# Mathematical Models of Disease Dynamics

Winfried Just
Department of Mathematics
Ohio University

June 9, 2014

# What processes are we modeling?

We are interested in diseases that are triggered when pathogens aka infectious agents such as viruses or bacteria (called microparasites) enter the organism of a host (human, animal, plant).

We are not interested here in the actual changes that the disease causes in the organism of the host, or how the pathogens multiply within the host. We only care about how the disease spreads between hosts of a given population.

In this lecture we will focus on diseases whose transmission requires direct contact (of a certain type) between hosts, as opposed to diseases that require a third type of organisms, called vectors for transmission between hosts (mosquitoes in the case of malaria), or diseases where the pathogens are taken up from the shared environment of the hosts (drinking water in the case of cholera).

# Which questions are we trying to answer?

- If some disease pathogens are introduced into a population of hosts that have not previously been exposed to the disease, will a major outbreak aka an epidemic result? That is, should we expect that a significant fraction of hosts in the population will eventually get infected?
- If an epidemic does result, what proportion of hosts will be infected? The proportion of hosts that will be infected at some (not necessarily the same) time during the epidemic is called the final size (of the outbreak).
- What control measures are most effective in either preventing an epidemic or reducing the final size as much as possible?
- Possible control measures include vaccination, quarantine, culling, (for animal and plant diseases) or behavior modifications (for human diseases).

# When is a mathematical model good (enough)?

Our goal is to construct mathematical models that give us correct answers to the questions on the previous slide. For that, we need the model to:

- Give us answers in the first place. Thus the model needs to be simple enough to be tractable either by mathematical analysis or computer simulations.
- Be sufficiently realistic. It needs to take into account sufficiently many biological details that influence the dynamics so as to make reasonably correct predictions.
- Be based on data that we actually can collect.

This may be too much to ask for. In practice, we may not know whether a given model is sufficiently realistic.

# What happens to a host during a disease?

Since we only aim at modeling the dynamics between hosts, we only make the following assumptions about the disease that gets transmitted to host number $i$ at some time $T_E^i$:

- All (potential) hosts start out being susceptible to the disease (at times $t < T_E^i$).
- At all times $T_E^i \leq t < T_I^i$ the host will not (yet) be able to infect others.
- At all times $t$ with $T_I^i \leq t < T_R^i \leq \infty$ the host will be infectious, that is, will transmit the disease with positive probability during contacts with susceptible hosts.
- At all times $t \geq T_R^i$ the host will neither be infectious nor susceptible.

**How are the onset and cessation of symptoms related to these times? What happens at time $T_R^i$?**

**What kind of diseases do not satisfy the above assumptions?**

# The building blocks of disease models: compartments

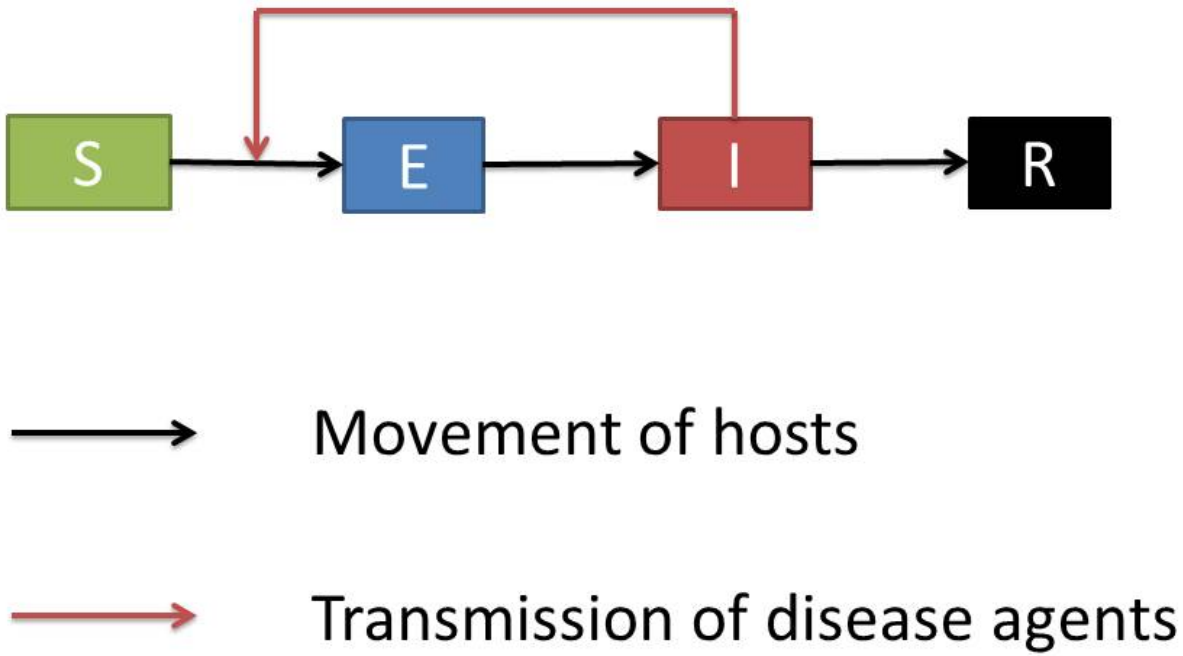
Let's summarize:

- $T_E^i$ is time of exposure of host number $i$ .
- $T_I^i$ is time of onset of infectiousness of host number $i$ .
- $T_R^i$ is time of removal of host number $i$, which means the host number $i$ either dies from the disease or acquires permanent immunity at time $T_R^i$.

This suggests a partition of host population at time $t$ into up to four compartments:
$S$ comprises all susceptible hosts (for which $t < T_E^i$),
$E$ comprises all exposed hosts with $T_E^i \leq t < T_I^i$),
$I$ comprises all infectious hosts with $T_I^i \leq t < T_R^i$, and
$R$ comprises all removed hosts with $T_R^i \leq t$.

# *SEIR*-models

If all four compartments are considered, we get an *SEIR*-model.



# Latency periods

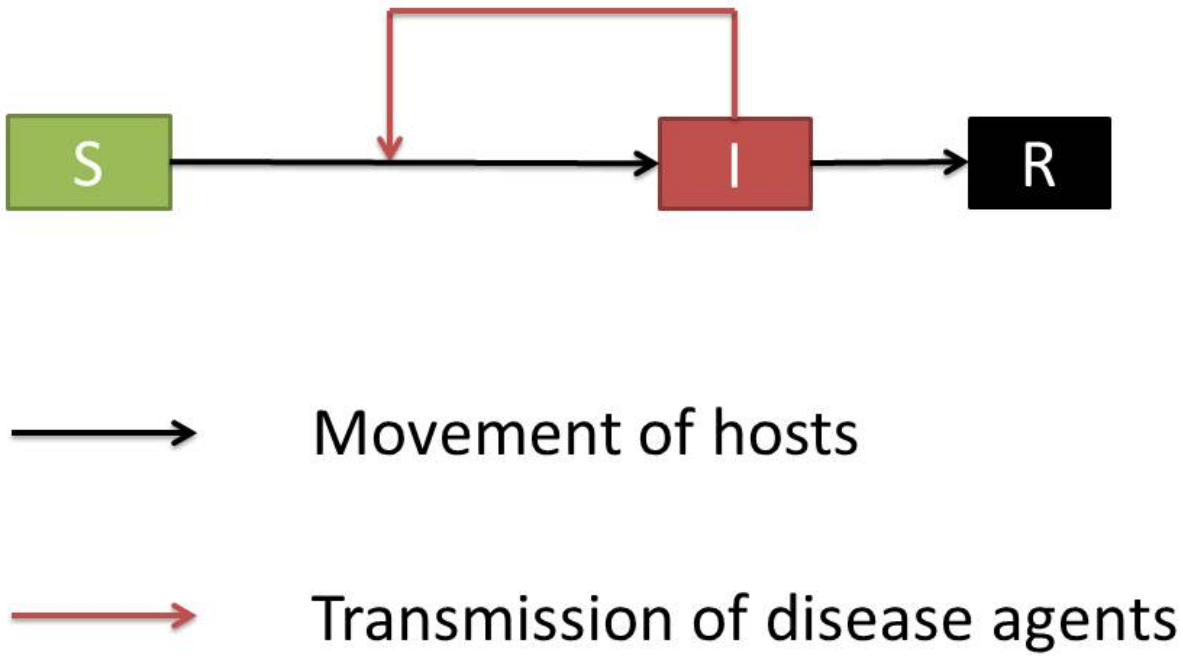
The time interval $[T_E^i, T_I^i)$ between exposure and onset of infectiousness is called the latency period. This may be different from the time interval between exposure and onset of symptoms, which is called the incubation period.

In many diseases, the length of the latency period is very short relative to the duration of infectiousness $T_i^R - T_I^i$.

**If the length of the latency period is very short relative to the duration of infectiousness, could we perhaps simplify the model? If so, how would we go about it?**

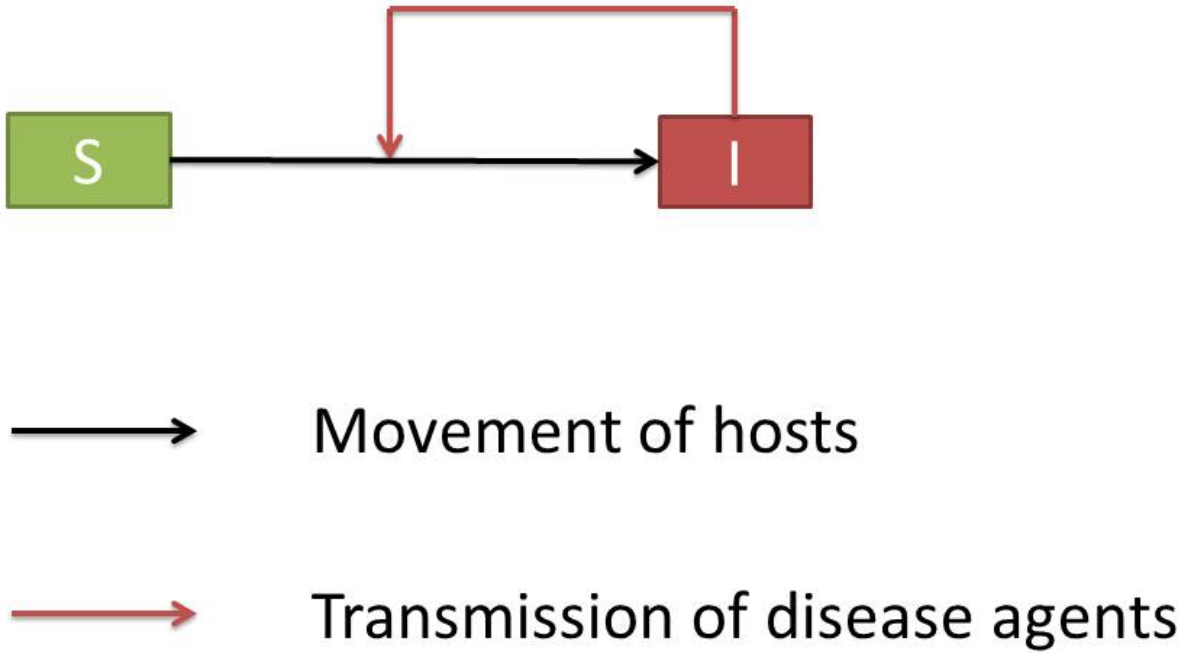
# *SIR*-models

The simplifying assumption $T_E^i = T_I^i$ eliminates the $E$-compartment and we get an *SIR*-model.



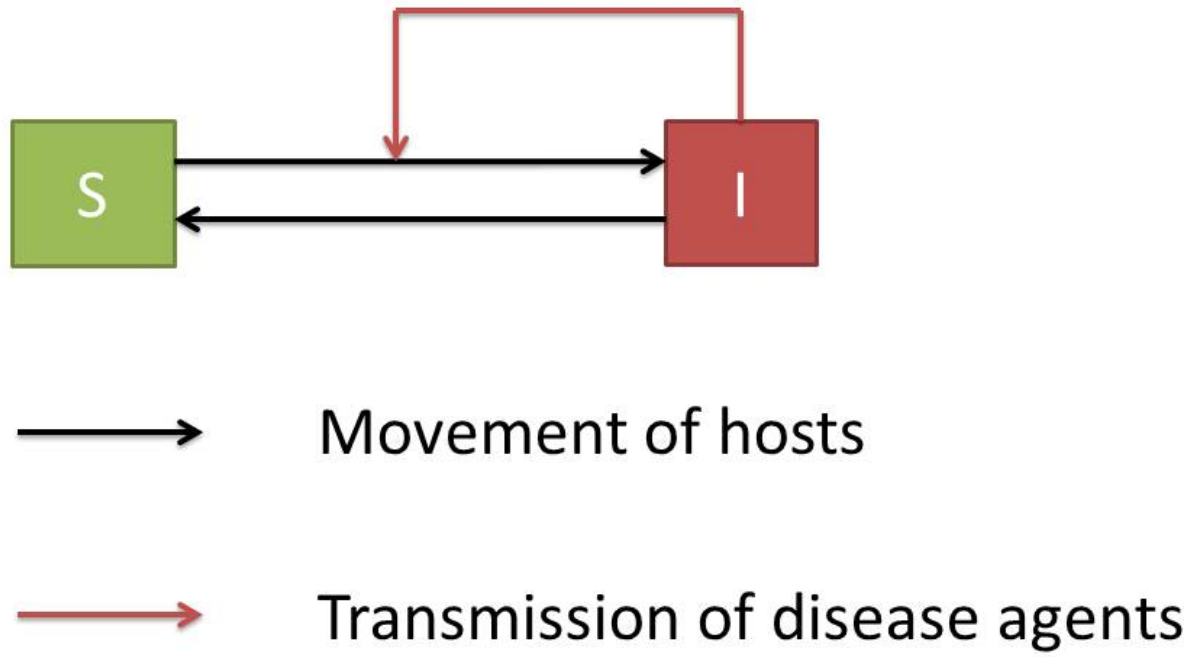
# *SI*-models

If in addition $T_R^i = \infty$, then the $R$-compartment becomes redundant and we get an *SI*-model.



# *SIS*-models

If we assume instead that at time $T_R^i$ hosts recover and become susceptible to reinfection instead of acquiring immunity, then we get an *SIS*-model.



# Four basic types of compartment models

$S$ comprises all susceptible hosts (for which $t < T_E^i$),
$E$ comprises all exposed hosts with $T_E^i \leq t < T_I^i$),
$I$ comprises all infectious hosts with $T_I^i \leq t < T_R^i$, and
$R$ comprises all removed hosts with $T_R^i \leq t$.
The times $T_E^i, T_I^i, T_R^i$ are specific to each host. Membership in the compartments changes over time, and we can think of hosts moving (but not in space!) from $S$ to $E$ to $I$ to $R$.

- The above is called an *SEIR*-model.
- The simplifying assumption $T_E^i = T_I^i$ eliminates the $E$-compartment and gives an *SIR*-model.
- If in addition $T_R^i = \infty$, then the $R$-compartment becomes redundant and we get an *SI*-model.
- If we assume instead that $T_E^i = T_I^i$ and at time $T_R^i$ hosts simply recover and become susceptible to reinfection instead of acquiring immunity, then we get an *SIS*-model.

**The choice of model type should depend on the disease. Can you think of other types of meaningful compartment models?**

# Major *vs.* minor outbreaks

Let us see whether we can translate our guiding questions into compartmentalese.

Assume an $SIR$-model. Consider a population of $N$ individuals, all initially in $S$, and assume $K$ of them become infected from outside sources at time $T_E$, where $K \ll N$.

For each time $t \geq 0$ consider the fraction $\frac{N - \#S(t)}{N}$.

The limit $F(K, N) = \lim_{t\to\infty} \frac{N - \#S(t)}{N}$ is the final size, *i.e.* fraction of hosts that don't escape infection.

If for fixed $K$ we have $\lim_{N\to\infty} F(K, N) = 0$, then the disease will affect only a negligible fraction of a large population. This signifies that all outbreaks are predicted to be minor.

If not, then there is danger of a major outbreak.

# Translating control measures into compartmentalese

**How would vaccination at time $T_v < T_E$ of a fraction $r$ of hosts translate into compartmentalese?**

As moving $rN$ hosts into $R$ at time $T_v$.

**How would culling at time $T_v < T_E$ of a fraction $r$ of an animal herd translate into compartmentalese?**

Again as moving $rN$ hosts into $R$ at time $T_v$.

Compartmentalese seems to be a convenient language for us.

But some important aspects of reality may get lost in the translation.

# Let's build and explore some models!

Disease transmission can be modeled within various mathematical frameworks, including ODE, PDE, difference equation, and stochastic process models.

If the population size $N$ is sufficiently large, then one can try to build ODE models with variables $S, I, R$ (if we do have removed individuals), and perhaps $E$. These variables could either represent the proportions of hosts in the respective compartments, or their numbers.

In order to make the use of derivatives somewhat respectable in this application, one can think of population size being expressed in units of a thousand or a million individuals so that at least some fractional values of the variables make sense.

# Simplifying assumption of our ODE models

- We will moreover assume that the population size $N$ is fixed, which ignores demographics, that is, births, deaths from unrelated causes, immigration and emigration.
- We will ignore the $E$-compartment and assume that the onset of infectiousness coincides with the time of exposure.
- In the resulting ODE models the state of the population is represented by three variables $S, I$, and $R$.
- All ODE models implicitly assume that the future course of an outbreak depends only on the current values of the variables that represent counts or proportions of hosts in the compartments.

One always needs to carefully consider to what extent the assumptions of the model might distort its predictions. We will examine the last of these assumptions in some detail later in this talk and in the lab.

It remains to derive expressions for $\frac{dS}{dt}$, $\frac{dI}{dt}$, and $\frac{dR}{dt}$ (if applicable).

# Building an ODE model for *SI*

The only variables are $S$ and $I$. New infections of susceptible individuals are the only mechanism by which the variables change.

Consider an (average) susceptible host. The "rate at which this host acquires the infection" is called the force of infection and will be denoted here by $f$.

In ODE models we assume that $f(t)$ is proportional to the number of currently infectious hosts, that is, $f(t) = \beta I(t)$ for some constant $\beta$. This is aversion of the uniform mixing assumption.

The rate at which hosts (plural) move out of the $S$-compartment and into the $I$-compartment is equal to $f(t)S(t)$. This gives

$$\frac{dS}{dt}(t) = -\beta S(t)I(t); \qquad \frac{dI}{dt}(t) = \beta S(t)I(t).$$

The constant $\beta$ needs to be estimated from data. It may or may not depend on $N$.

# What does this *SI*-model predict?

Consider the *SI*-model

$\frac{dS}{dt} = -\beta IS; \qquad \frac{dI}{dt} = \beta IS.$

Then $\frac{d(S+I)}{dt} = 0.$ **What does this tell us?**

We can eliminate one of the variables; let us set $S = I - N$:

$\frac{dI}{dt} = \beta I(N - I).$

**Looks familiar?**

The right-hand side is positive for all $I \in (0, N)$, which implies that all trajectories that start with $I(0) \in (0, N]$ will approach the steady state in which all hosts are infected. The disease-free equilibrium at which no hosts are infectious is unstable.

**What does this model predict about the final size of the outbreak?**

**Is this prediction realistic? If not, how should we modify the model?**

# An ODE version of the *SIS*-model

The only variables are $S$ and $I$. The rate of new infections is expressed as in the $SI$ model, but now we also need to consider the rate at which infectious hosts recover and move back into the $S$-compartment.

This rate is assumed proportional to $I(t)$. We obtain:

$$\frac{dS}{dt} = -\beta IS + \alpha I;$$
$$\frac{dI}{dt} = \beta IS - \alpha I.$$

Again, the constants $\alpha$ and $\beta$ need to be estimated from data.

# What does this *SIS* model predict?

$\frac{dS}{dt} = -\beta IS + \alpha I;$
$\frac{dI}{dt} = \beta IS - \alpha I.$

Since $\frac{dS+I}{dt} = 0$, by letting $S = N - I$, this simplifies to:

$\frac{dI}{dt} = \beta I(N - I) - \alpha I = \beta I(N - \frac{\alpha}{\beta} - I).$

**Looks familiar?** This is a logistic growth model!

It follows that the model has a disease-free equilibrium $I^* = 0$ and an endemic equilibrium $I^{**} = N - \frac{\alpha}{\beta}$.

If $R_0 := \frac{\beta N}{\alpha} < 1$, then the disease-free equilibrium is the only biologically meaningful one and is globally stable; if $R_0 > 1$, all trajectories that start with $I(0) \in (0, N]$ converge to the endemic equilibrium.

It seems that the number that we labeled $R_0$ is very important. We will explore its precise meaning later in this talk.

# What happens if $N$ increases?

Consider a very large population size $N$. The location of the endemic equilibrium $I^{**} = N - \frac{\alpha}{\beta}$ indicates that in the long run, at all times every host except a lucky $\frac{\alpha}{\beta}$ few will be infected.

This seems weird. **Is this prediction correct and meaningful?**

Only if $\beta$ does not depend on $N$, that is, if transmission is strictly density-dependent.

If, for example, transmission is frequency-dependent, then $\beta = \frac{\beta'}{N}$ for some fixed $\beta'$. In this case we get:

$$I^{**} = N - \frac{\alpha N}{\beta'} = N(1 - \frac{\alpha}{\beta'}).$$

This means that the endemic equilibrium is predicted to exist only if $\beta' > \alpha$ and will comprise a fixed fraction of the population in this case.

# An ODE version of the *SIR* model

The variables are $S, I$ and $R$. The difference from the *SIS*-model is that upon recovery, hosts move into the $R$-compartment instead of back into the $S$-compartment. This gives:

$\frac{dS}{dt} = -\beta IS$; (as in the *SI*-model)
$\frac{dI}{dt} = \beta IS - \alpha I = I(\beta S - \alpha)$; (as in the *SIS*-model)
$\frac{dR}{dt} = \alpha I$.

If very few infectious hosts are introduced into a large and otherwise susceptible population, then $S \approx N$.

For $R_0 := \frac{\beta N}{\alpha} < 1$, we get $\frac{dI}{dt} < 1$, and the model predicts that $I$ will *decrease at all times* and the outbreak will be *minor.*

If $R_0 > 1$, then the model predicts that $I$ will *initially increase,* peak when $\beta S(t) - \alpha = 0$, and then decrease. Every outbreak will be a *major* one in this case. The final size will be a proportion that is strictly between 0 and 1.

Again, the number $R_0$ turns out to be very important.

# Predictions of this *SIR*-model about control measures

$\frac{dS}{dt} = -\beta IS; \quad \frac{dI}{dt} = \beta IS - \alpha I = I(\beta S - \alpha); \quad \frac{dR}{dt} = \alpha I.$

If we start from *any* initial condition with $I(0) > 0$, then $I$ will *decrease at all times* and the outbreak will be minor *as long as* $\beta S(0) - \alpha < 0$. If $\beta S(0) - \alpha > 0$, then $I$ will *initially increase*, and the outbreak will be major.

**How can we design effective control measures based on this prediction?**

Think about vaccinating a certain proportion $HIT$ of hosts prior to an outbreak.

This will change $S(0)$ from $\approx N$ into $(1 - HIT)N$.

If we set $HIT = 1 - \frac{1}{R_0} = 1 - \frac{\alpha}{\beta N}$, then

$\beta S(0) - \alpha = \beta(1 - HIT)N - \alpha = \beta \frac{\alpha}{\beta N} N - \alpha = 0.$

The fraction $HIT$ is called the herd immunity threshold. It is the minimum fraction of hosts that need to be vaccinated to prevent major outbreaks.

# Herd immunity and public policy

Herd immunity is a counterintuitive concept: By vaccinating enough hosts, we protect the entire population from major outbreaks, not only the hosts that got vaccinated. It may be a lot less costly to vaccinate only a fraction instead of all hosts at risk. But think about a severe, possibly life-threatening disease.

**Would it be ethical to vaccinate only a fraction *HIT* of hosts? Can you think of some practical obstacles to implementing such a policy?**

One important concern is that the calculation of *HIT* needs to be based on a mathematical model. If we want to base a serious public policy decision on this number, we need to make sure, as best as we can, that the assumptions of the model do not significantly distort the predictions about the value of *HIT*.

**How much can we trust the predictions of our ODE models?**

# The ODE version of the *SIR*-model revisited

$$\frac{dS}{dt} = -\beta IS; \quad \frac{dI}{dt} = \beta IS - \alpha I = I(\beta S - \alpha); \quad \frac{dR}{dt} = \alpha I.$$

If one or very few infectious hosts are introduced into a large and otherwise susceptible population, then for $R_0 := \frac{\beta N}{\alpha} < 1$, the model predicts that $I$ will decrease at all times and the outbreak will be minor. If $R_0 > 1$, then the model predicts that $I$ will initially increase and every outbreak will be a *major* one.

Also recall the underlying assumption of all ODE models that

" the future course of an outbreak depends only on the current values of the variables that represent counts or proportions of hosts in the compartments."

**Is there anything suspicious about these predictions and the underlying assumption?**

# The problem with deterministic models

Suppose one infectious host with a new strain of the flu, called an index case, is introduced into an otherwise susceptible population. You will not automatically catch the flu if you stand in a bus next to this person. If that person sneezes into your face though, you most likely will.

What is true of you is true of everybody else in the population. The index case may or may not infect anybody else and may or may not cause any kind of outbreak. Disease transmission is inherently a stochastic process.

ODE and other types of deterministic models ignore the stochastic nature of disease transmission.

# In what sense can we predict the final size?

On the previous slide we assumed $K$ out of $N$ individuals became initially infected and considered
$F(K,N) = \lim_{t\to\infty} \frac{N - \#S(t)}{N}$.
This limit exists all right in each repetition of the "experiment," but the outcome will differ between repeated runs of the "experiment."
To make the above definition of final size meaningful for a given compartmental model we need to treat both sides as expected values.

Even in this interpretation though, $F(K,N)$ will not in general be determined by $K$ and $N$ alone. **Does it matter which hosts are initially infected?**

ODE models ignore heterogeneities between individual hosts, as "the future course of an outbreak depends only on the current values of the variables that represent counts or proportions of hosts in the compartments."

