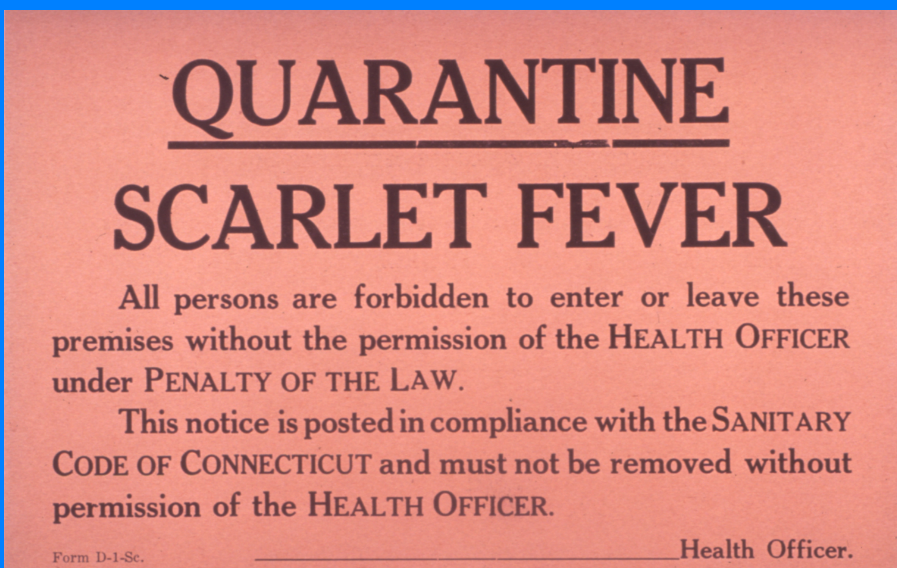
•What if you only get infected with a certain probability if you meet an infected person?

•What if vaccines only work with a certain probability?

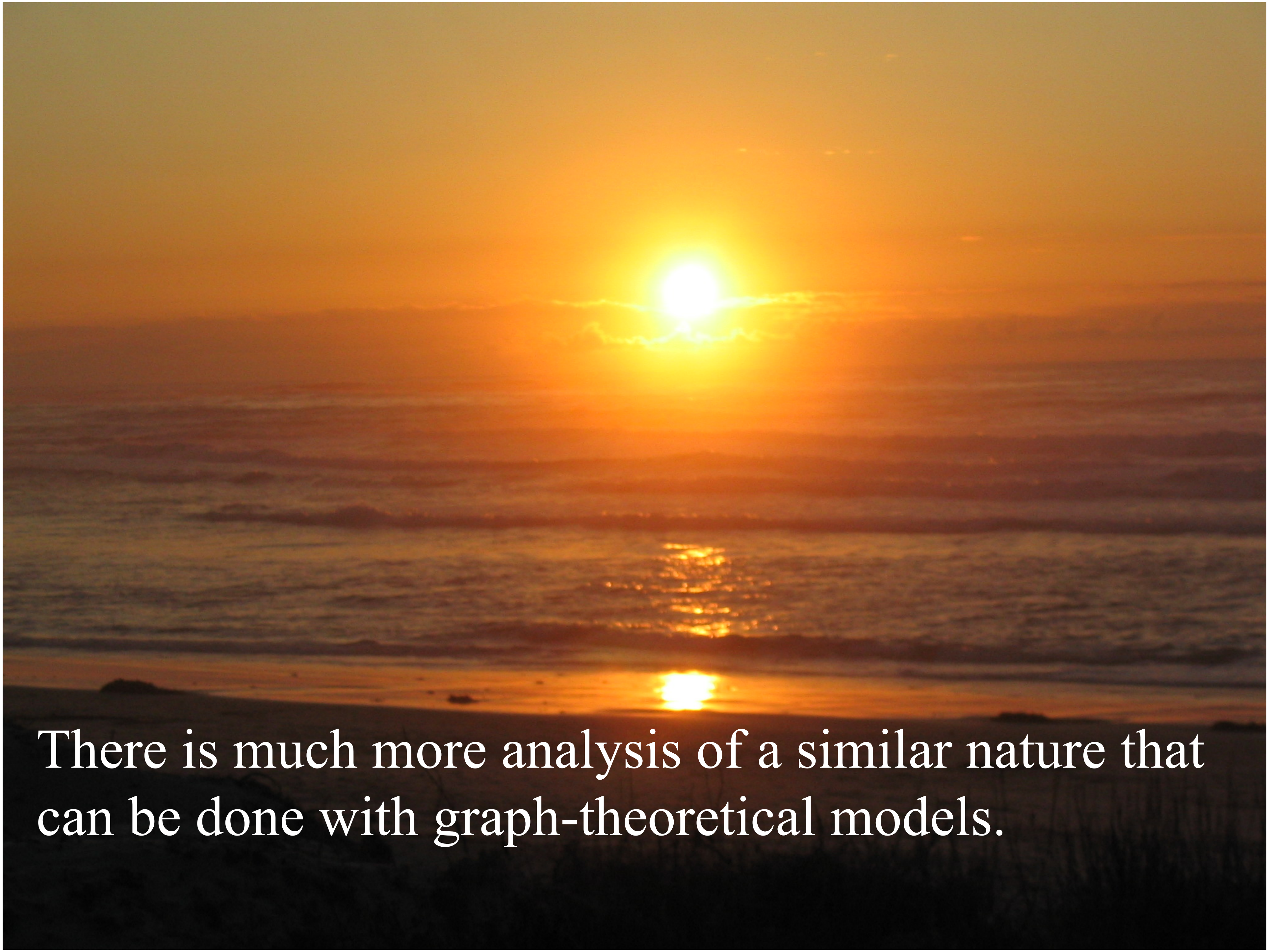
•What if the amount of time you remain infective exhibits a probability distribution?

Other Questions

Can you use graph-theoretical models to analyze the effect of different quarantine strategies?



Don't forget diseases of plants.


There is much more analysis of a similar nature that can be done with graph-theoretical models.