# An example

Consider the spread of flu in a dorm with initially $K = 1$ student infected. If we ignore heterogeneities, then we can estimate the probability that a major outbreak will result if the infected student is "average." Look at two scenarios:

**Scenario 1:** The infected student caught it in a bar.

**Does the probability estimate based on the assumption of an "average" student appear to apply in this scenario, or does it appear to be too high or too low?**

**Scenario 2:** The infected student caught it from the janitor.

**Does the probability estimate based on the assumption of an "average" student appear to apply in this scenario, or does it appear to be too high or too low?**

# Advantages and disadvantages of ODE models

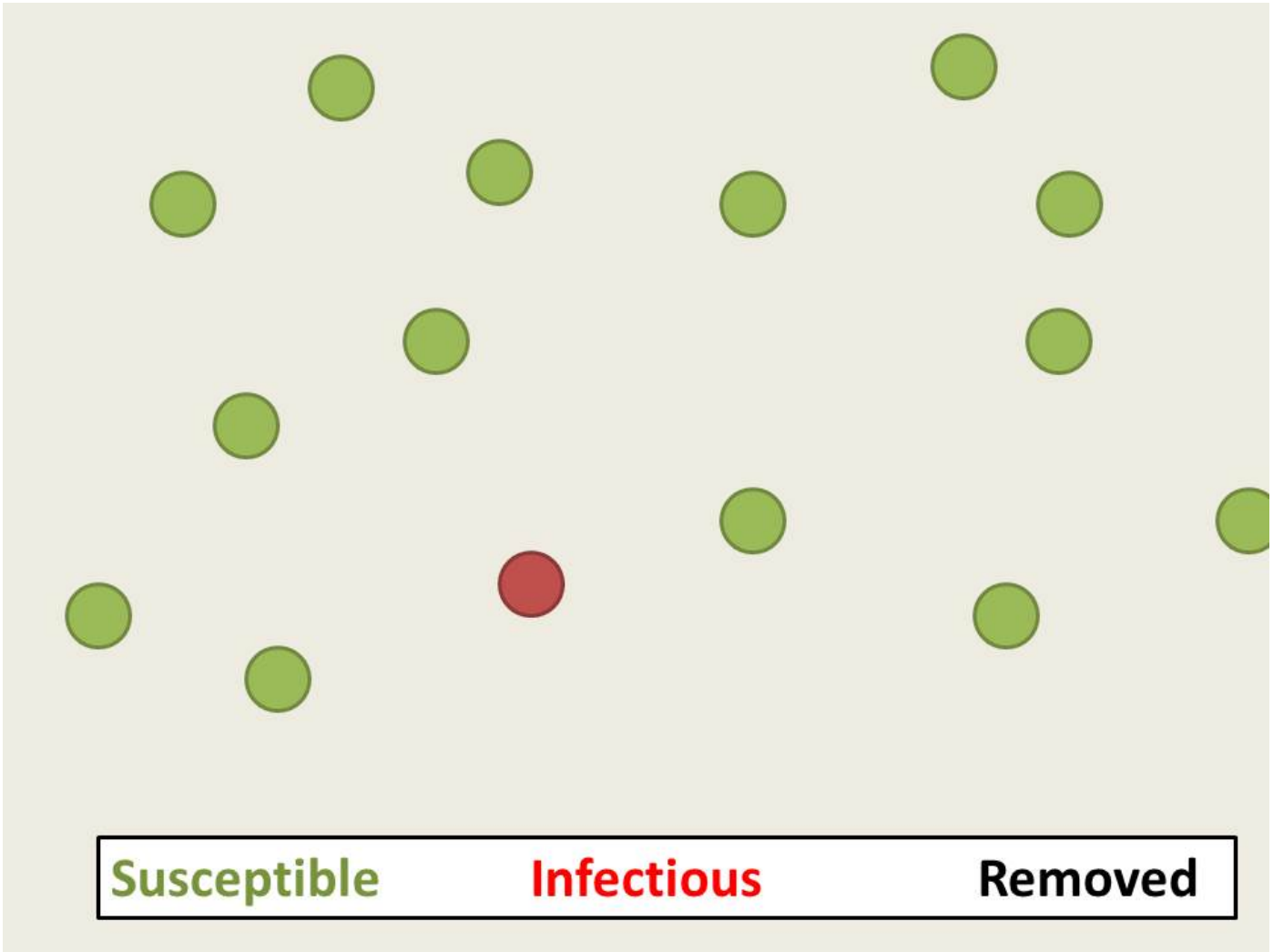
ODE models often make fairly realistic predictions. They have certain advantages and certain disadvantages.

- ODE models are relatively easy to study.
- They involve few variables and require estimation of very few parameters.
- ODE models ignore the stochastic nature of disease transmission.
- ODE models ignore heterogeneities between individual hosts.
- ODE models are based on the often unrealistic assumption of uniform mixing between individual hosts.

We will present a different type of models that can alleviate these three disadvantages to some extent. But first let us look at some examples of how outbreaks get started.

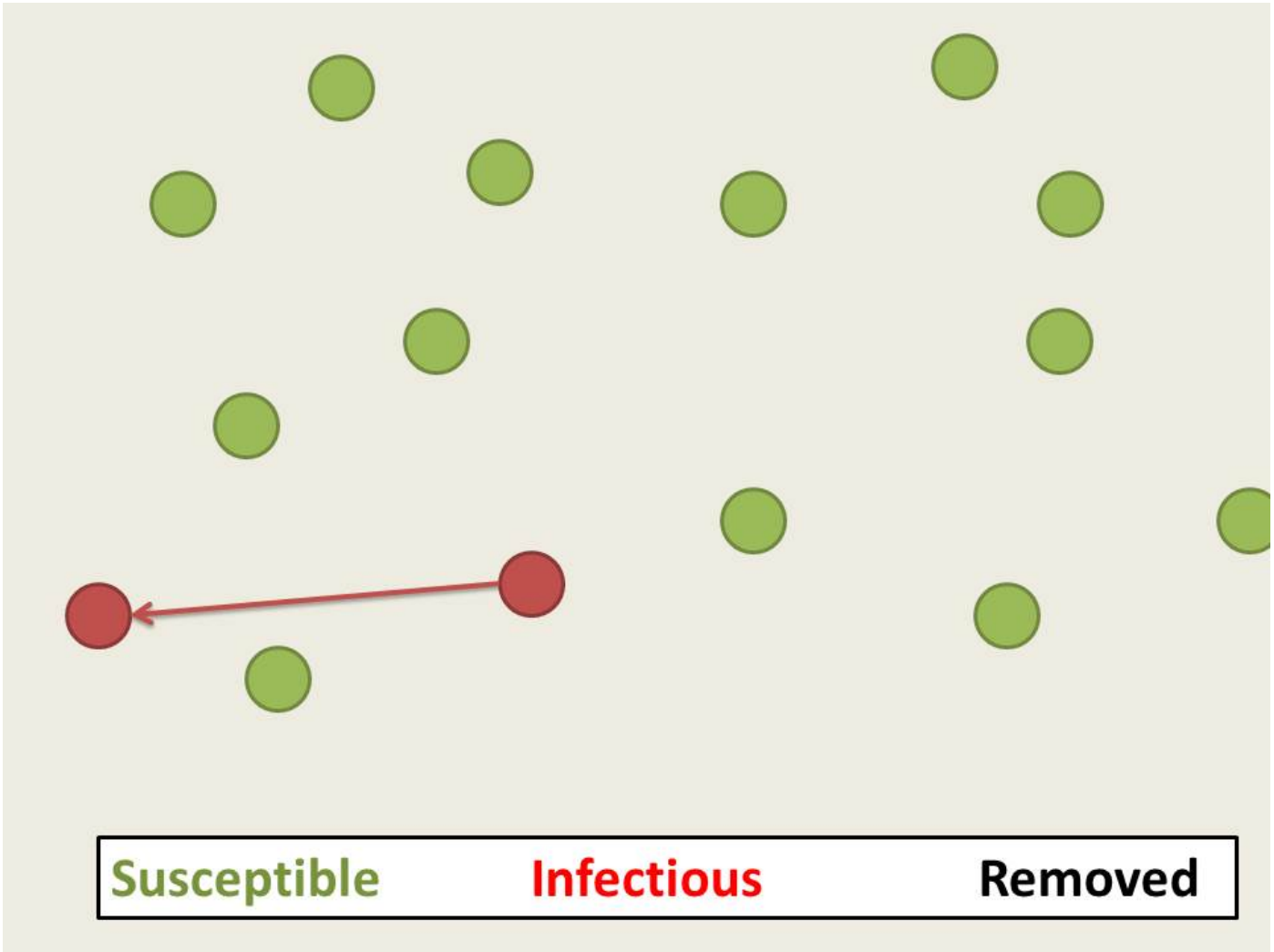
# How would an outbreak get started anyway? Example 1

A single infected host is introduced into a large population of susceptibles.



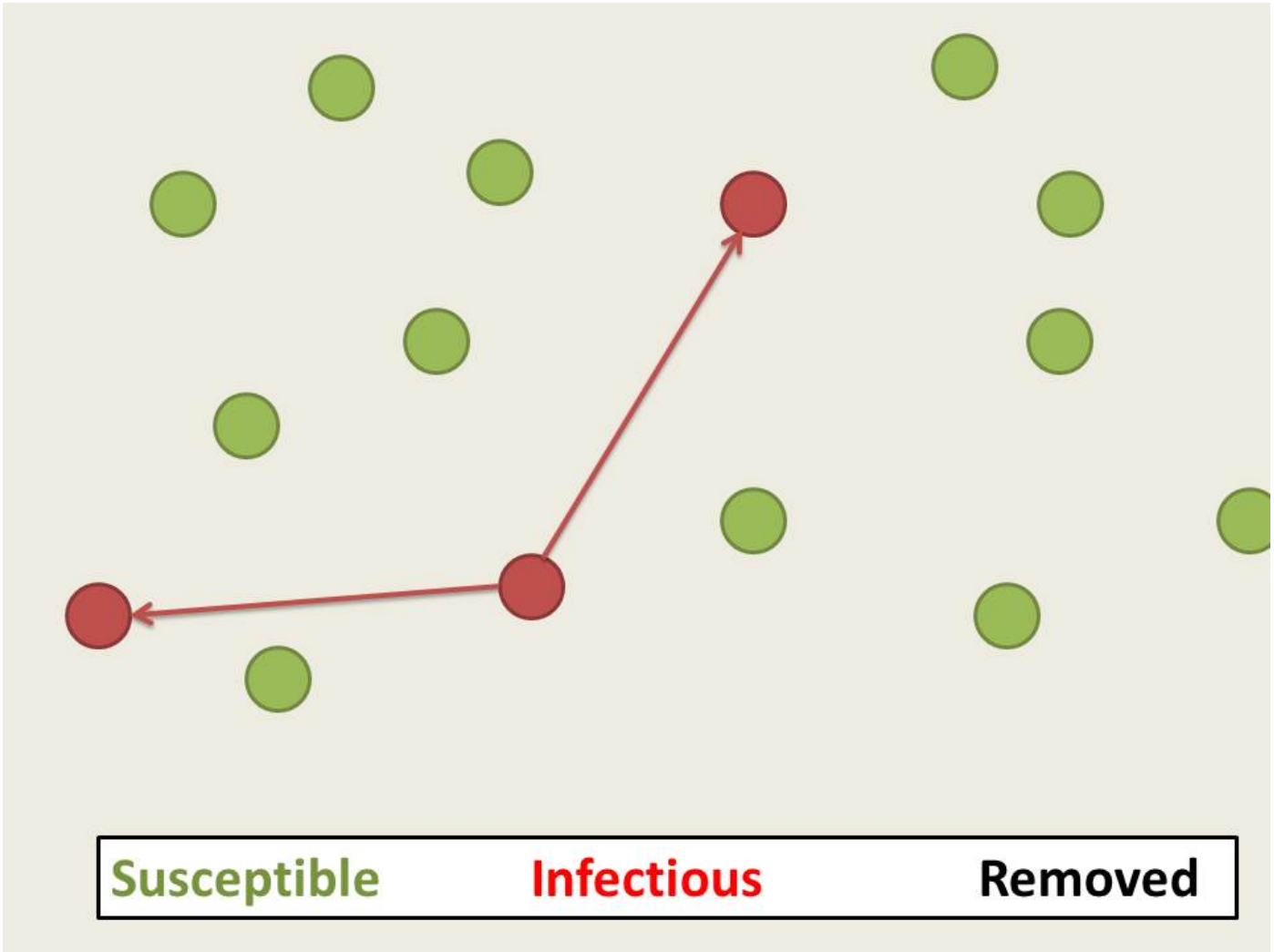
# How would an outbreak get started anyway? Example 1

A new infection occurs.



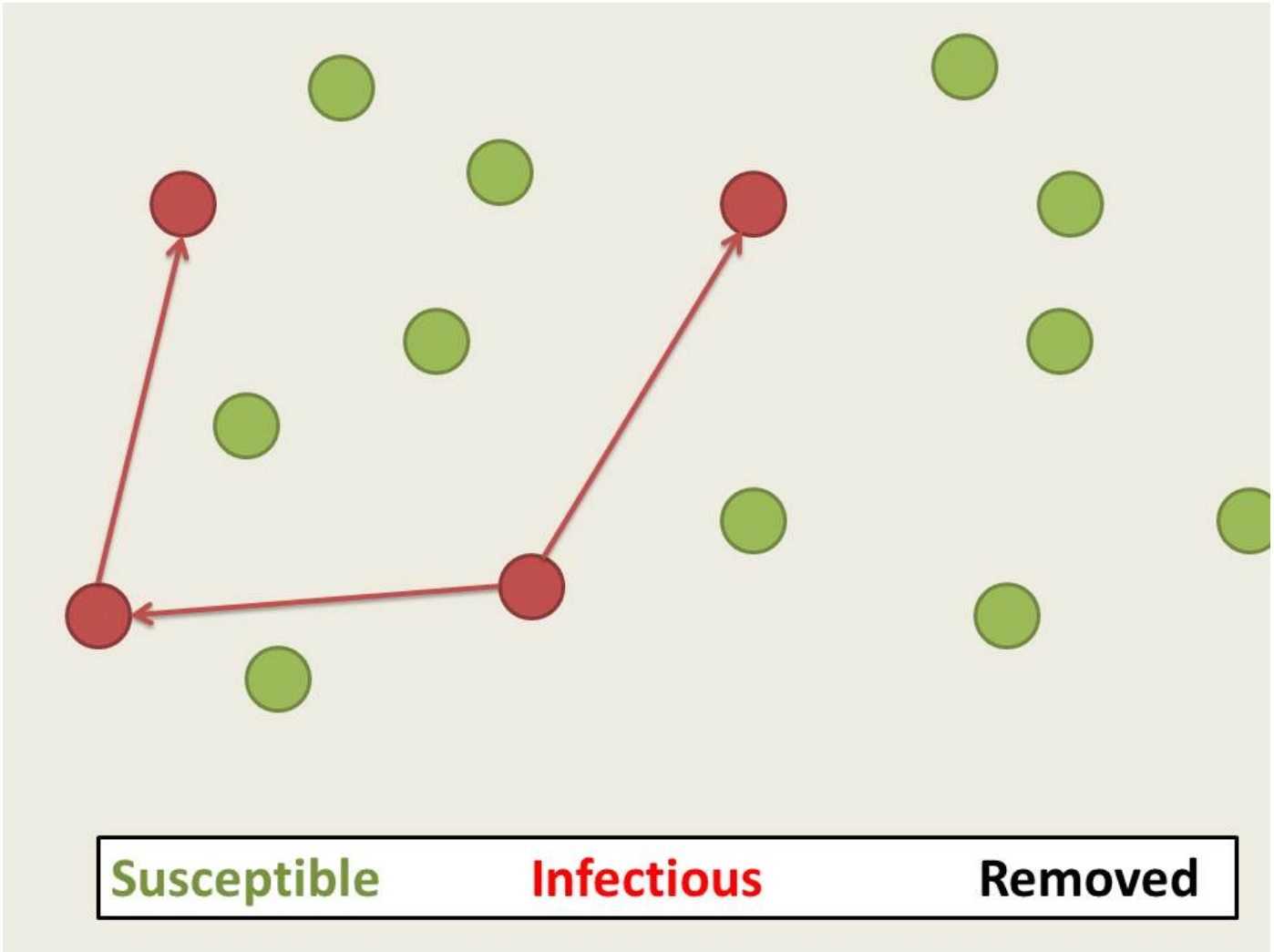
# How would an outbreak get started anyway? Example 1

A new infection occurs.



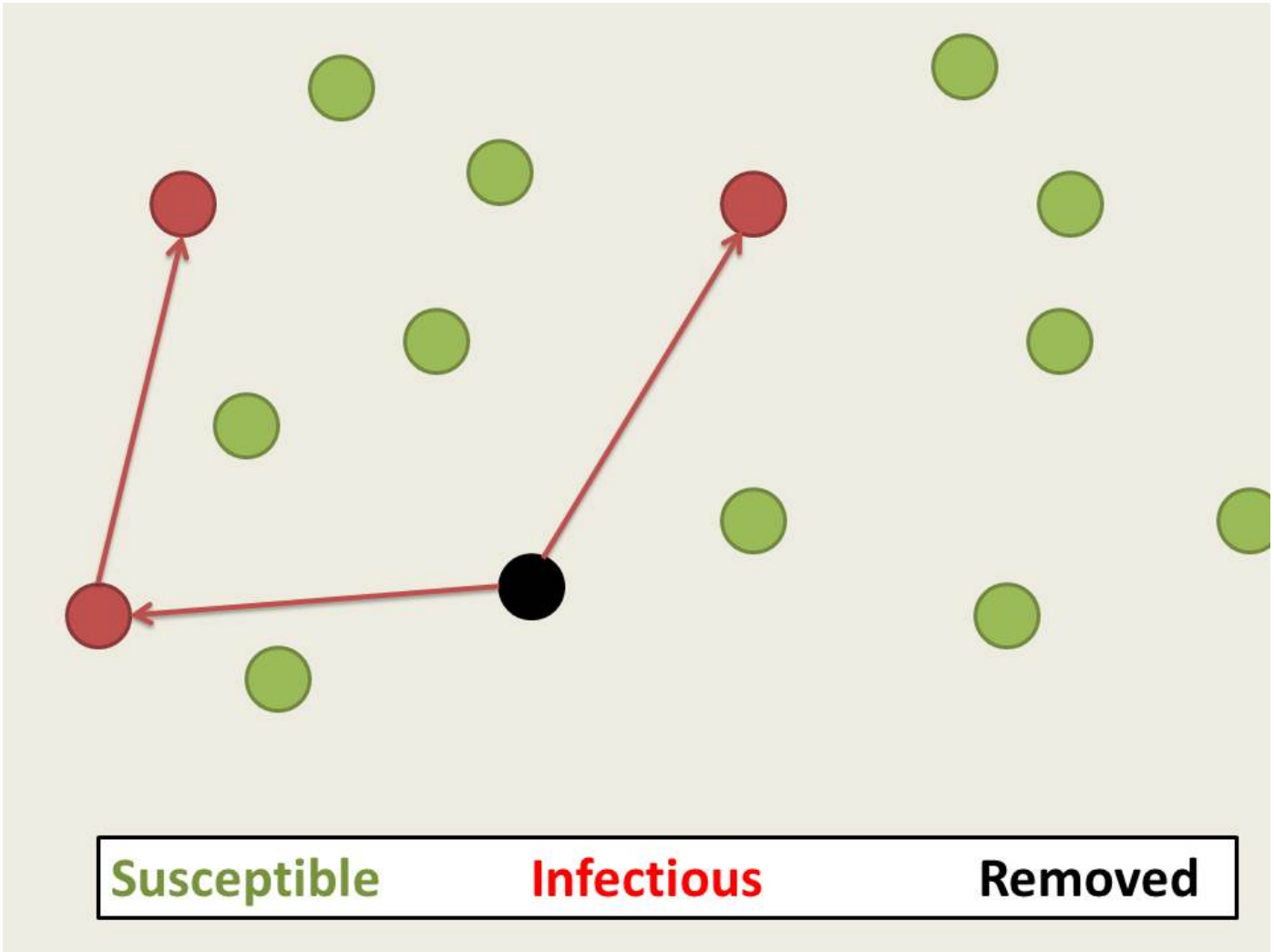
# How would an outbreak get started anyway? Example 1

A new infection occurs.



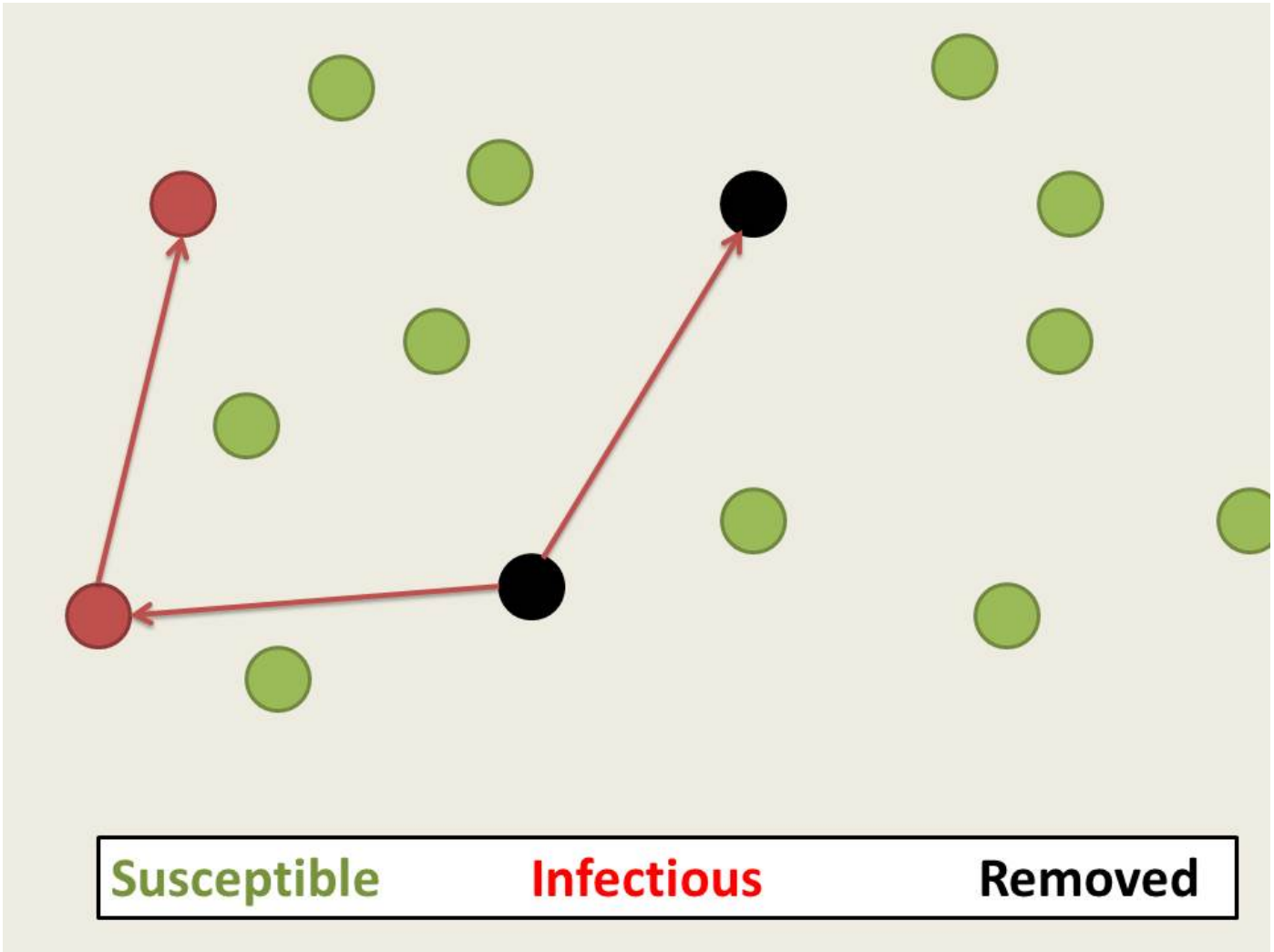
# How would an outbreak get started anyway? Example 1

An infectious host is removed.



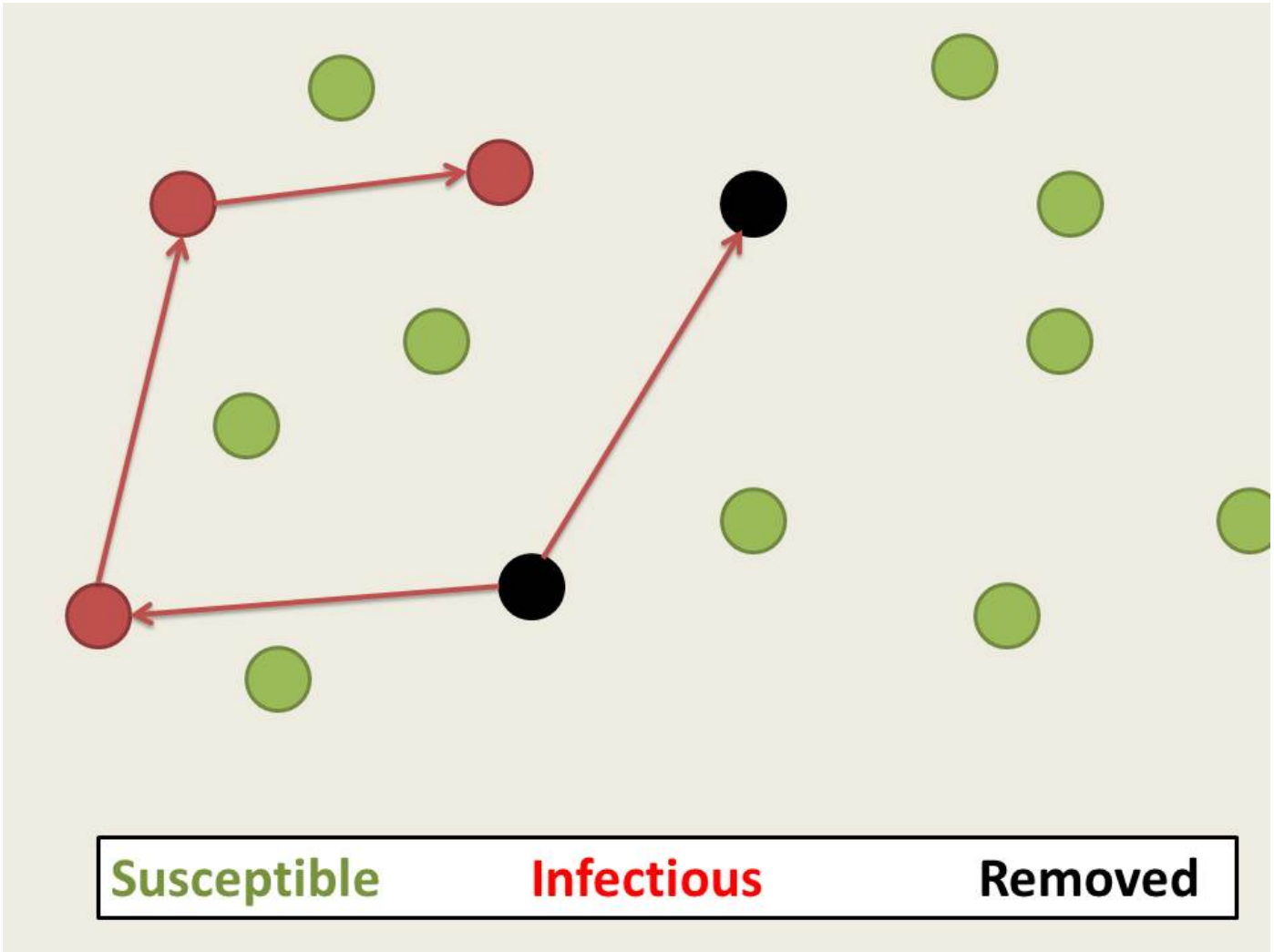
# How would an outbreak get started anyway? Example 1

An infectious host is removed.



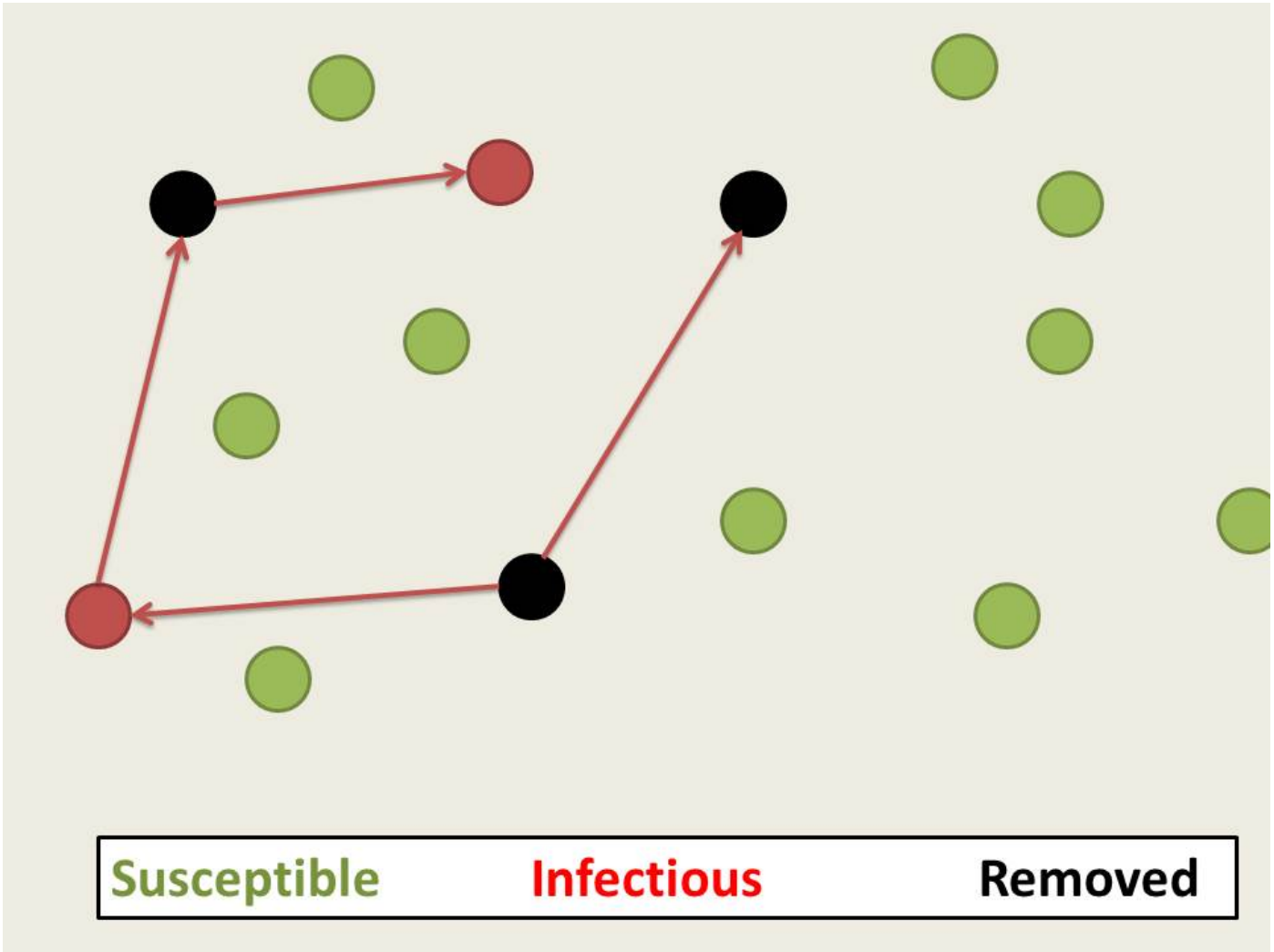
# How would an outbreak get started anyway? Example 1

A new infection occurs.



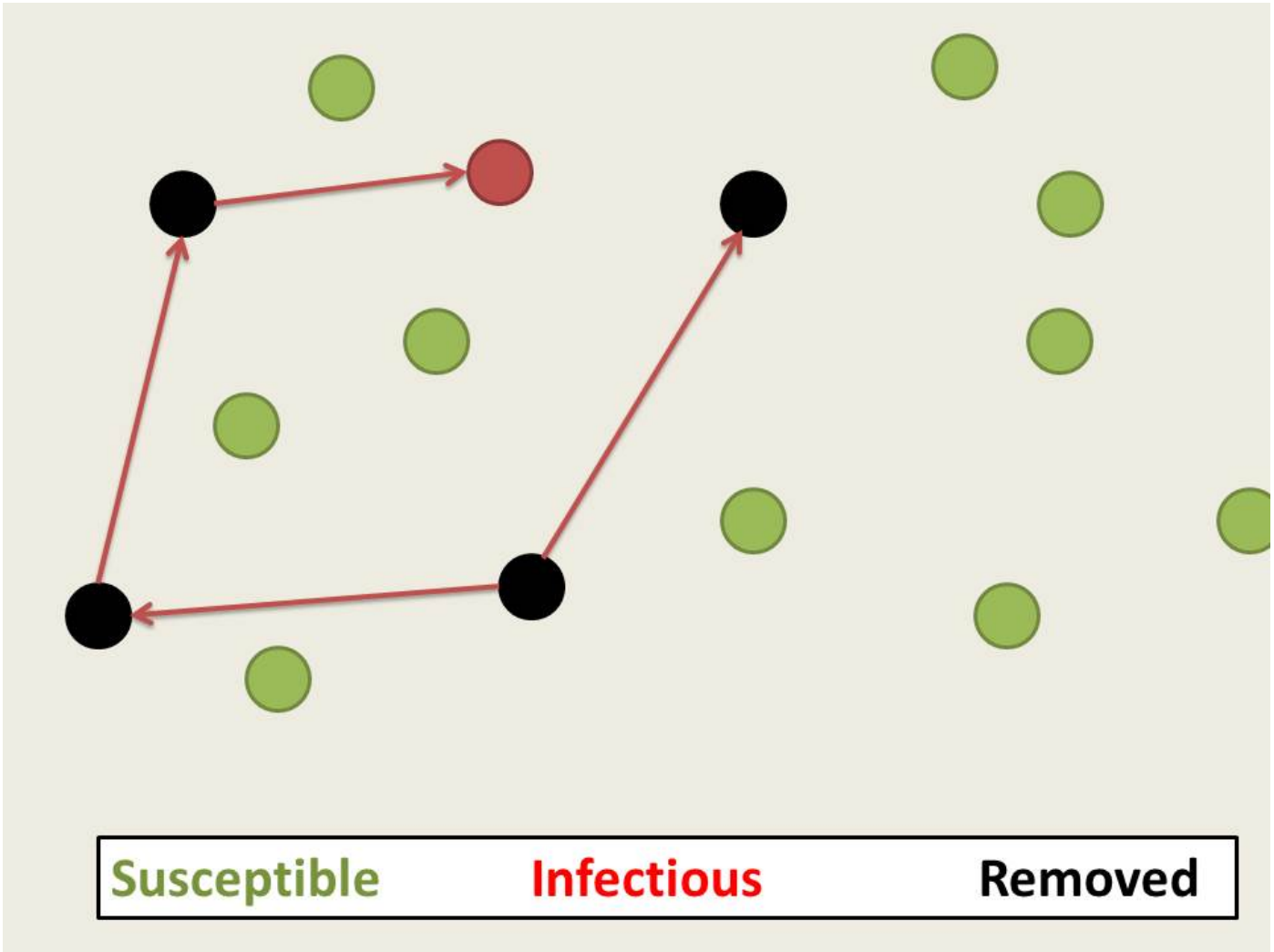
# How would an outbreak get started anyway? Example 1

An infectious host is removed.



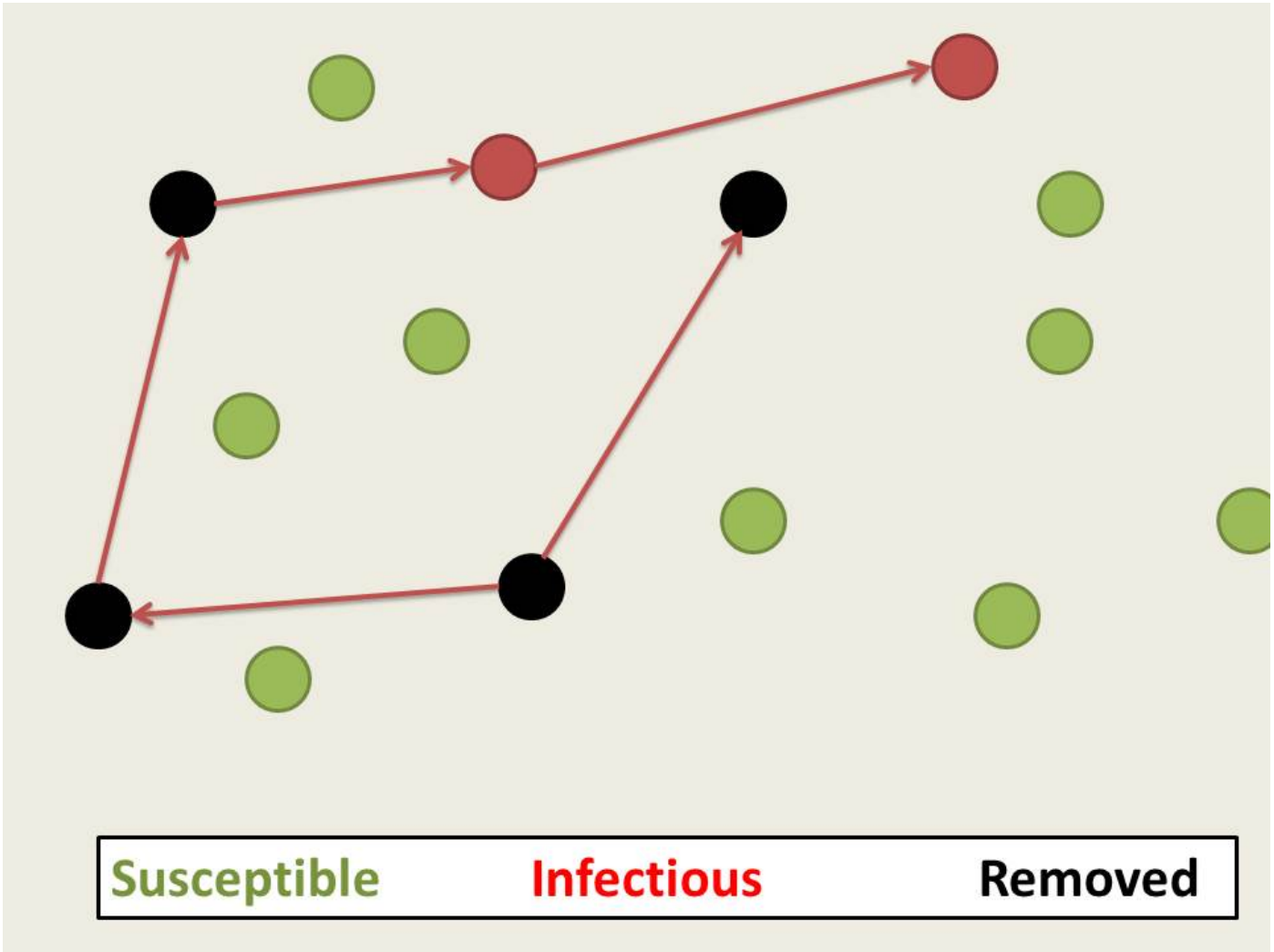
# How would an outbreak get started anyway? Example 1

An infectious host is removed.



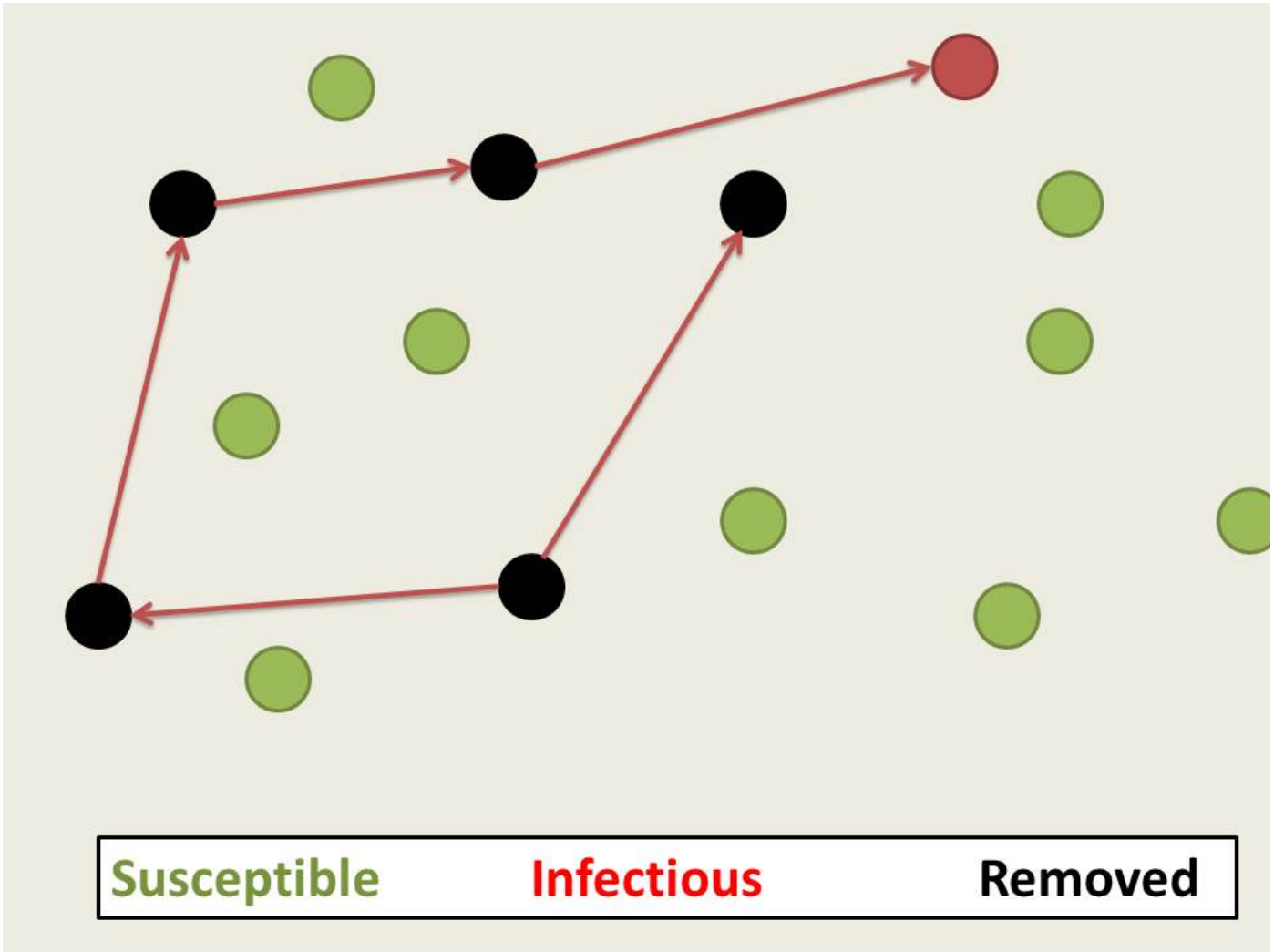
# How would an outbreak get started anyway? Example 1

A new infection occurs.



# How would an outbreak get started anyway? Example 1

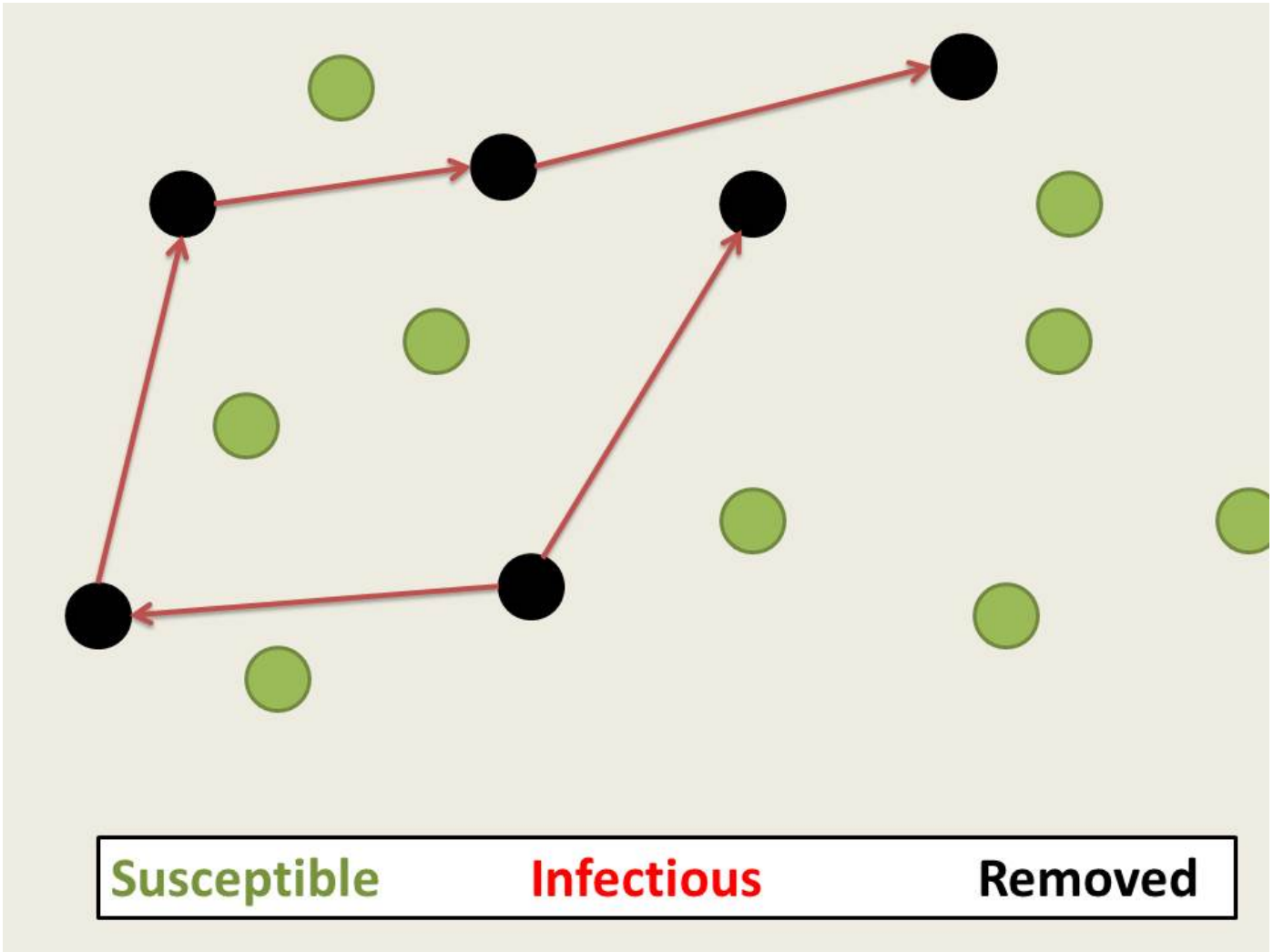
An infectious host is removed.



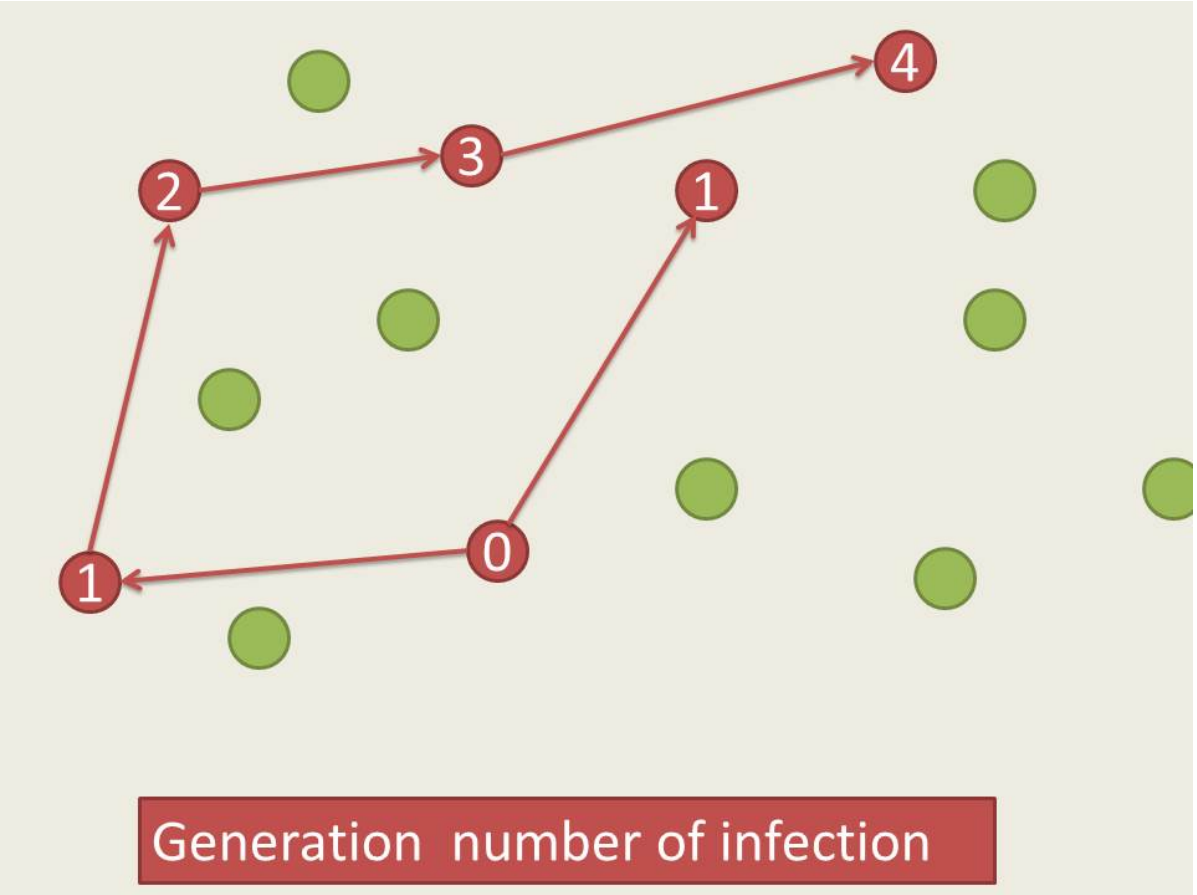
# How would an outbreak get started anyway? Example 1

An infectious host is removed. The infection has died out.

We have a minor outbreak in this example.



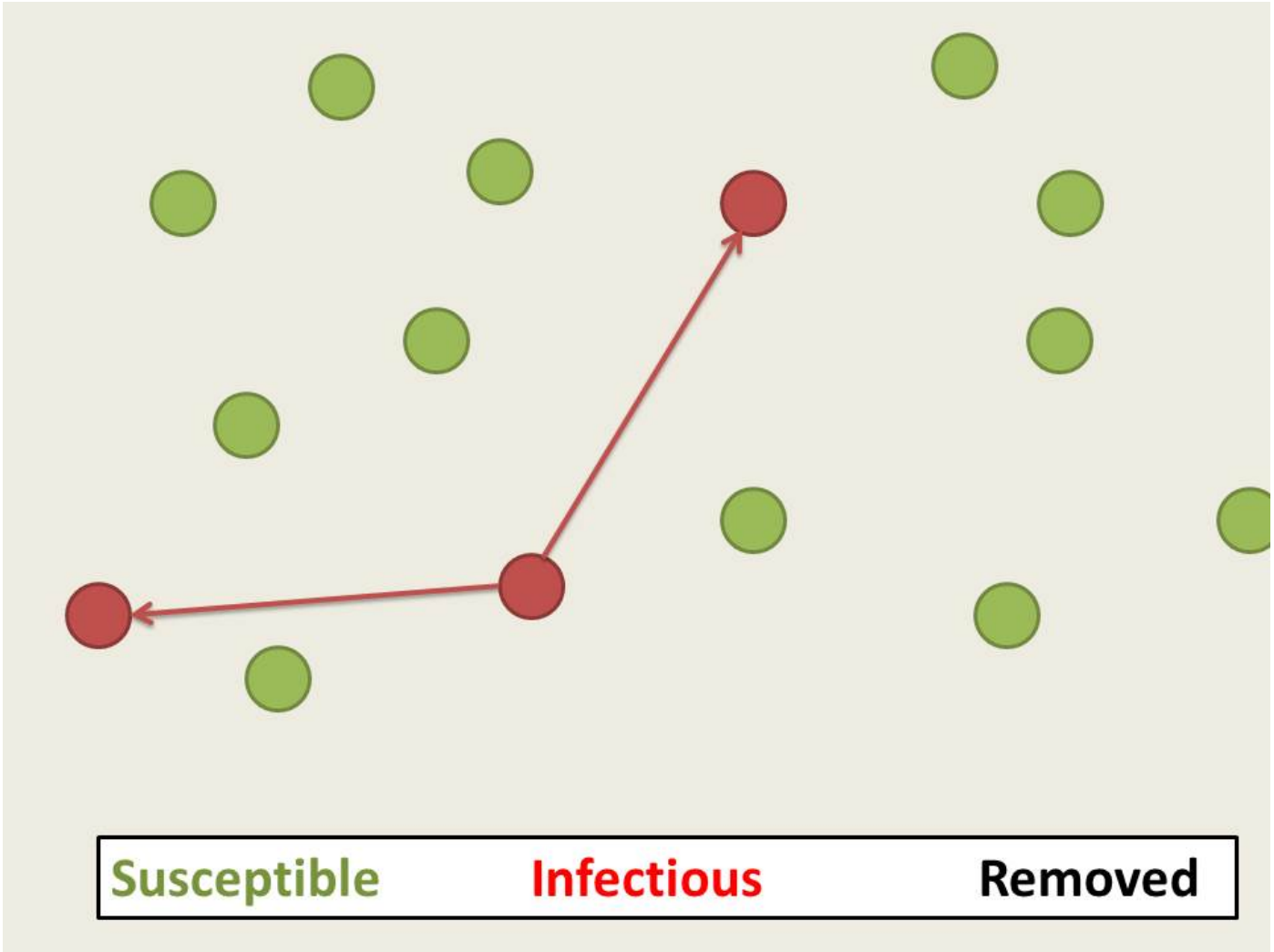
# The generations of the infection in Example 1




The average number of secondary infections per infectious host in this example is $\frac{2+0+1+1+1+0}{6} = \frac{5}{6}$.

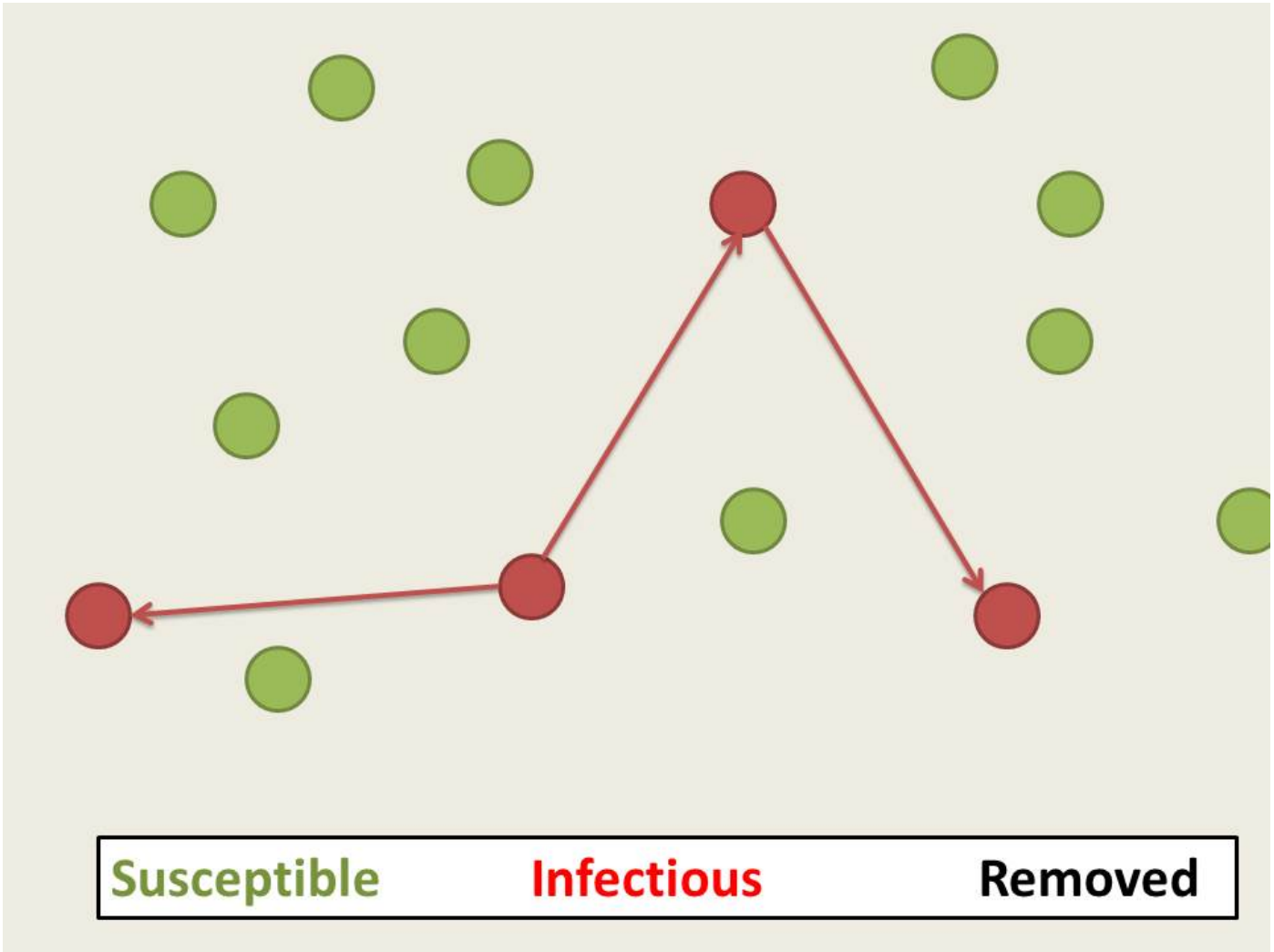
# How would an outbreak get started anyway? Example 2

Generations 0 and 1 might look as in Example 1.



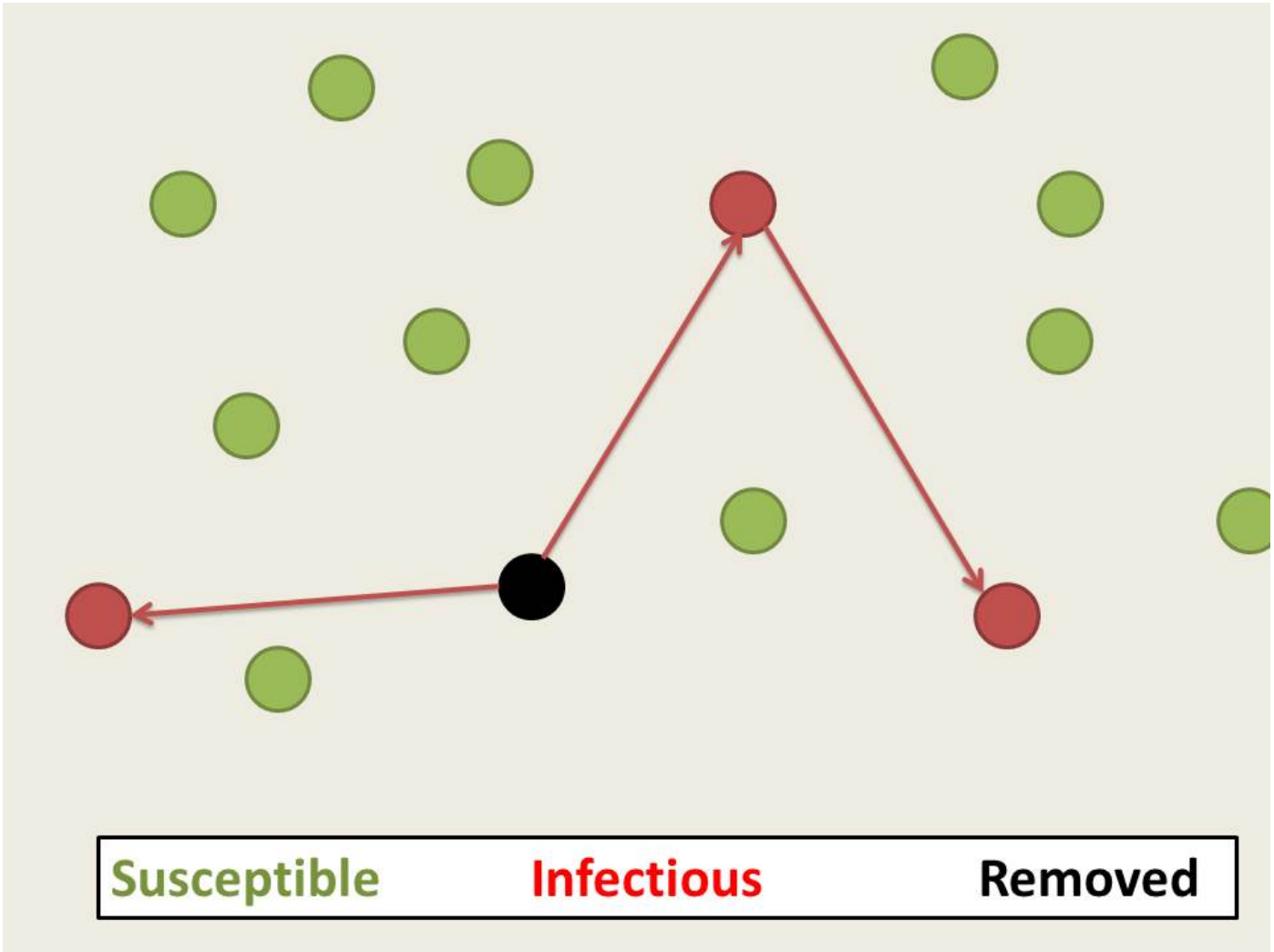
# How would an outbreak get started anyway? Example 2

A new infection occurs.



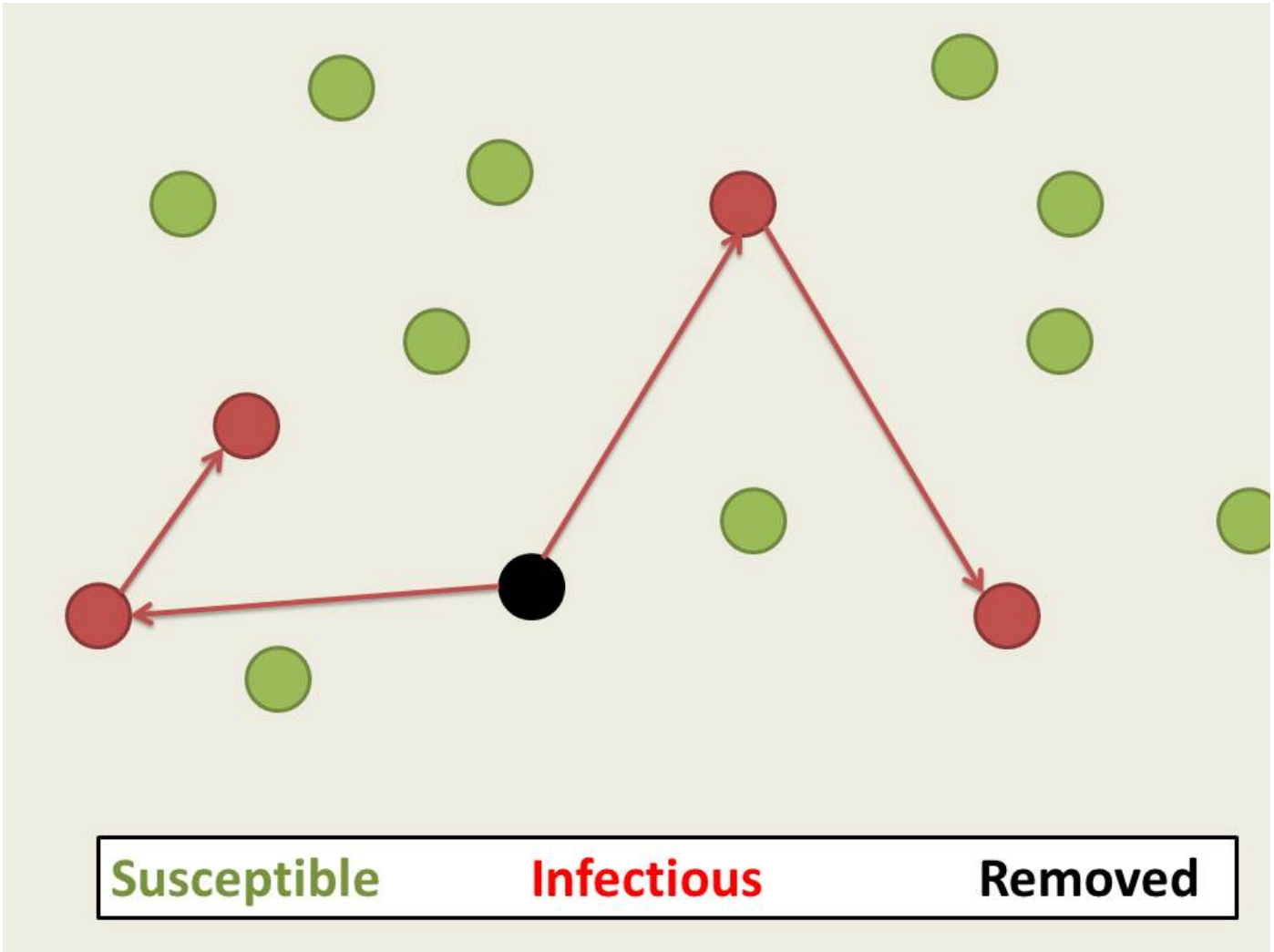
# How would an outbreak get started anyway? Example 2

An infectious host is removed.



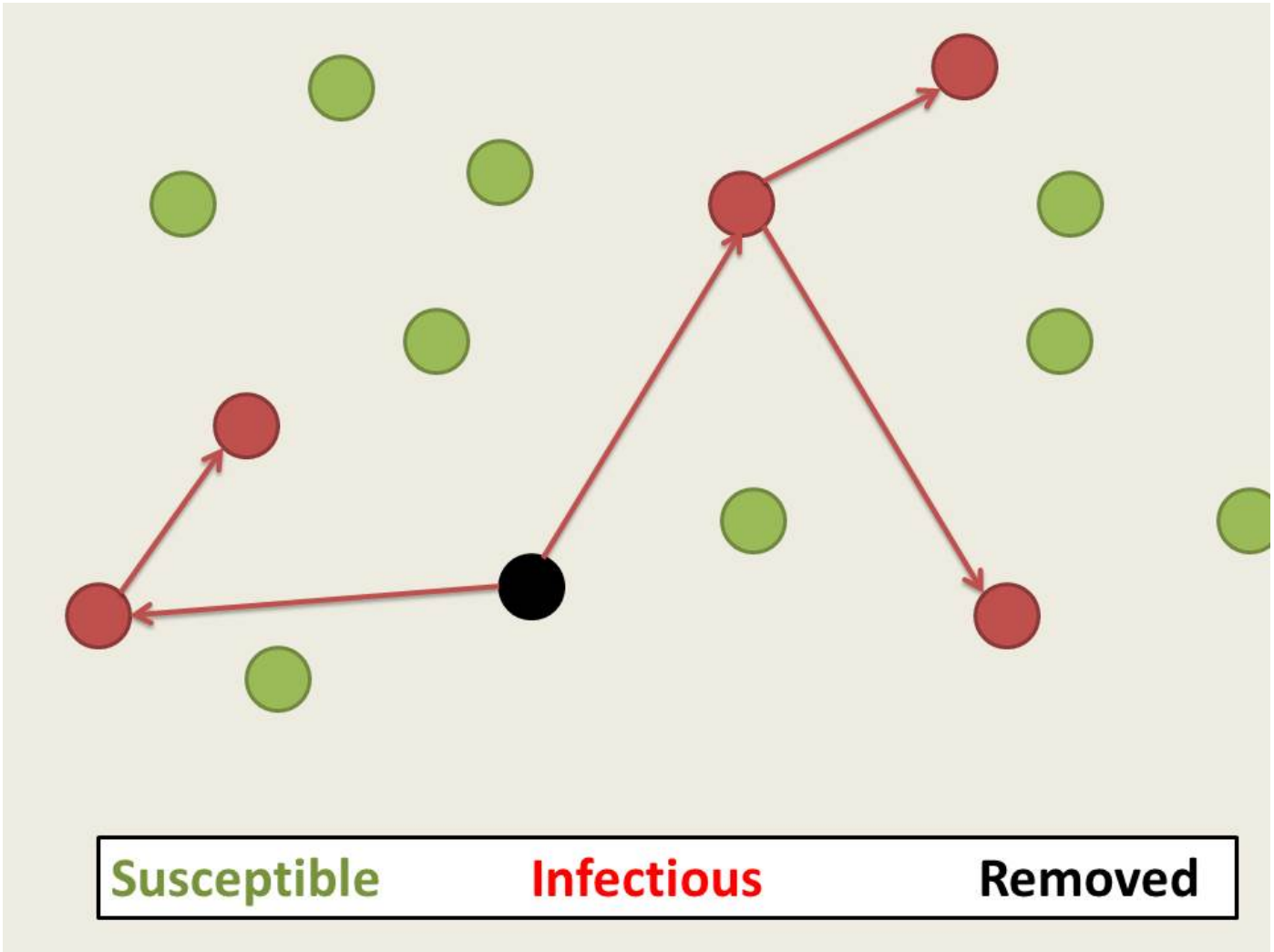
# How would an outbreak get started anyway? Example 2

A new infection occurs.



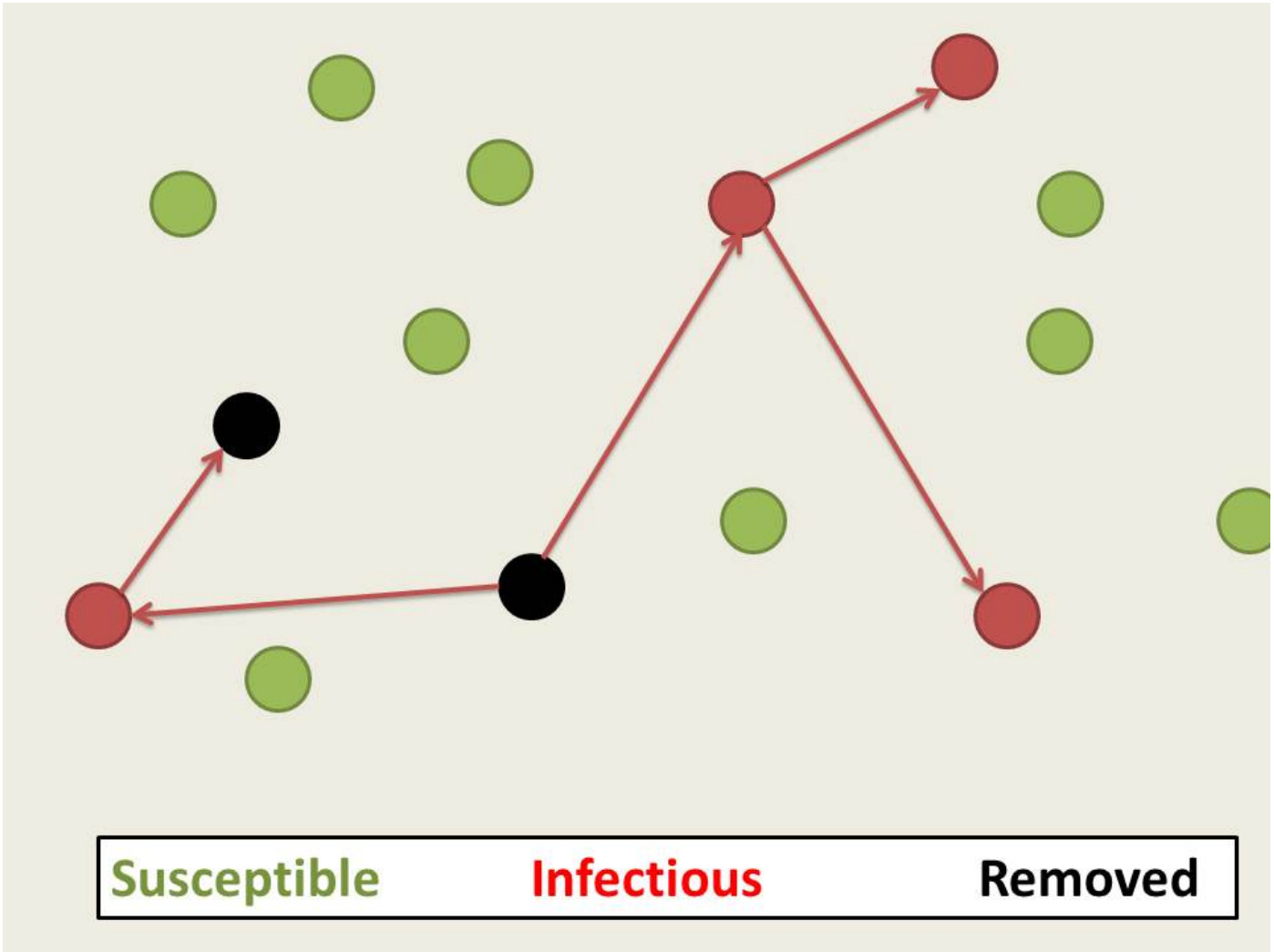
# How would an outbreak get started anyway? Example 2

A new infection occurs.



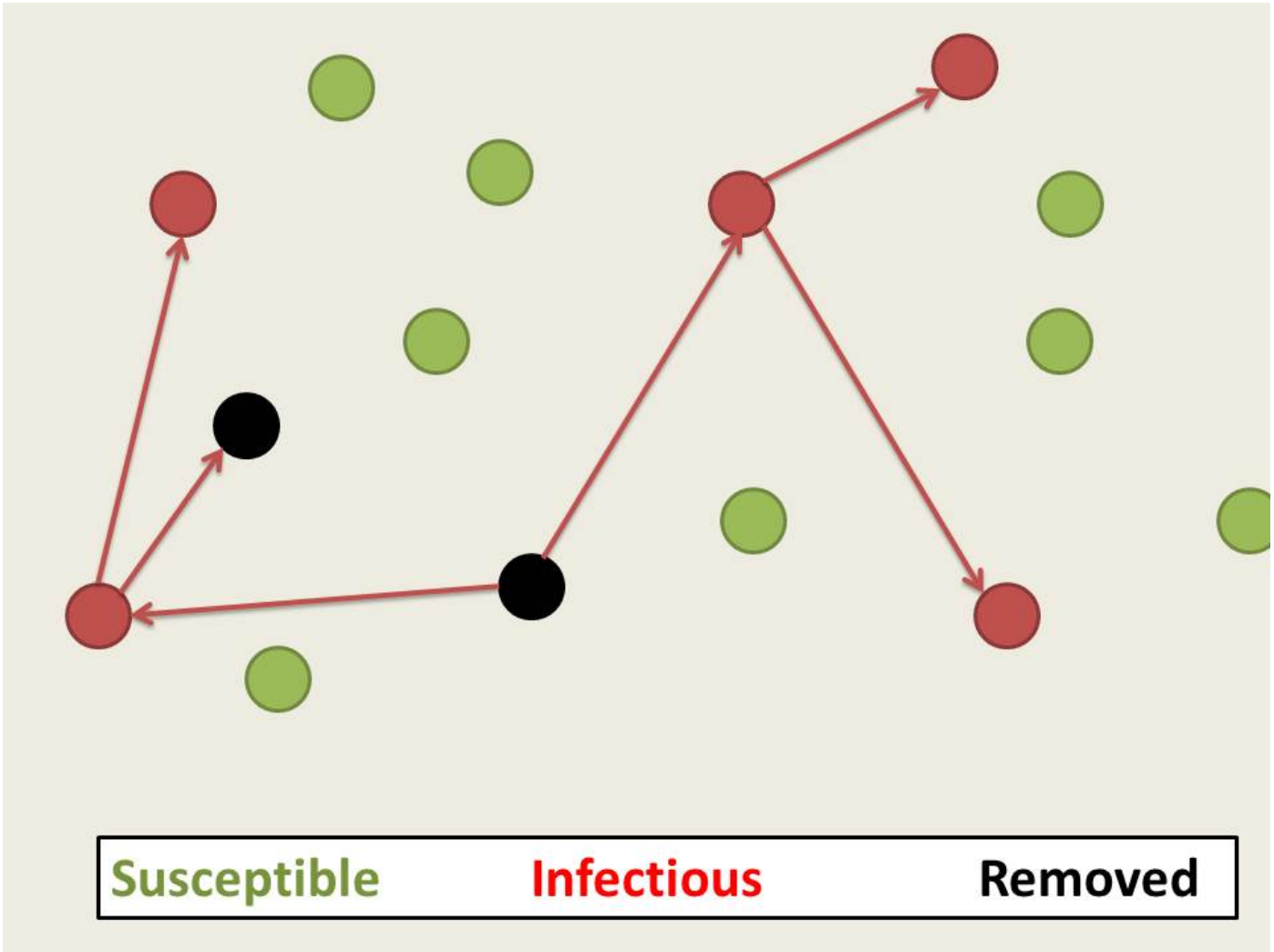
# How would an outbreak get started anyway? Example 2

An infectious host is removed.



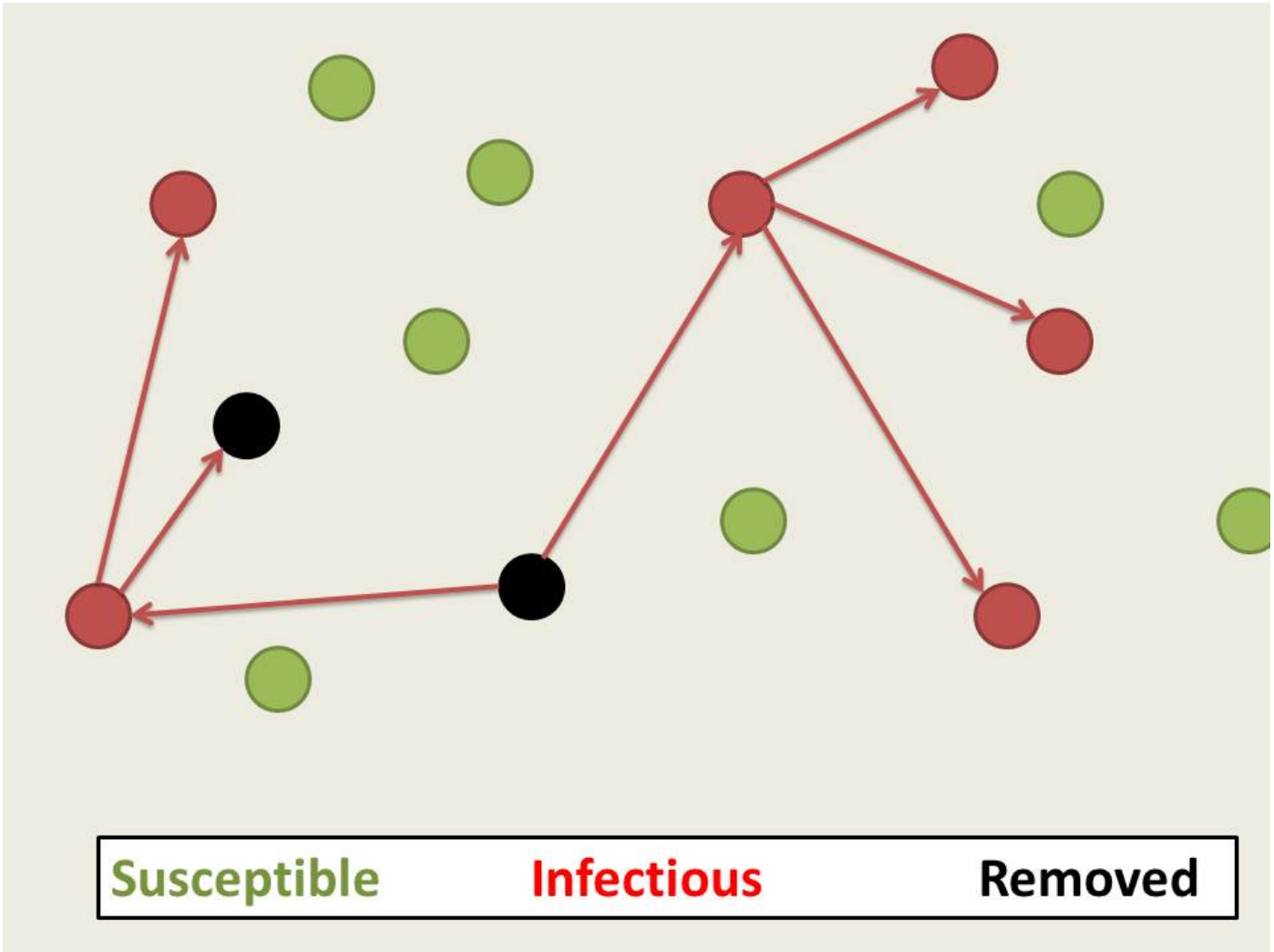
# How would an outbreak get started anyway? Example 2

A new infection occurs.



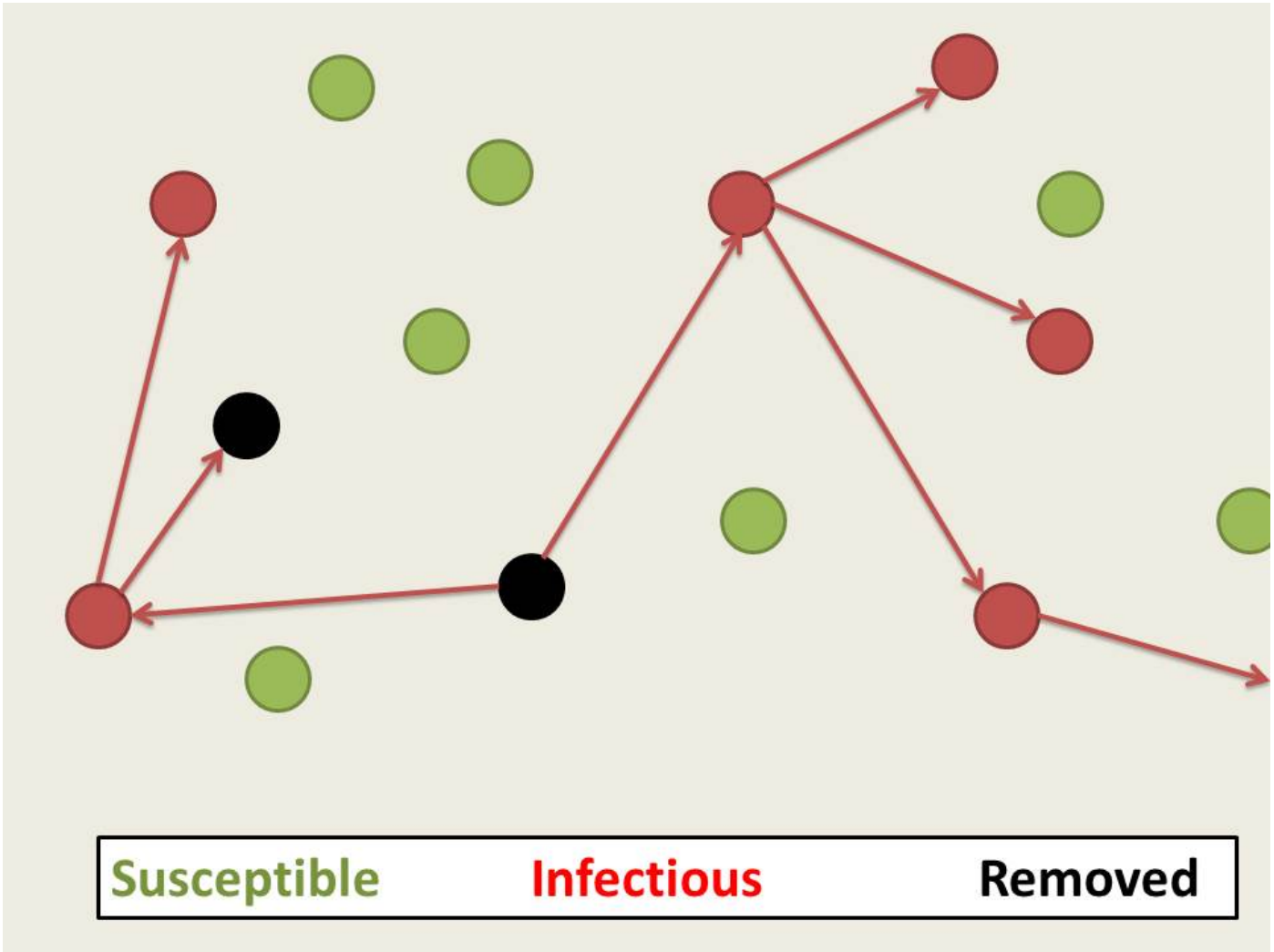
# How would an outbreak get started anyway? Example 2

A new infection occurs.



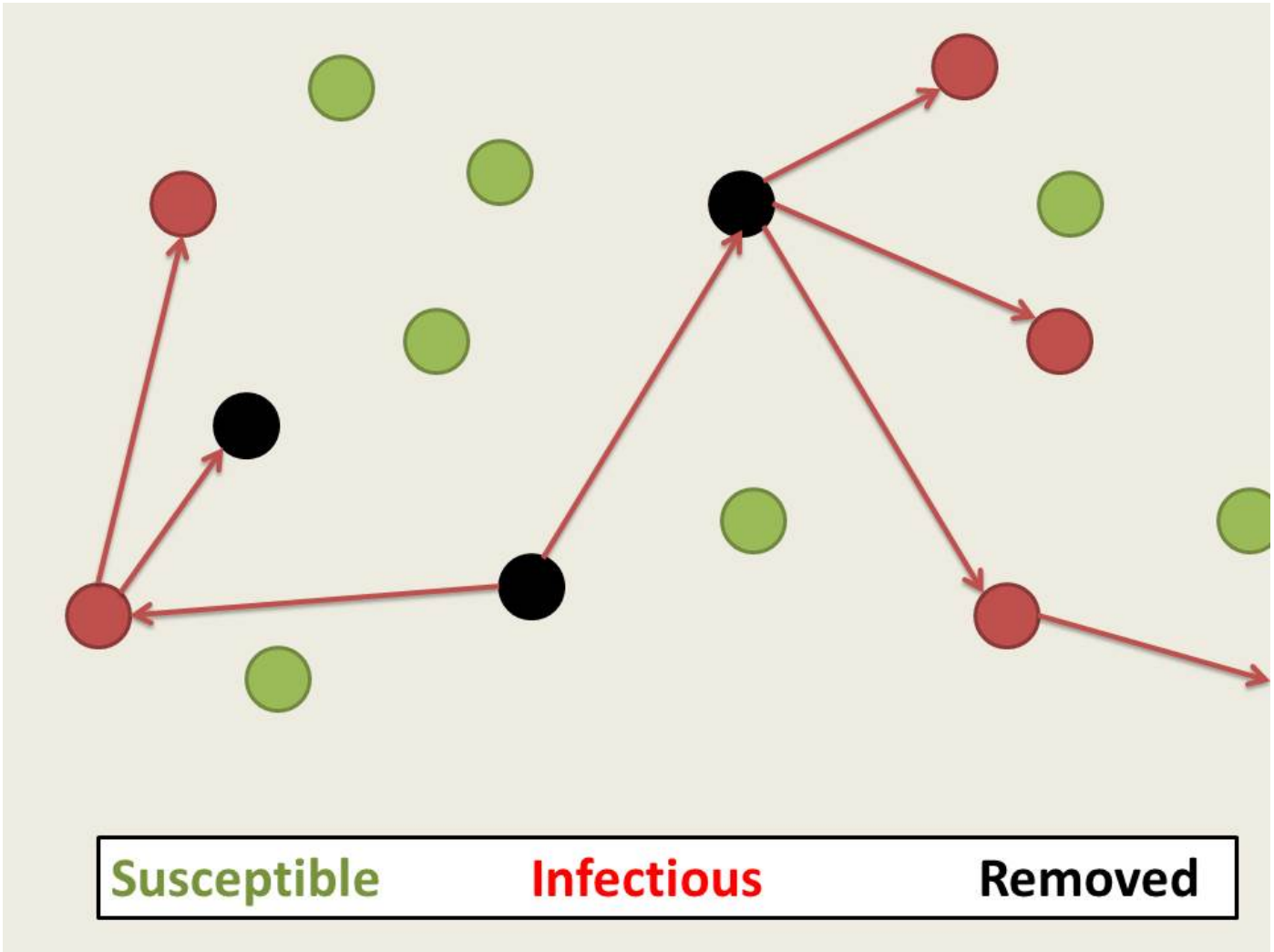
# How would an outbreak get started anyway? Example 2

A new infection occurs.



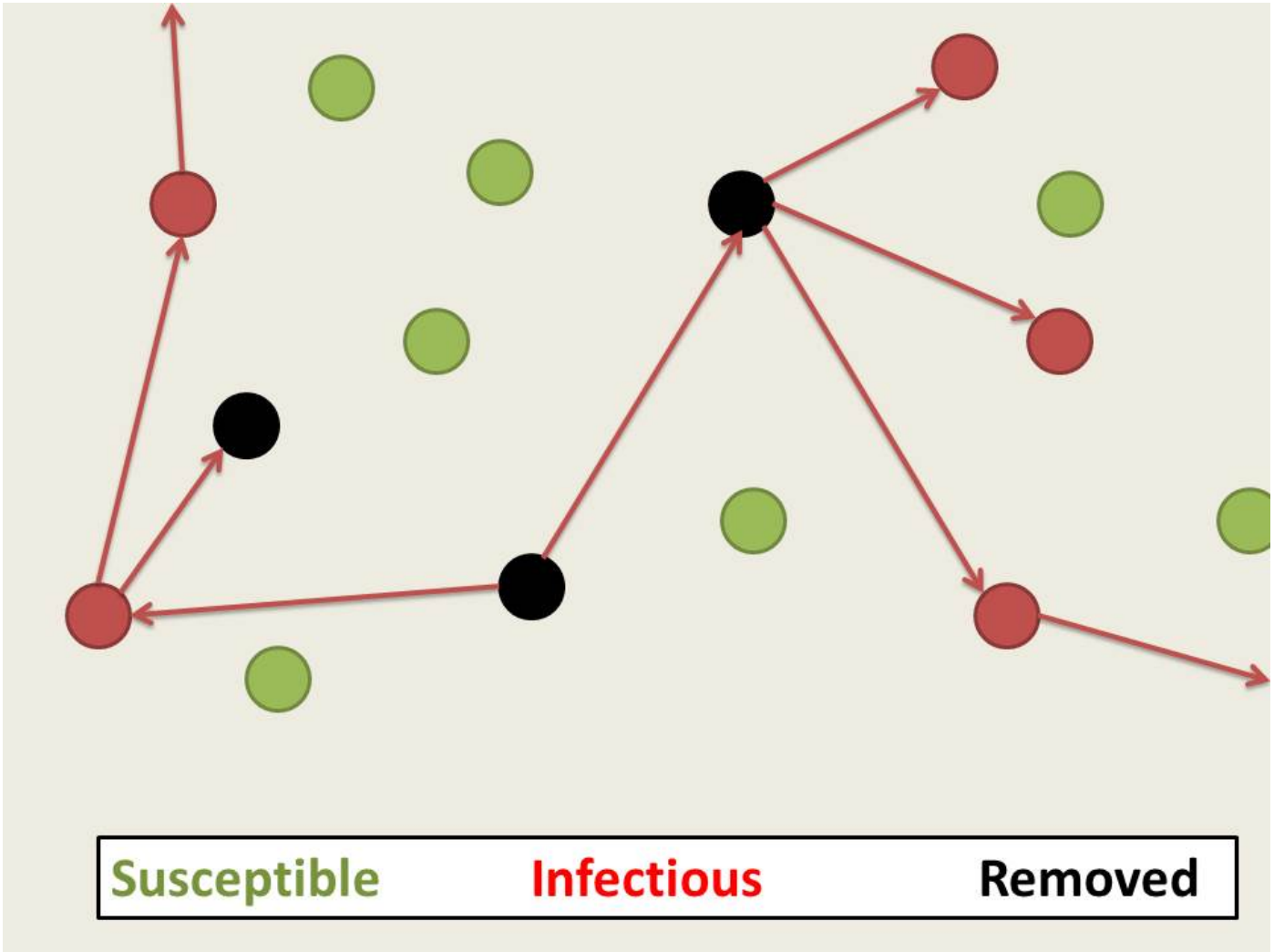
# How would an outbreak get started anyway? Example 2

An infectious host is removed.



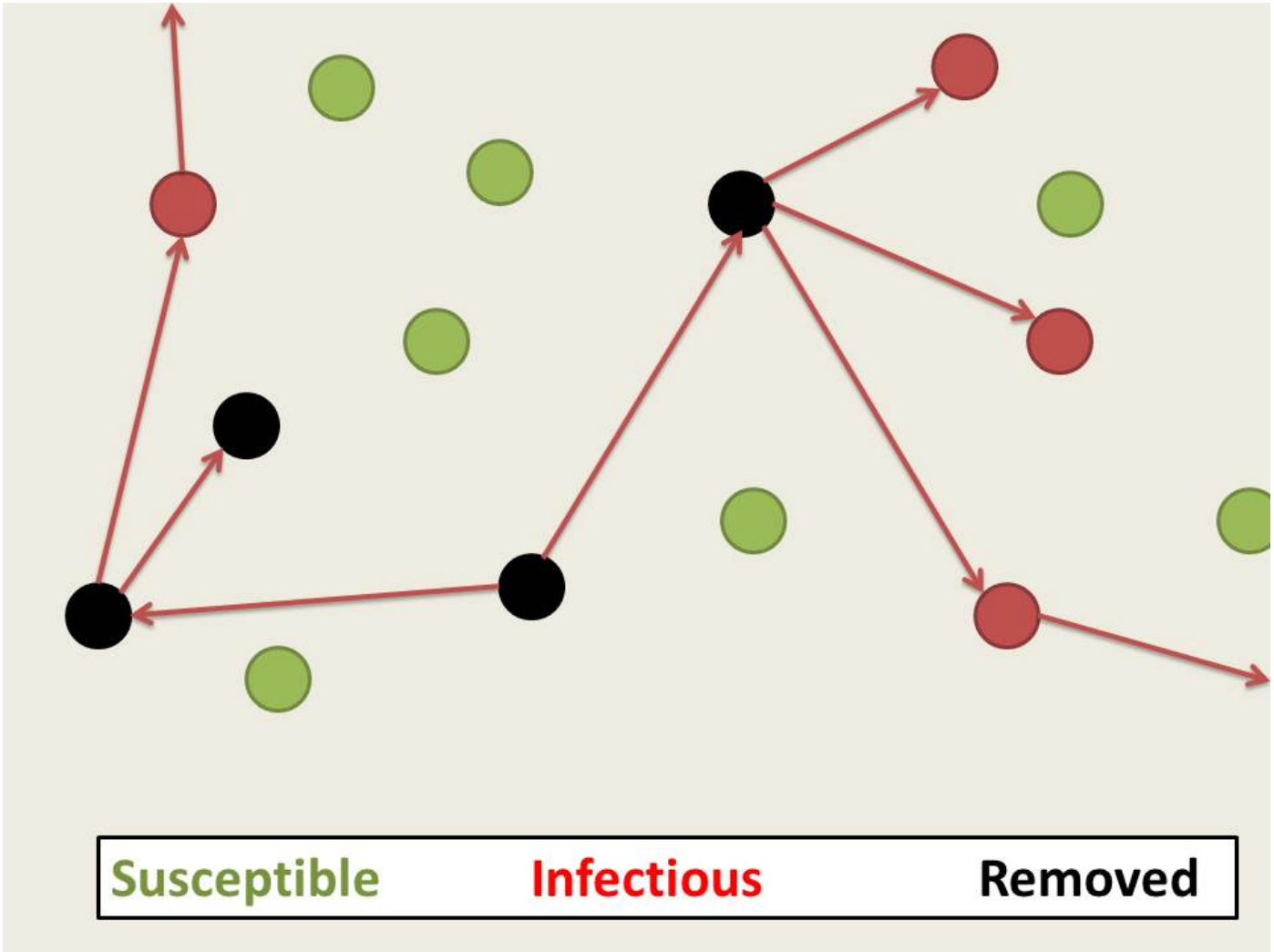
# How would an outbreak get started anyway? Example 2

A new infection occurs.



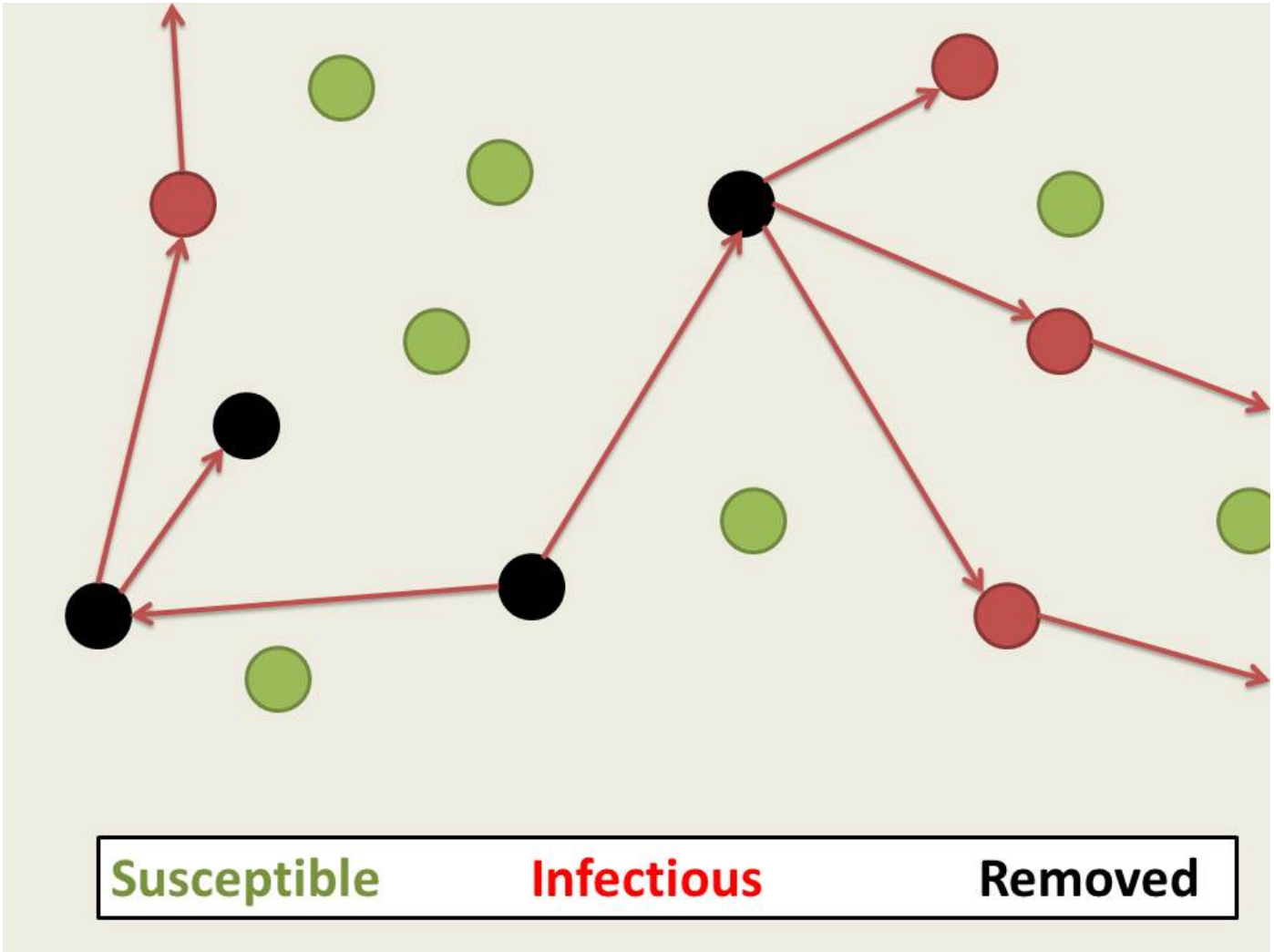
# How would an outbreak get started anyway? Example 2

An infectious host is removed.



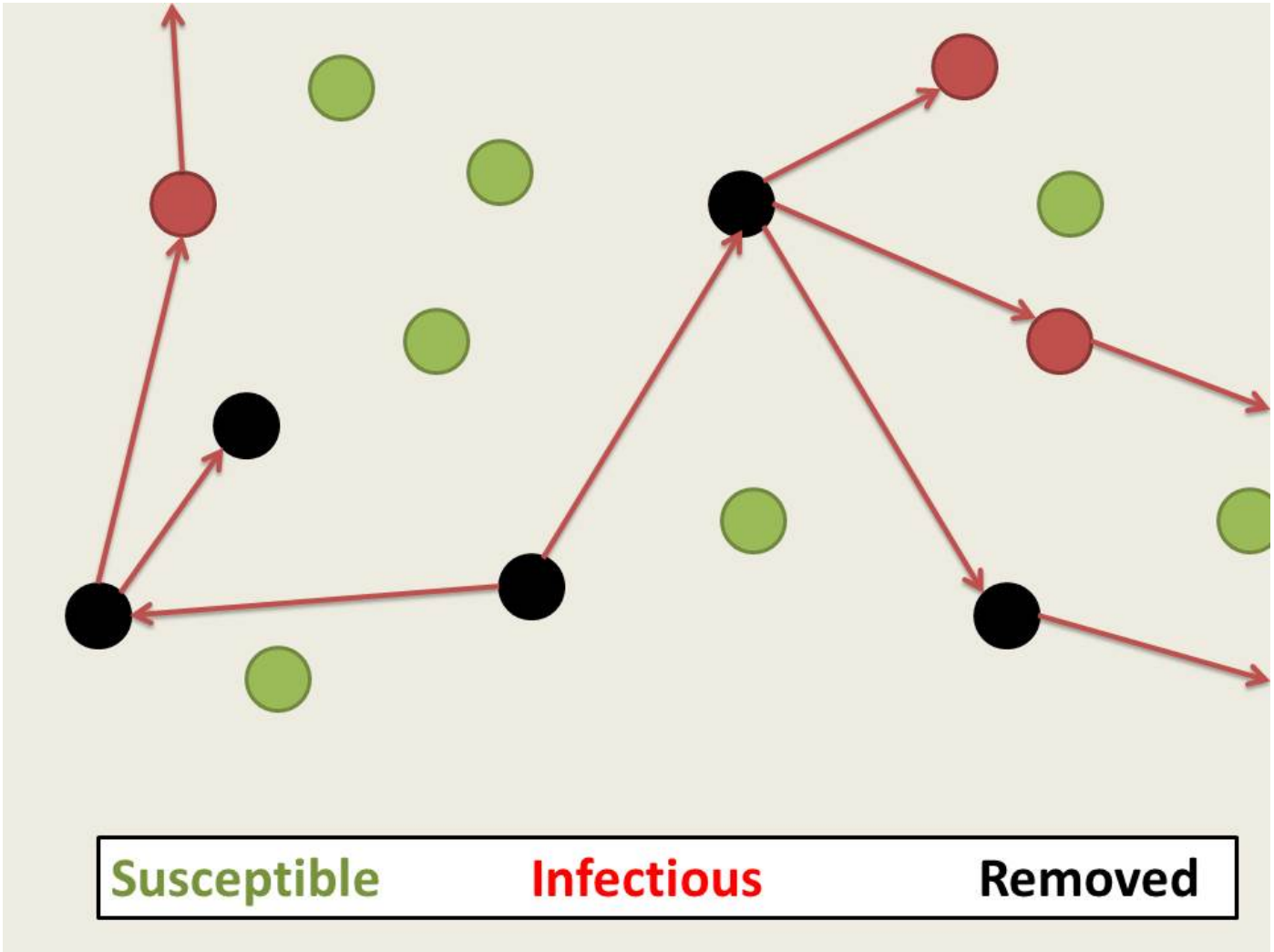
# How would an outbreak get started anyway? Example 2

A new infection occurs.



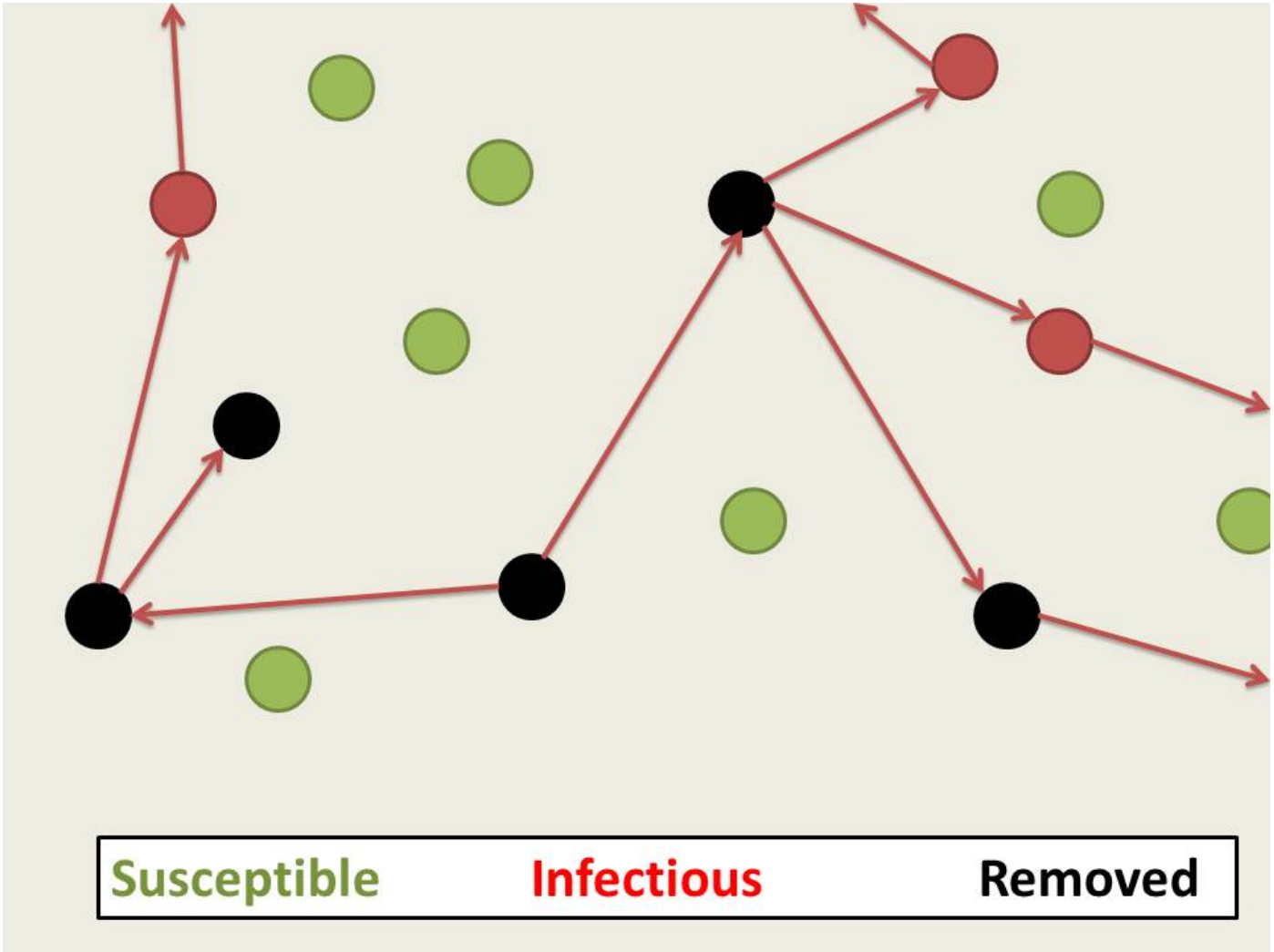
# How would an outbreak get started anyway? Example 2

An infectious host is removed.



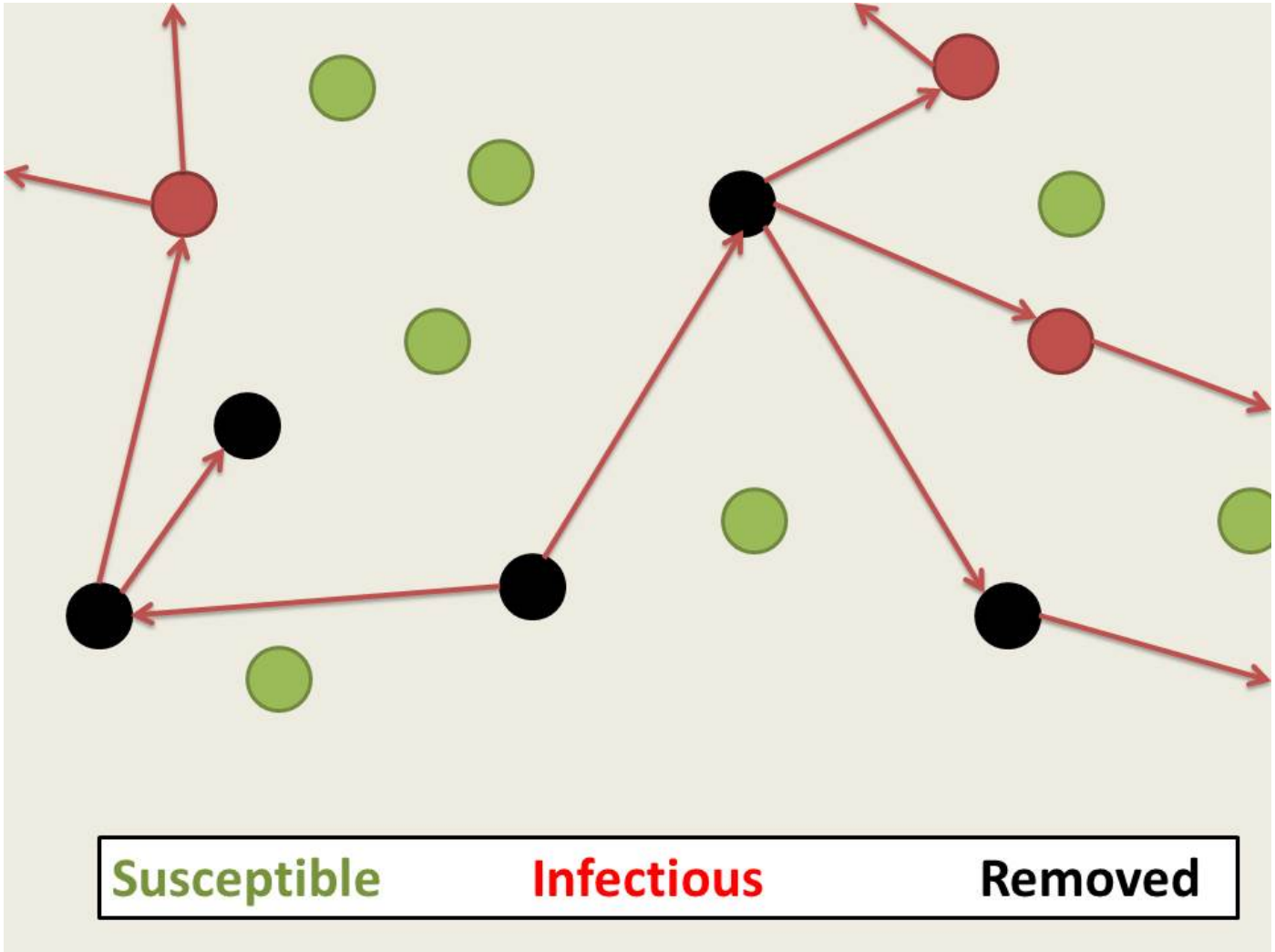
# How would an outbreak get started anyway? Example 2

A new infection occurs.



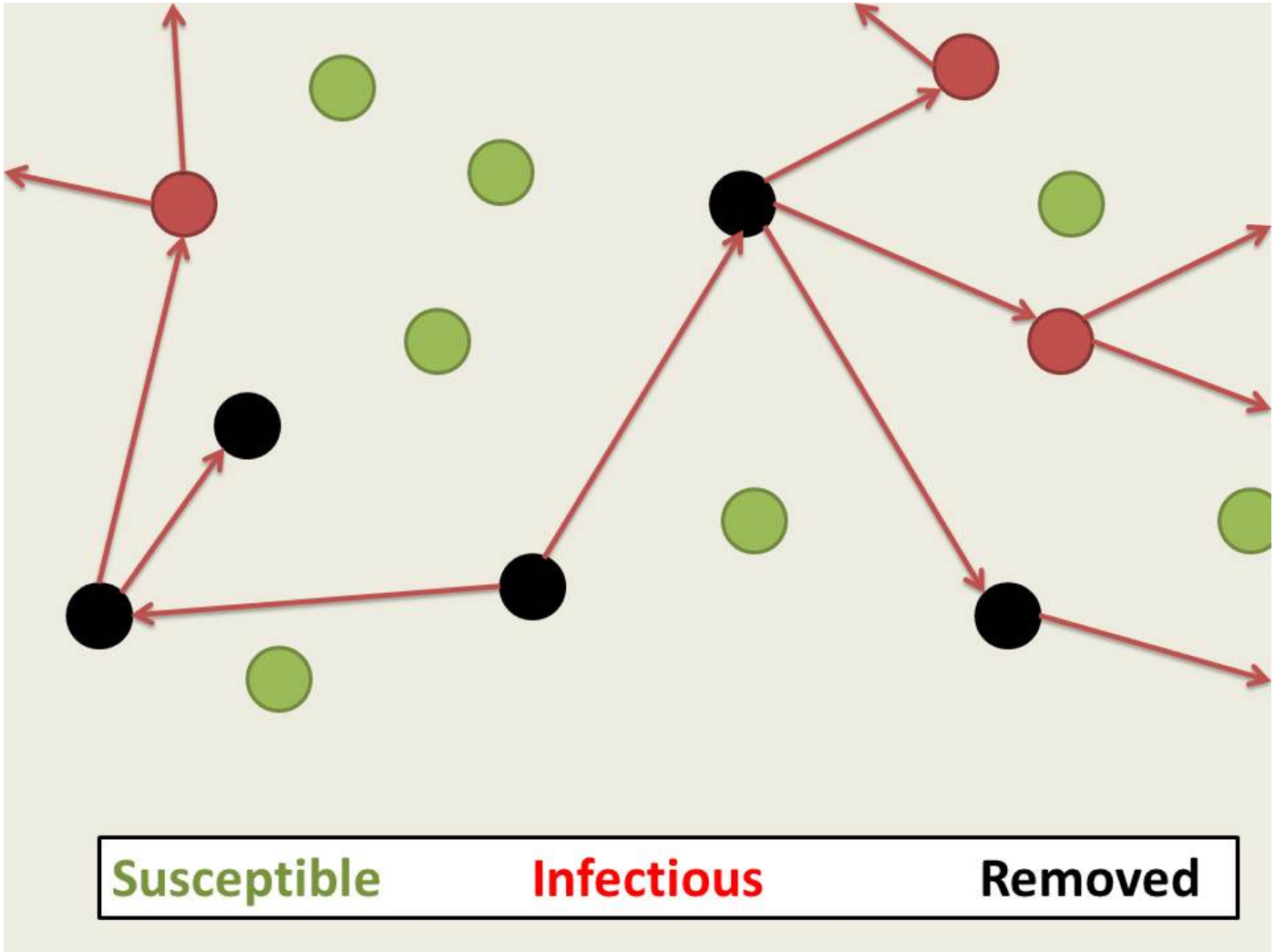
# How would an outbreak get started anyway? Example 2

A new infection occurs.



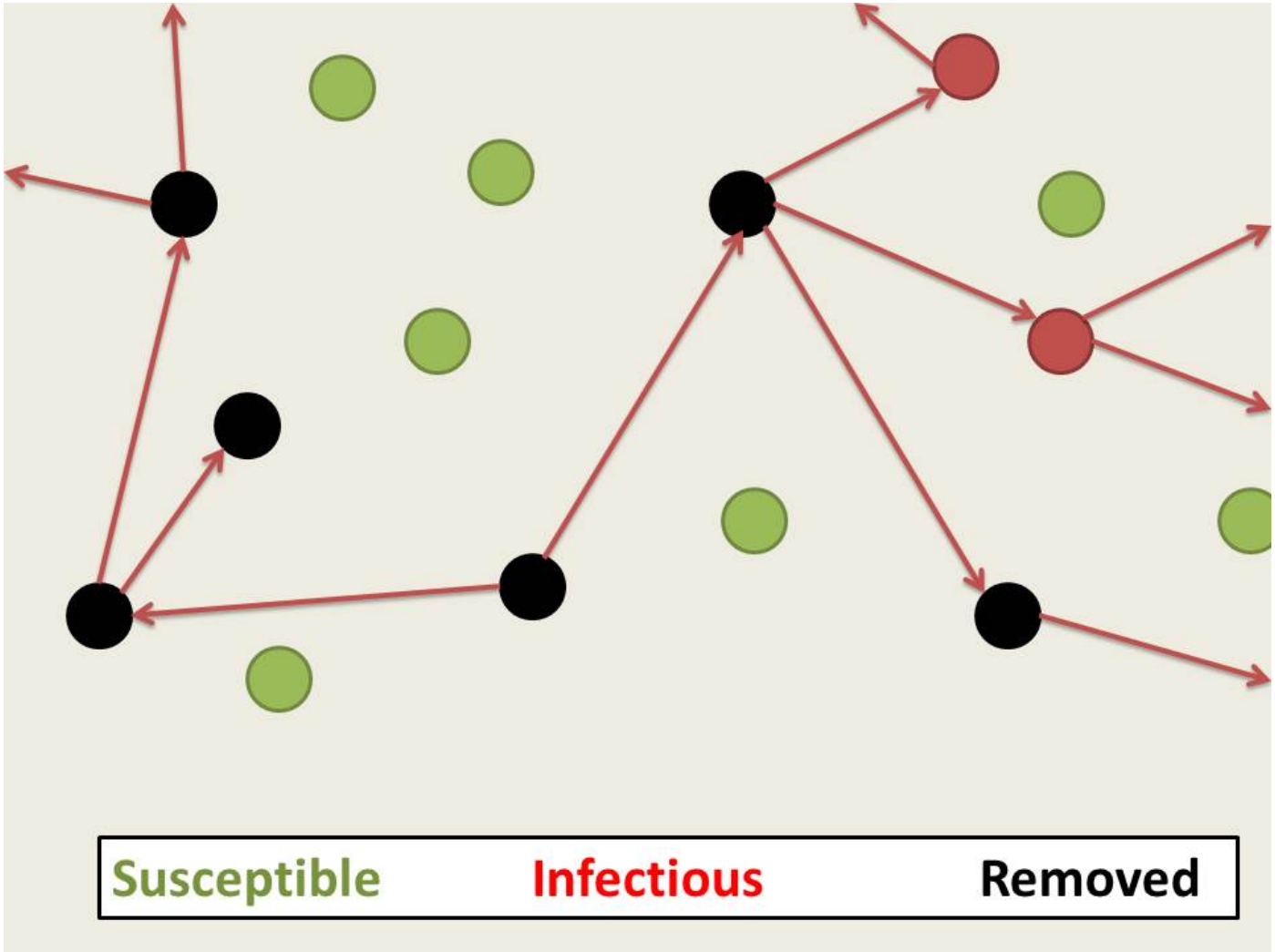
# How would an outbreak get started anyway? Example 2

A new infection occurs.



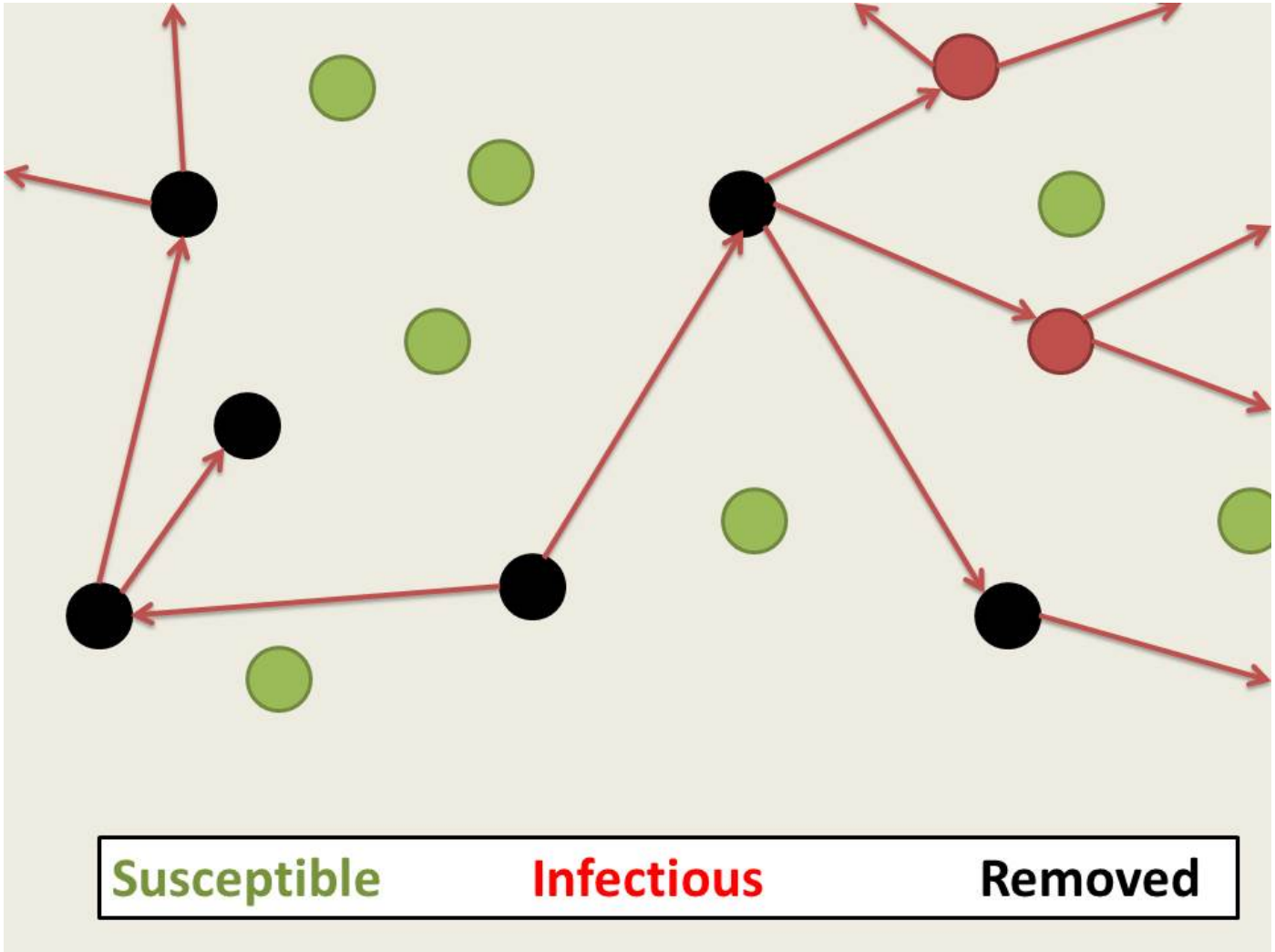
# How would an outbreak get started anyway? Example 2

An infectious host is removed.



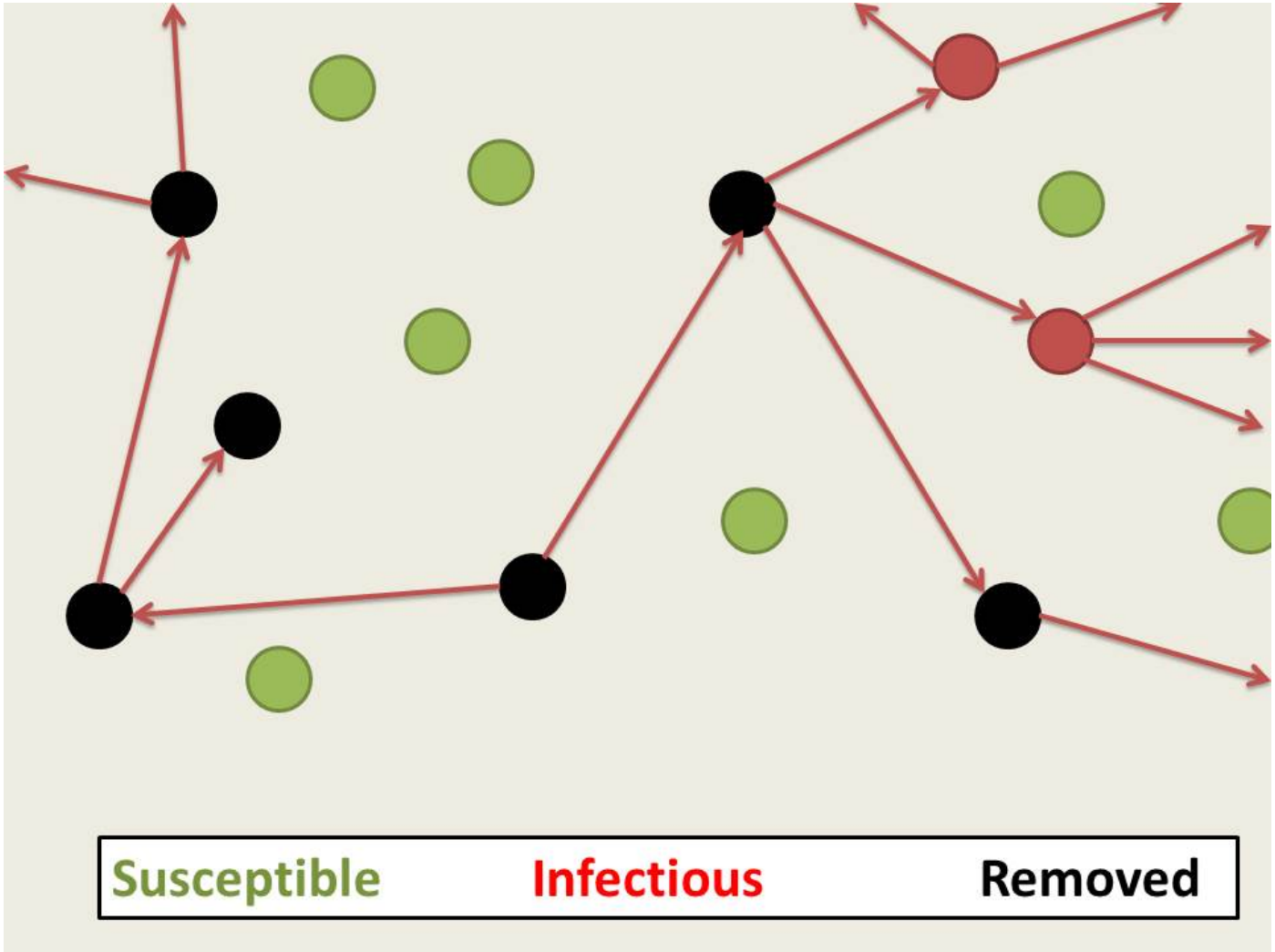
# How would an outbreak get started anyway? Example 2

A new infection occurs.



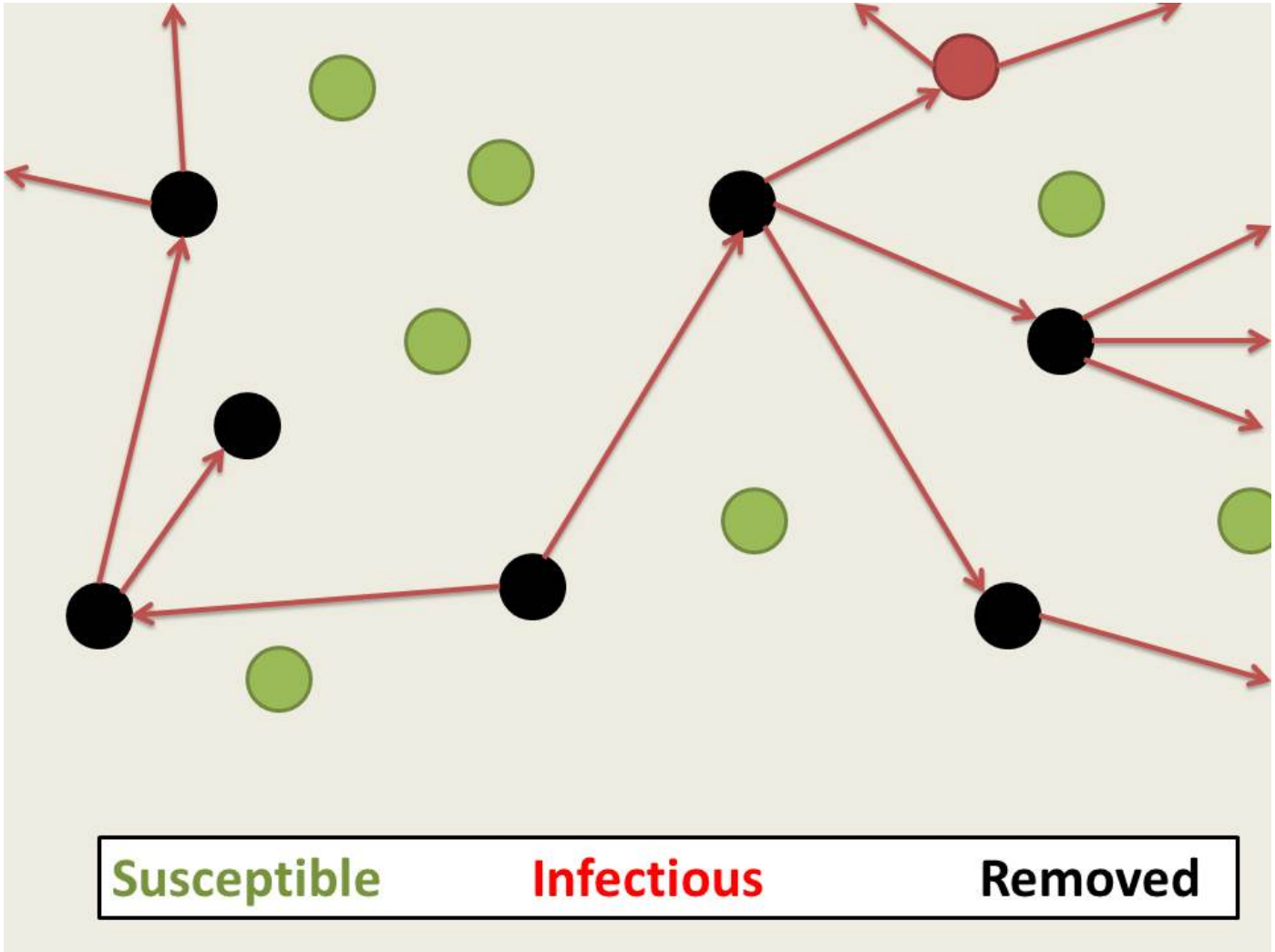
# How would an outbreak get started anyway? Example 2

A new infection occurs.



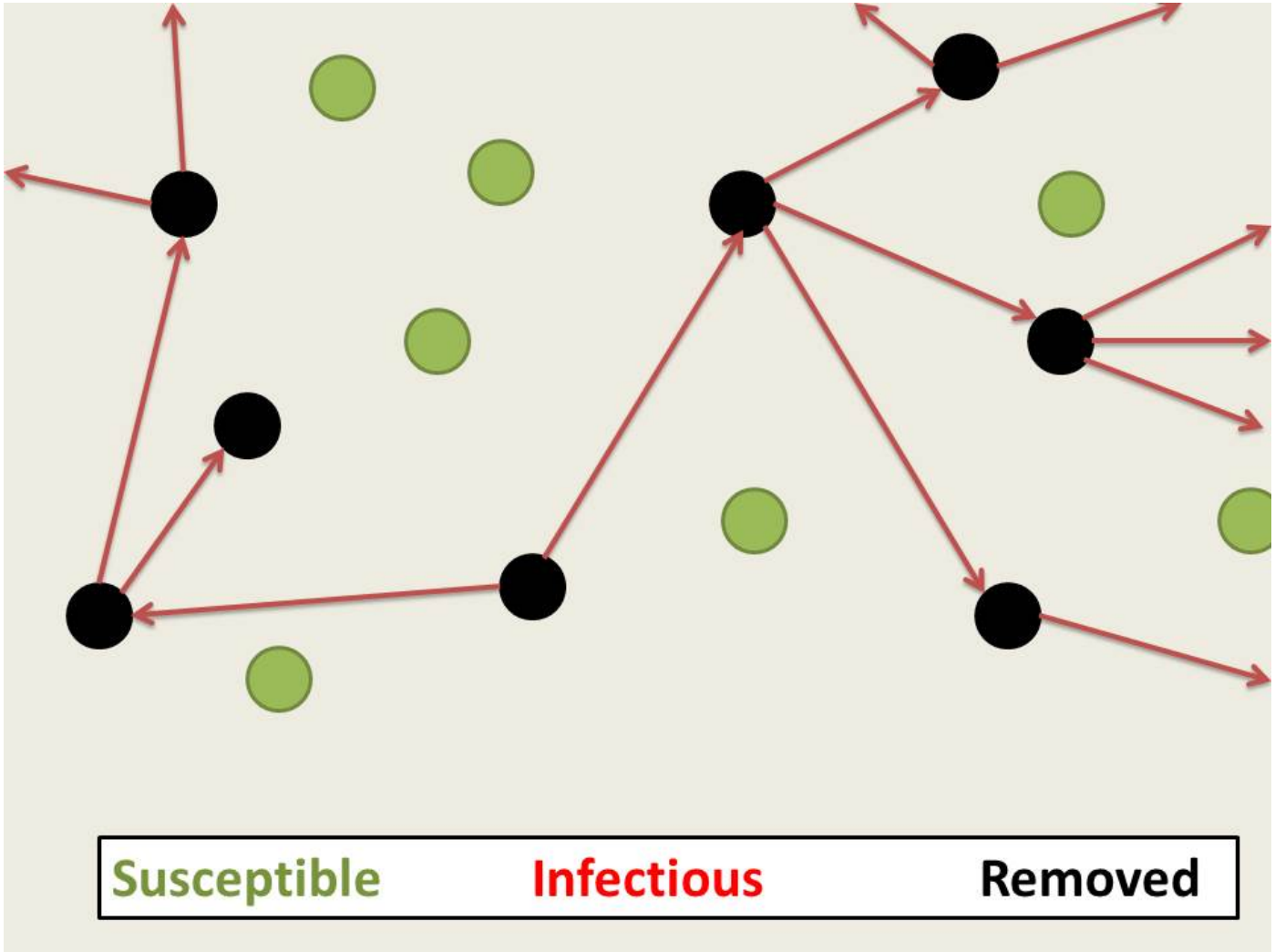
# How would an outbreak get started anyway? Example 2

An infectious host is removed.

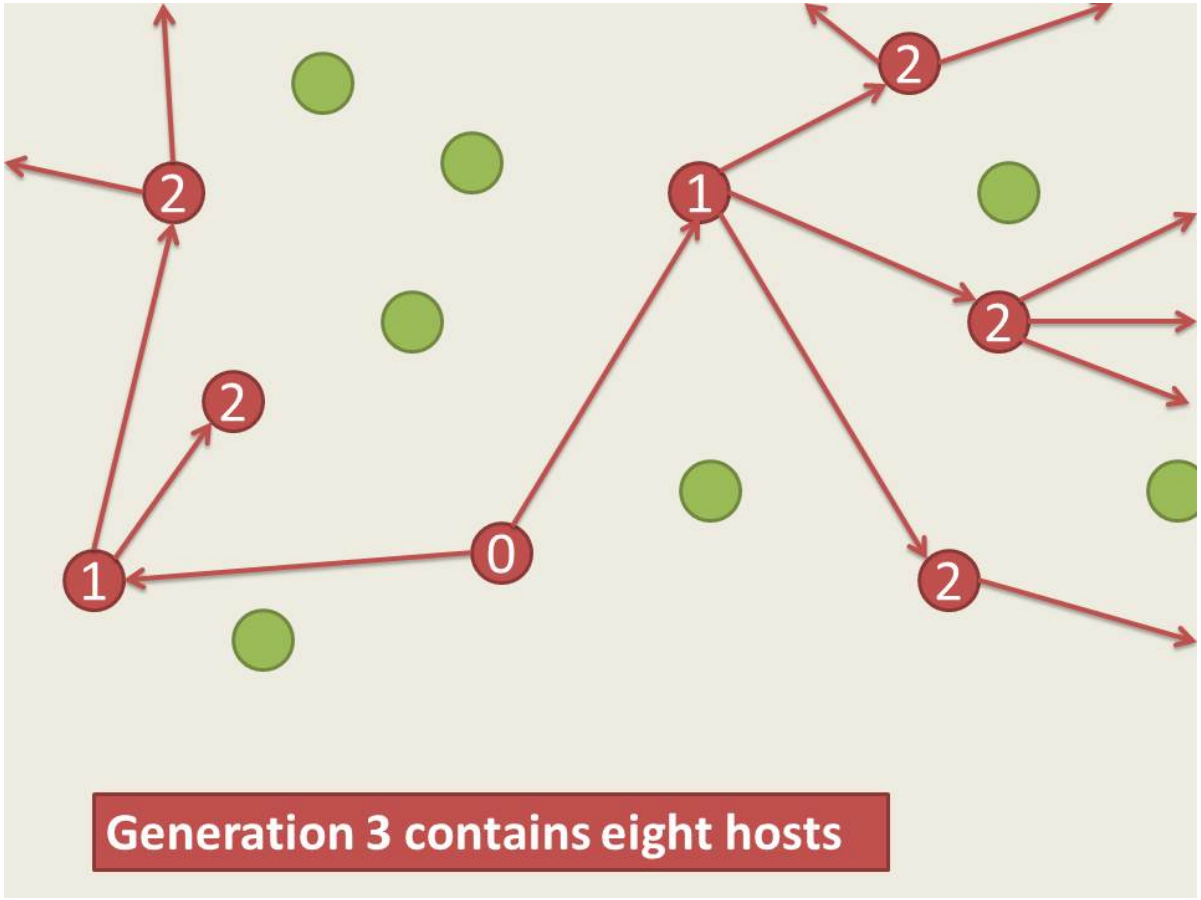


# How would an outbreak get started anyway? Example 2

An infectious host is removed.

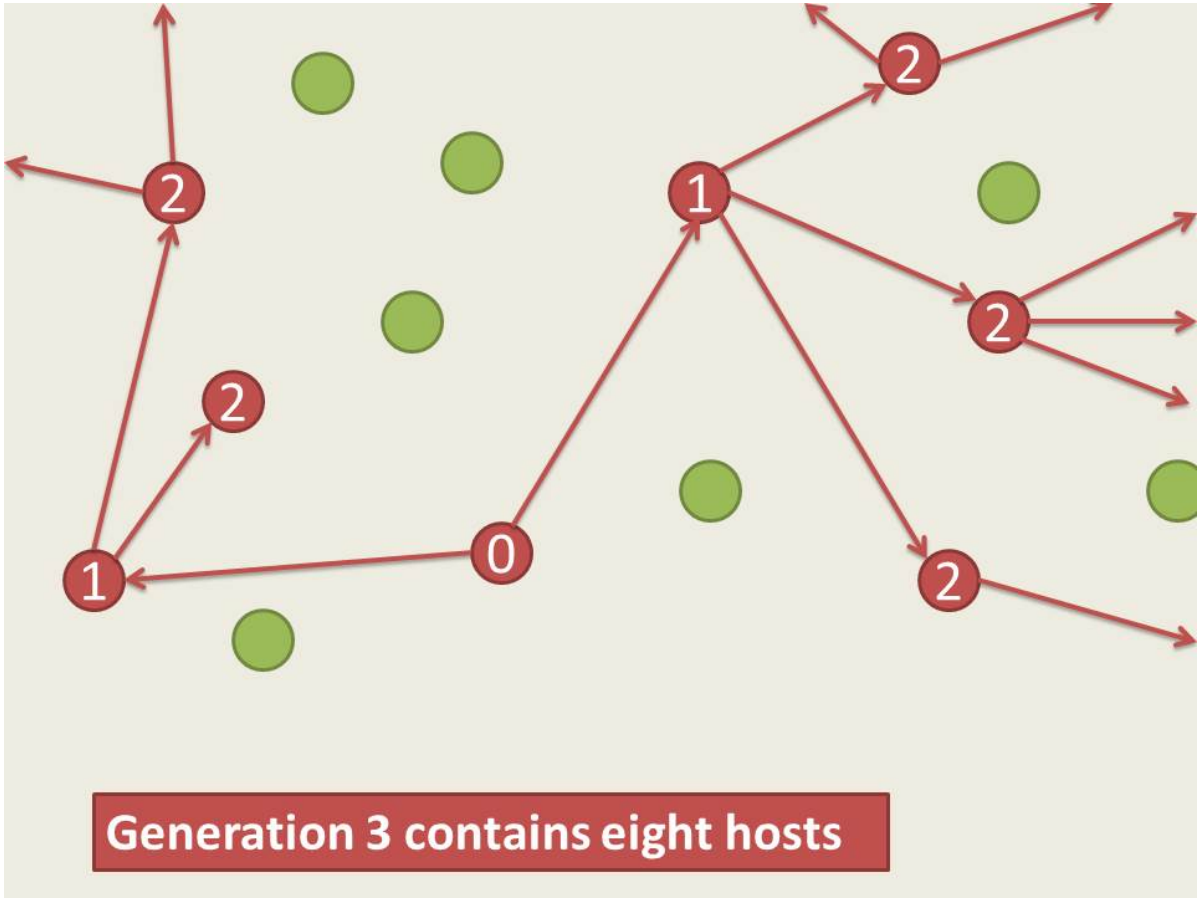


# Some generations of the infection in Example 2




The average number of secondary infections per infectious host in generations 0 to 2 in this example is $\frac{2+2+0+2+3+2+3+1}{8} = \frac{15}{8}$.

# Does Example 2 indicate the start of an epidemic?




Most likely. From generation 0 to generation 3 the number of infectious hosts has increased by a factor of 8, and one might expect similar increases in subsequent generations.

# What makes the difference?

**Definition**

The expected number of secondary infections that will be caused by a single infectious host that is introduced into a large and entirely susceptible population is denoted by $R_0$ and called the basic reproductive ratio or basic reproductive number.

If $R_0 \ll N$ and if we assume uniform mixing of the population, then practically all contacts of infectious hosts during the first few generations will be with susceptibles, and we can assume, as long as $k$ is sufficiently small, that $R_0 \approx R_k$, where $R_k$ denotes the mean number of secondary infections caused by a host in the $k$-th generation.

Thus our best guess at $R_0$ would be $R_0 \approx \frac{5}{6} < 1$ in Example 1 and $R_0 \approx \frac{15}{8} > 1$ in Example 2.

# $R_0$ makes the difference

**Theorem**

*Assume uniform mixing and introduction of a single infectious host into an entirely susceptible large population. Assume, moreover, that $R_0$ does not depend on $N$.*

*If $R_0 < 1$, then the expected number of hosts that eventually experience infection is bounded by a constant that depends only on $R_0$ but not on $N$, and the disease is predicted to quickly die out.*

*If $R_0 > 1$, then with probability $> 0$ an epidemic whose final size is at least a fraction $F(1, N) > 0$ that depends only on $R_0$ will occur.*

**"Proof":** Under the assumption the expected number of infecteds in generation $k$ satisfies $E(g_k) = R_0 R_1 \dots R_{k-1} \leq R_0^k$, since $R_k \leq R_0$. Thus $E(\lim_{t\to\infty} N - S(t)) \leq \sum_{k=0}^{\infty} R_0^k = \frac{1}{1-R_0}$.

If $R_0 > 1$, then $E(g_k) \approx R_0^k$ for small $k$. More generally, $R_k \geq 1$ until a significant fraction of susceptibles move to the $I$- or $R$-compartments. An epidemic will occur with positive probability.

# Our next goals

We want to build models that don't suffer from the drawbacks of ODE models but still allow us to make meaningful predictions. We are aiming at models that

- Take into account the stochastic nature of disease transmission.
- Allow us to consider differences between individual hosts instead of assuming that each host is "average."
- Are still tractable.
- Can be defined in terms of relatively few parameters that can be reasonably well estimated from data.

This may be too much to ask for. We will need to strike a reasonable compromise between the first two and the last two items on the wish list.

In particular, we will still need to make some simplifying assumptions.

# Building stochastic process models: basics

Let us assume a fixed population of size $N$ and restrict our description to $SIR$-models.

Examples 1 and 2 suggest that stuff happens at times $T_I^i$ and $T_R^i$, which are random variables. One can conceptualize disease dynamics as a stochastic process that moves hosts around between the compartments.

At any given time, a r.v. $x_i(t)$ associated with host number $i$ can take values $x_i(t) \in \{S, I, R\}$, depending on whether $t < T_I^i$, or $t_I^i \leq t < T_R^i$, or $T_R^i \leq t$.

The state of the population at time $t$ is the vector $\vec{x}(t) = (x_1(t), \ldots, x_N(t))$.

The state of the population changes randomly over time. We will assume that for any given $\vec{x}(t)$ and any $\Delta t > 0$, the probability distribution of states $\vec{x}(t + \Delta t)$ that depends only on $\vec{x}(t)$ and $\Delta t$. Our processes will have the Markov Property.

# Mathematical *vs.* agent-based stochastic process models

Under these assumptions, building a mathematical model of the stochastic process boils down to finding formulas that specify how the probability distributions of future states $\vec{x}(t + \Delta t)$ depend on the current state $\vec{x}(t)$ and on $\Delta t$.

Instead of building mathematical models, in our case one can construct so-called agent-based models. These are computer programs that embody the stochastic process and allow us to simulate it instead of deriving mathematically rigorous predictions. This is often much easier. In the lab, we will use the code to explore this type of models.

In our agent-based models, the representation of the r.v. $x_i$ in the computer is called an agent, and building the model boils down to finding precise rules, usually **If** ... **then** ... rules, for how the states of the agents change over time.

# Building agent-based models: $x_i(t) \in I \cup R$

**If** $x_i(t) = R$, **then** $x_i(t + \Delta t) = R$. No randomness here.

**If** $x_i(t) = I$, **then** $x_i(t + \Delta t) = I$ or $x_i(t + \Delta t) = R$.

This is not a complete rule. We have two possibilities that will occur with some probability. In each run of a simulation, the computer needs to make a decision. A more complete rule would look like this:

Randomly generate a time $T_R^i$.
**If** $x_i(t) = I$ and $t + \Delta t < T_i^R$, **then** $x_i(t + \Delta t) = I$.
**If** $x_i(t) = I$ and $t + \Delta t \geq T_i^R$, **then** $x_i(t + \Delta t) = R$.

**I called this a "more complete" rule. Is there still something missing?**

# What, exactly, does "randomly generate" mean?

The phrase "randomly generate" does not have an absolute meaning. We need to specify the probability distribution from which the random object (the time of removal $T_i^R$ in our case) is to be drawn.

**Which distribution should we use here?**

By our assumption of the Markov property, the distribution of the time $T_R^i - t$ until removal should depend only on the current state, not on how long ago host $i$ became infectious. Mathematicians would say that the r.v. $T_R^i - T_I^i$ is memoryless.

Memoryless continuous r.v.s have exponential distributions. Thus we will assume that

$$P(T_R^i \geq T_I^i + \Delta t) = e^{-\alpha_i \Delta t}.$$

Here $\alpha_i$ is a parameter of the model. It is equal to the reciprocal of the mean value of $T_i^R - T_I^i$.

# A challenge question

The assumption that $T_R^i - T_I^i$ is a memoryless r.v. turned out to be very convenient, but it is blatantly wrong for most diseases. Recovery times show usually a distribution that peaks at some modal value. For example, you are much more likely to recover during day 7 of a bout of the flu than during day 2, while an exponential distribution would predict the opposite.

**Challenge Question:** How can we modify the model so that the distribution of recovery times becomes more realistic without sacrificing the Markov Property of the process?
*Hint:* Introduce additional compartments.

# Building agent-based models: $x_i(t) = S$

**If** $x_i(t) = S$, **then** $x_i(t + \Delta t) = S$ or $x_i(t + \Delta t) = I$ or $x_i(t + \Delta t) = R$.

We can turn this into a more complete rule as follows:

Randomly generate a time $T_I^i$.
Randomly generate a time $T_R^i > T_I^i$.
**If** $x_i(t) = S$ and $t + \Delta t < T_I^i$, **then** $x_i(t + \Delta t) = S$.
**If** $x_i(t) = S$ and $T_I^i \leq t + \Delta t < T_i^R$, **then** $x_i(t + \Delta t) = I$.
**If** $x_i(t) = S$ and $T_i^R \leq t + \Delta t$, **then** $x_i(t + \Delta t) = R$.

The hard part of course is how to randomly generate the time $T_I^i$ when host $i$ becomes infectious.

# How to randomly generate $T_I^i$

If $x_i(0) = S$, then host $i$ can become infectious only through direct contact with another host $j$ who will be infectious at the time of the contact, while $i$ is still susceptible. Not all such contacts will lead to infection of host $i$; transmission of a sufficient number of pathogens during the contact is required.

Consider a hypothetical situation where $i$ and $j$ are the only hosts, $x_i(T_I^j) = S$, and $T_R^j = \infty$ (as in an $SI$-model). If the probability of subsequent "successful" contact between $i$ and $j$ is positive, then there will be a time $T_I^{i,j} > T_I^j$ when host $i$ gets infected by host $j$.

The Markov property implies that $T_I^{i,j} - T_I^j$ is a memoryless r.v. with an exponential distribution.

$$P(T_I^{i,j} \geq T_I^j + \Delta t) = e^{-\beta_{i,j}\Delta t}.$$

The number $\beta_{i,j}$ is another parameter of the model. It is equal to the reciprocal of the mean value of $T_I^{i,j} - T_I^j$.

# How to randomly generate $T_I^i$, continued

In actual simulations there will usually be more than two hosts, and infectious hosts will eventually be removed. However, we can still use the formula

$$P(T_I^{i,j} \geq T_I^j + \Delta t) = e^{-\beta_{i,j}\Delta t}$$

to generate hypothetical times $T_I^{i,j}$.

If $x_i(0) = S$ and host $i$ becomes eventually infectious, then we must have
$T_I^i = T_I^{i,j}$ for some $j$.

**How can we identify the correct $T_I^{i,j}$?**

It has to be the smallest among those $T_I^{i,j}$ that satisfy $T_I^{i,j} < T_R^j$.

# We have built a model!

The rules we have described in the previous slides define an agent-based model. Recall that this is an algorithm that embodies a stochastic process and allows to simulate it.

Essentially, the algorithm will randomly generate transition times $T_I^i$, $T_R^i$ and simulate the process by changing the states of the agents at these times accordingly.

These transition times need to be generated in a certain order and sometimes updated. This involves careful bookkeeping, which is not entirely trivial but fairly routine. I will omit these details here.

# Some properties of our model

- The model is a continuous-time stochastic process. Transition times $T_I^i$, $T_R^i$ can take any nonnegative real values, but nothing of interest happens in between these times.
- We have made a number of simplifying assumptions along the way; not all of which were spelled out explicitly. When trying to make inferences about actual outbreaks of diseases, one always needs to carefully examine these assumptions.
- The parameters of the model are $\alpha_i$ and $\beta_{i,j}$. The former allow us to incorporate some amount of heterogeneity of hosts, while the latter give us considerable flexibility in modeling the mixing pattern.

**What might influence $\alpha_i$?**

**What might influence $\beta_{i,j}$?**

**We have implicitly assumed that $\alpha_i$ and $\beta_{i,j}$ do not depend on the current state of the system. Is this assumption justified?**

# Have we accomplished our goal?

**Did we build models that:**

- Take into account the stochastic nature of disease transmission? Yes!
- Allow us to consider differences between individual hosts? Yes!
- Are still tractable. Yes, at least in terms of simulations.
- Can be defined in terms of relatively few parameters that can be reasonably well estimated from data? No!

There are just too many parameters. We would need $N$ parameters $\alpha_i$ plus $N(N-1)$ parameters $\beta_{i,j}$. For realistic population sizes, it is impossible to estimate that many parameters from any kind of data set.

How can we reduce the number of parameters that need to be estimated from the data without sacrificing the major advantages of our models?

# Reducing the number of parameters: homogeneity of hosts

Having a separate parameter $\alpha_i$ for each host allows us to consider heterogeneity of hosts in the sense of individual host's propensity to recover more or less quickly from the disease. Biologically speaking, the parameter $\alpha_i$ may correspond to the strength of host $i$'s immune system or general state of health.

However, even if $\alpha_i = \alpha_j$, the actual duration of infectiousness $T_R^i - T_I^i$ for host $i$ may be significantly different from $T_R^j - T_I^j$ for host $j$, as $\alpha_i$ only is a parameter of a probability distribution. There always will be some variability in actual durations of infectiousness.

It thus seems plausible that the predictions of the model will not change much if we assume that $\alpha_i = \alpha$ for some fixed single parameter $\alpha$. We will from now on make this assumption.

Since $\frac{1}{\alpha}$ represents the mean duration of infectiousness for the population, the parameter $\alpha$ should be relatively easy to estimate from data.

# Reducing the number of parameters: the uniform mixing assumption

The parameters $\beta_{i,j}$ depend both on how frequently hosts $i$ and $j$ tend to make contact, and on the intensity of the contacts, which translates into the probability that sufficiently many pathogens get transferred during a single contact.

These two aspects often, but not always go together. Mr. Jones may spend more time with his boss than with his wife, but will kiss and hug her more often than his boss, or so one would hope.

The uniform mixing assumption postulates that for any pair $(i, j)$ of different hosts both the frequency and intensity of contacts is the same. In the context of our models, this translates into assuming that $\beta_{i,j} = \beta$ for some fixed single parameter $\beta$.

The parameter $\beta$ should also be relatively easy to estimate from data. If we make both assumptions of homogeneity of hosts and uniform mixing, we end up with models that have only two parameters: $\alpha$ and $\beta$. **Looks familiar?**

# Is the uniform mixing assumption realistic? What if it isn't?

Mixing may be nearly uniform if hosts move around a lot relative to the size of the habitat, encounter each other rarely, and there is no social structure.

In populations with a well-defined social or territorial structure though, some pairs of individuals will have contact relatively frequently (think of co-workers or neighbors in human populations), while other pairs of individuals will almost certainly never encounter each other (think of your likelihood to ever meet the Supreme Leader of North Korea).

We can approximate the latter situation by assuming the existence of a contact network which determines whether it is even possible that the disease can be transmitted between two given hosts. The nature of the required contact, and thus the relevant contact network, may depend on the particular disease. Think of the flu *vs.* a computer virus *vs.* a sexually transmitted disease.

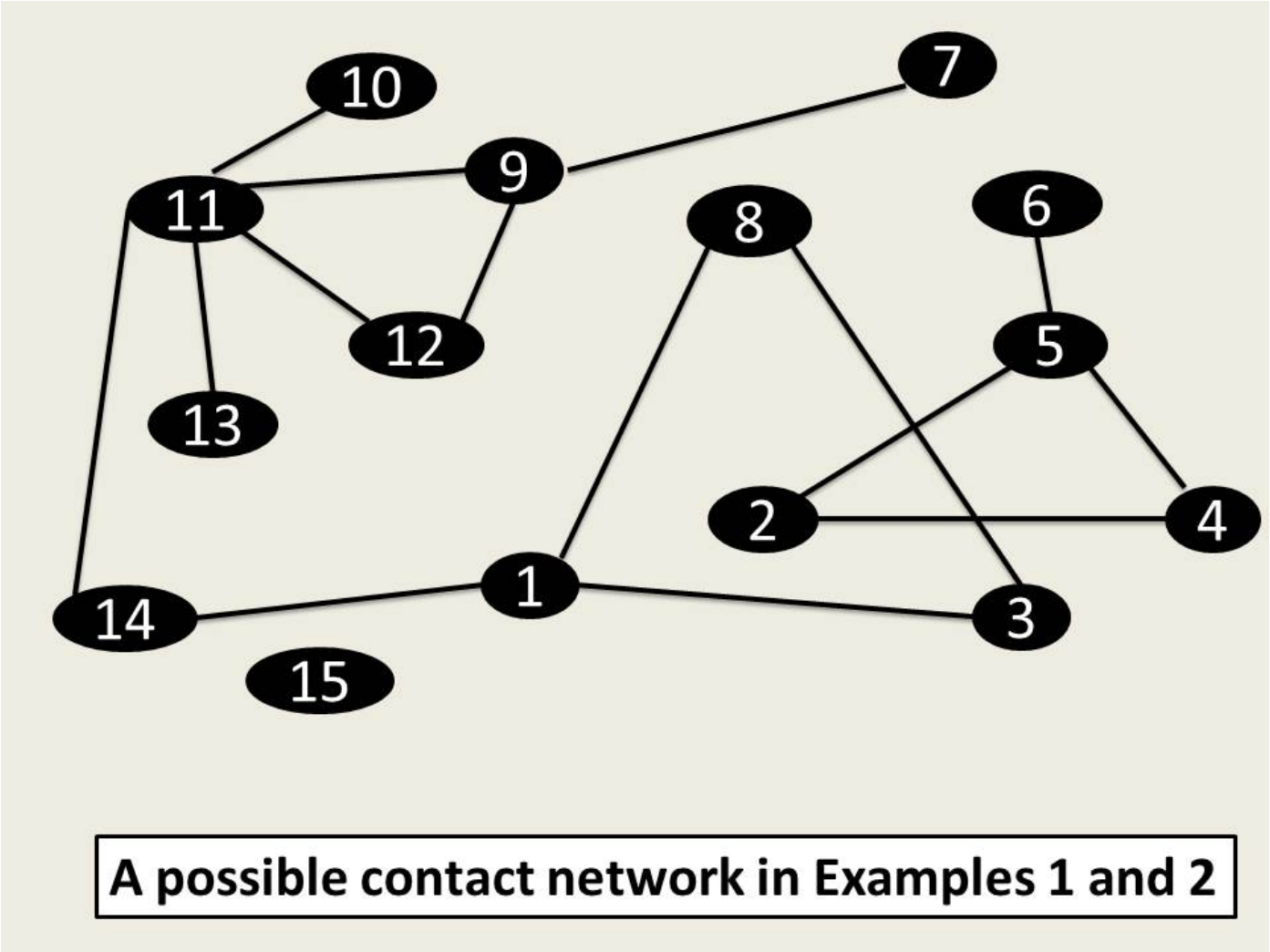
# Mathematical structures for modeling contact networks: graphs

A graph is an ordered pair $G = (V, E)$, where $V$ denotes the set of vertices, or nodes, and the set $E$ of edges of $G$ is a subset of the set of unordered pairs of nodes.

A contact network can be modeled as a graph whose vertices are the individual hosts in the population, and an edge between two hosts signifies an above-threshold probability of a relevant contact between these two hosts.

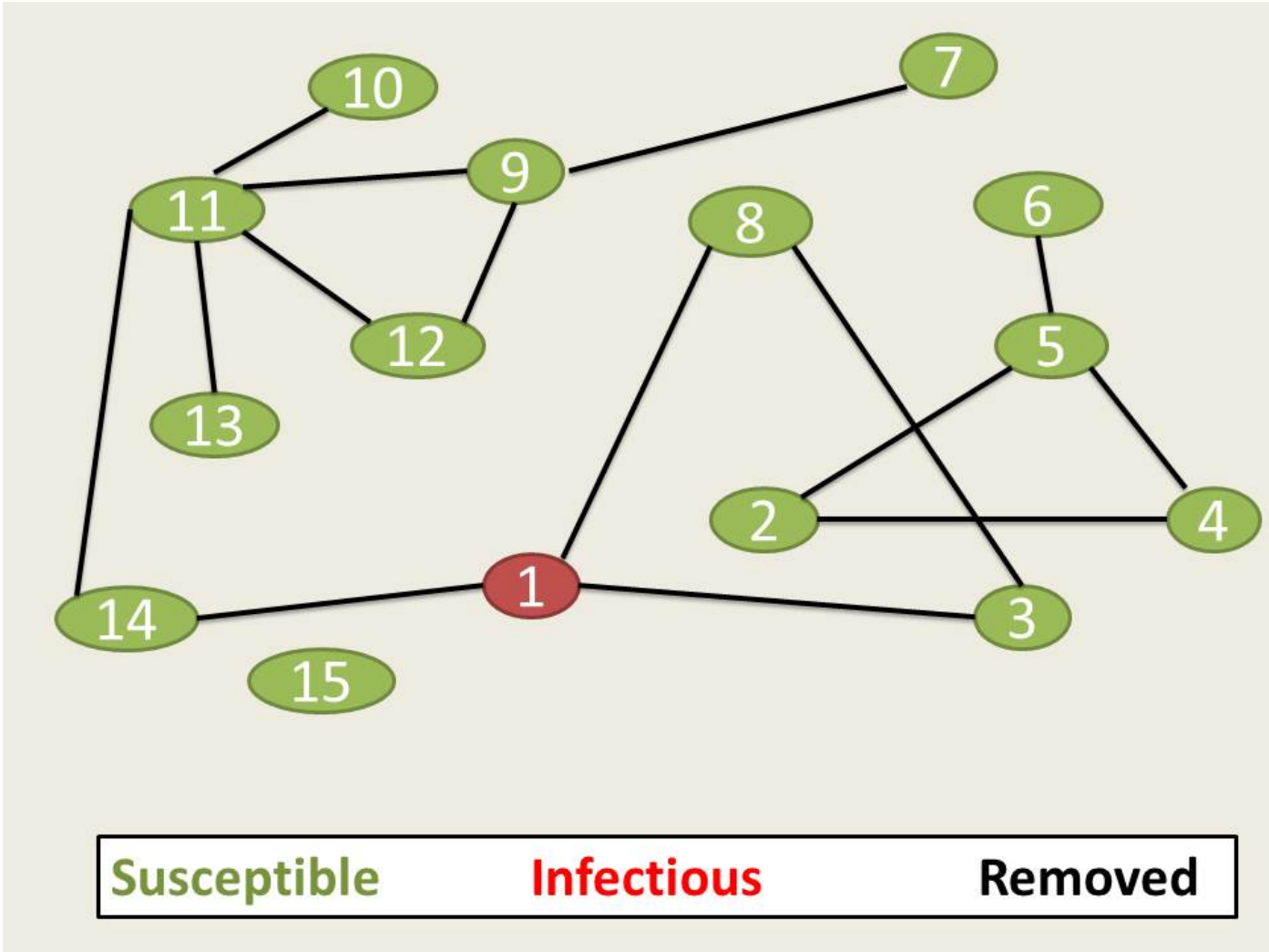
One can then assume that disease transmission can occur only between two hosts that are represented by adjacent nodes, that is, endpoints of a common edge, and study the possible or likely dynamics of the disease on the network.

# An example of a graph

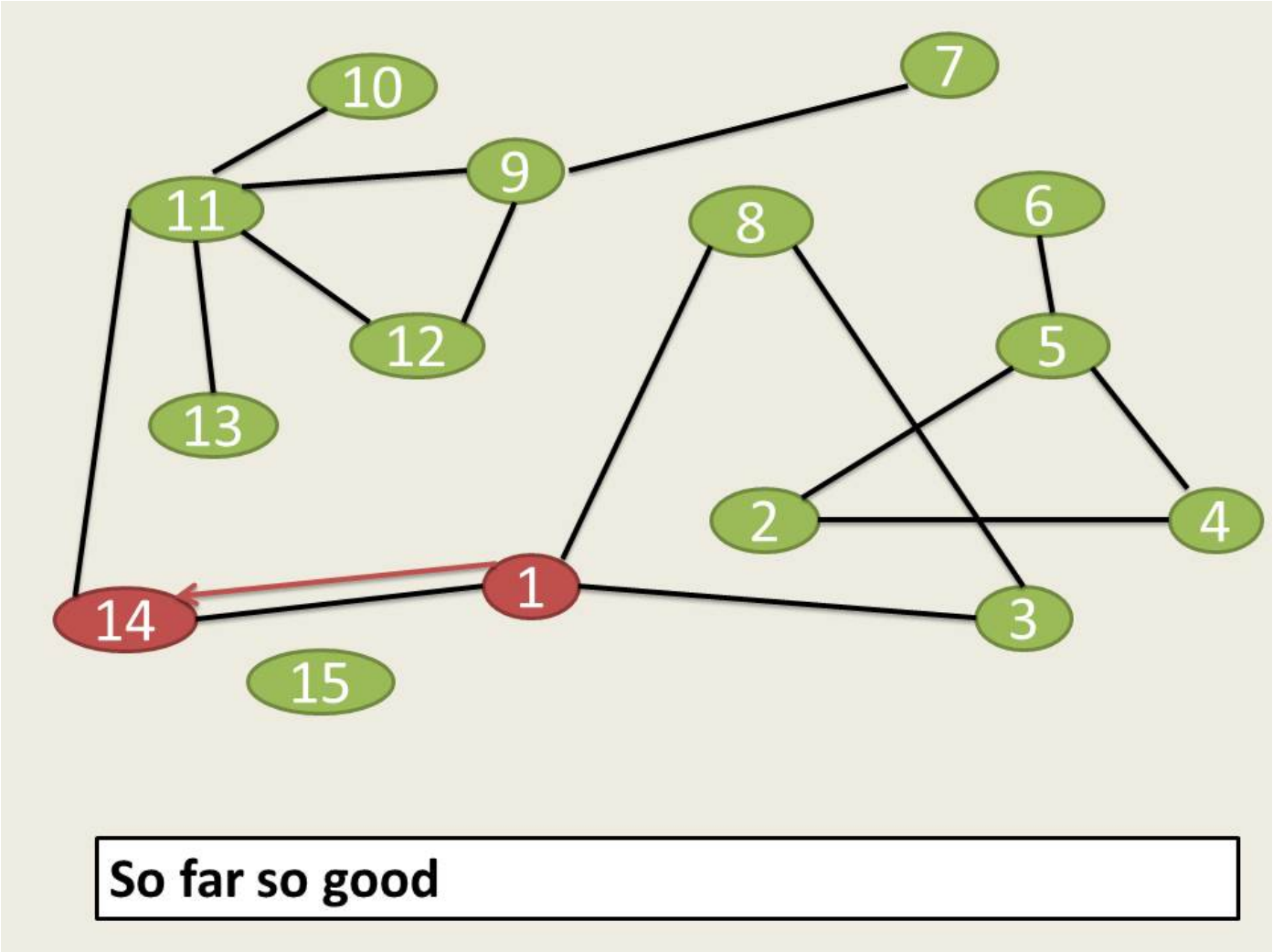



A possible contact network in Examples 1 and 2

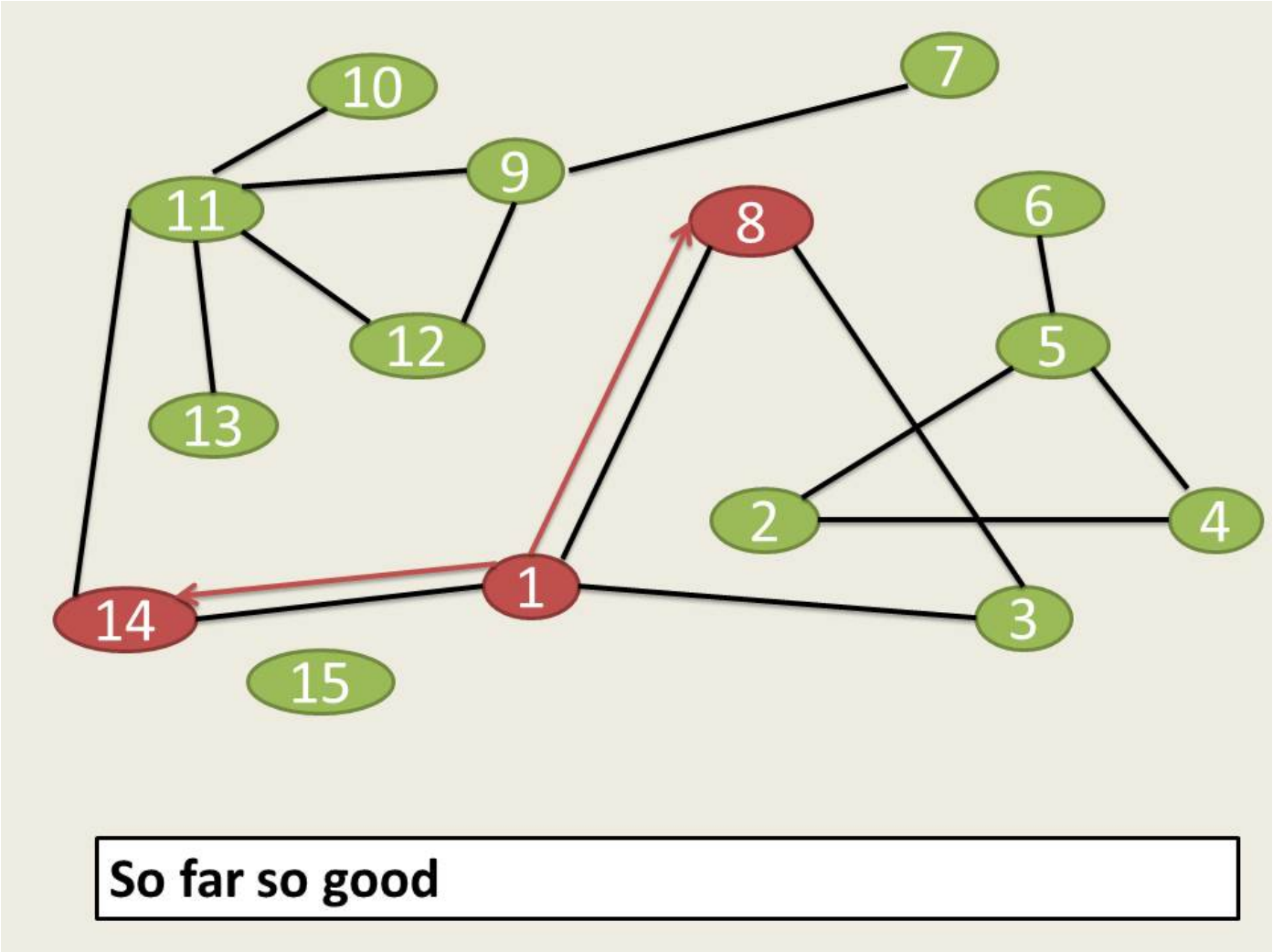
# Could Example 1 occur on this contact network?



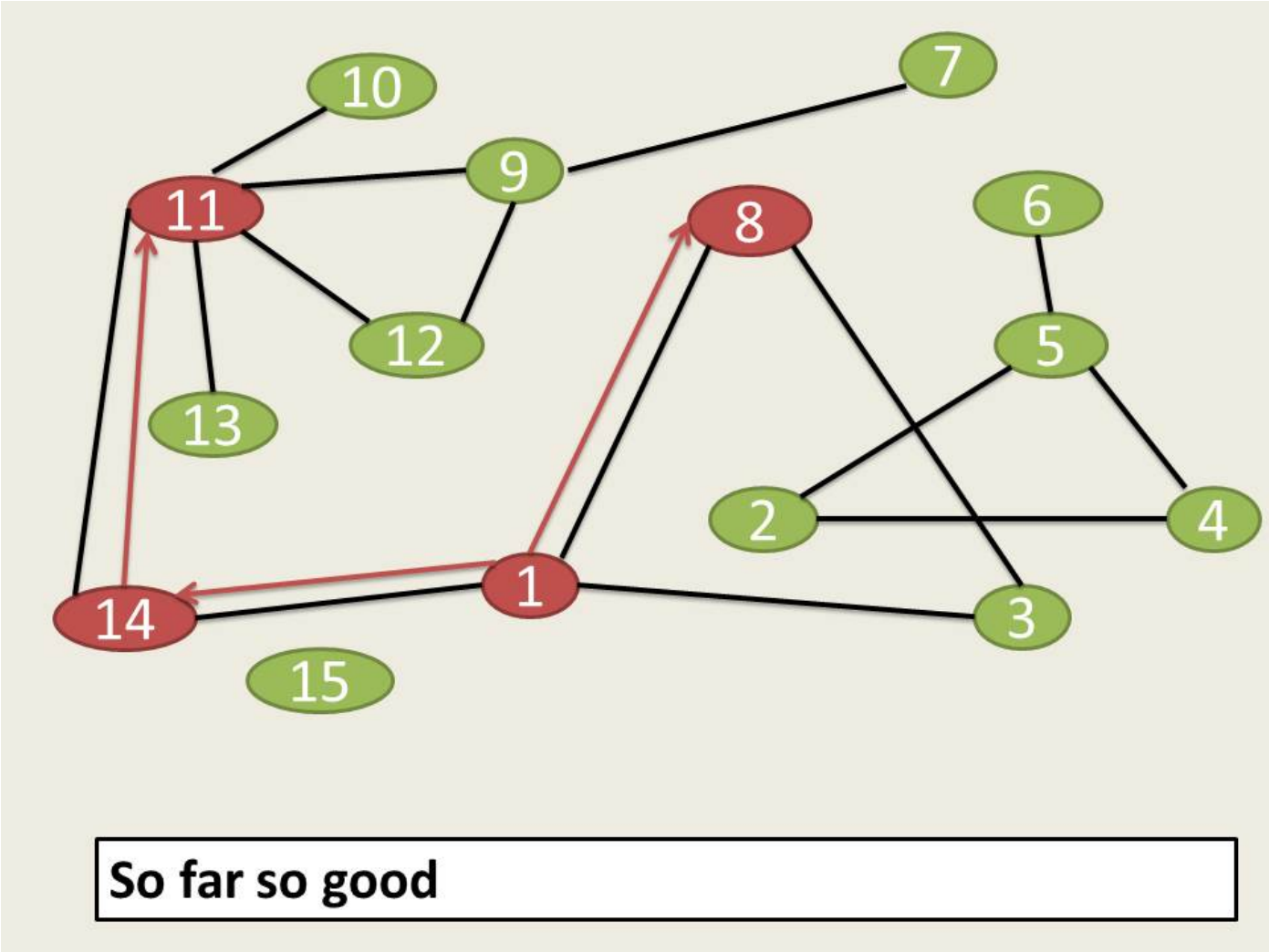
# Could Example 1 occur on this contact network?



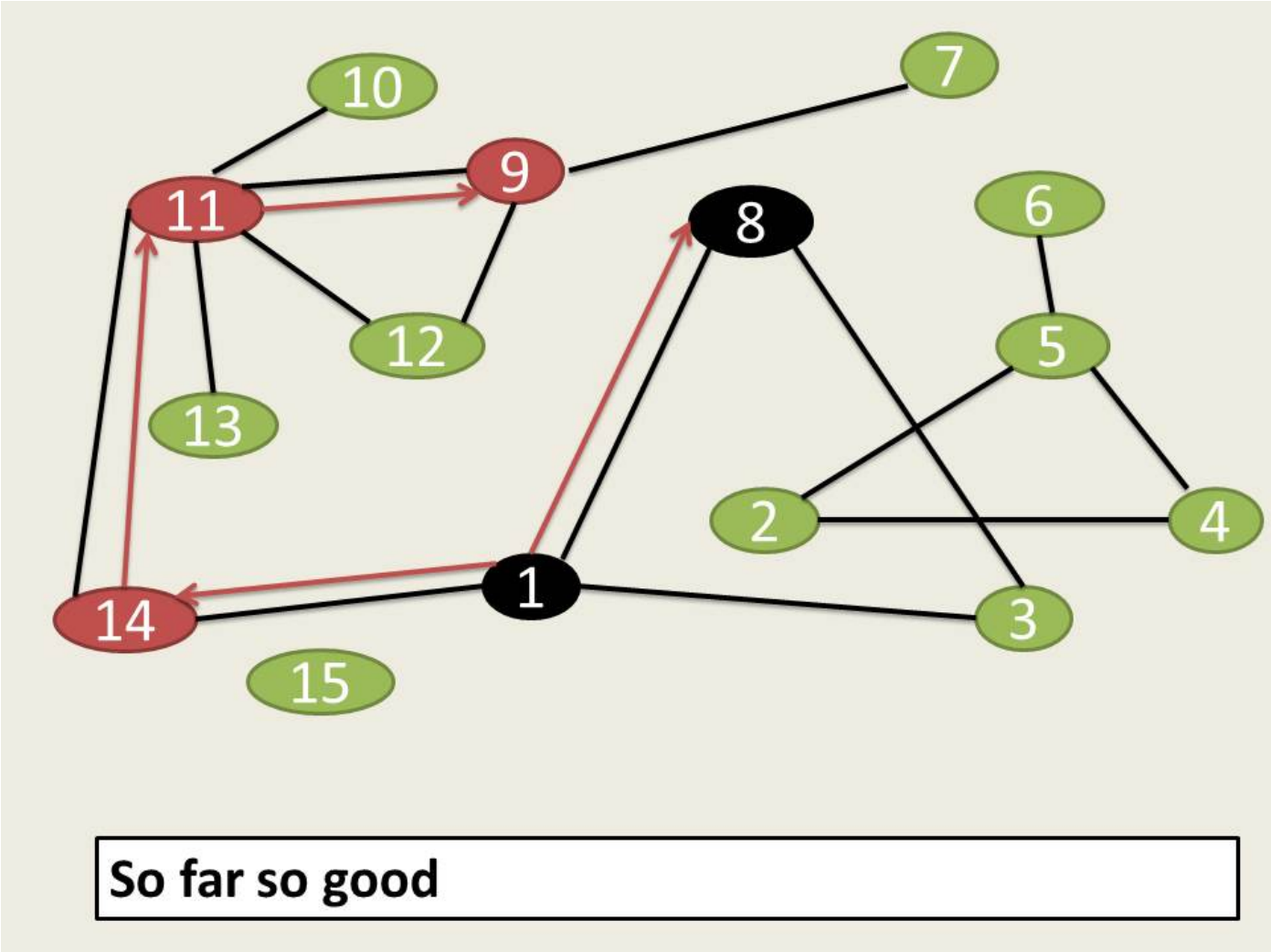
# Could Example 1 occur on this contact network?



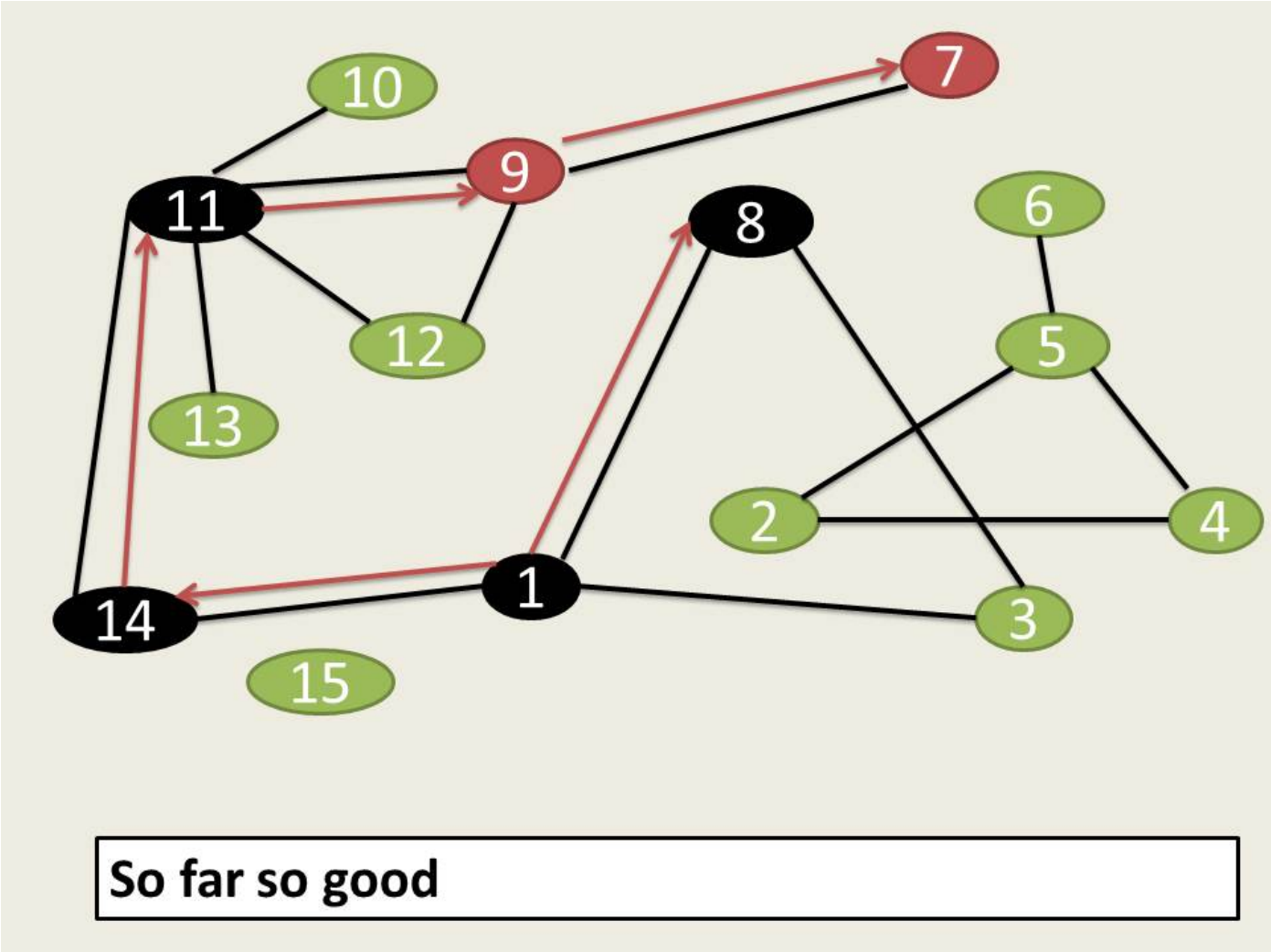
# Could Example 1 occur on this contact network?



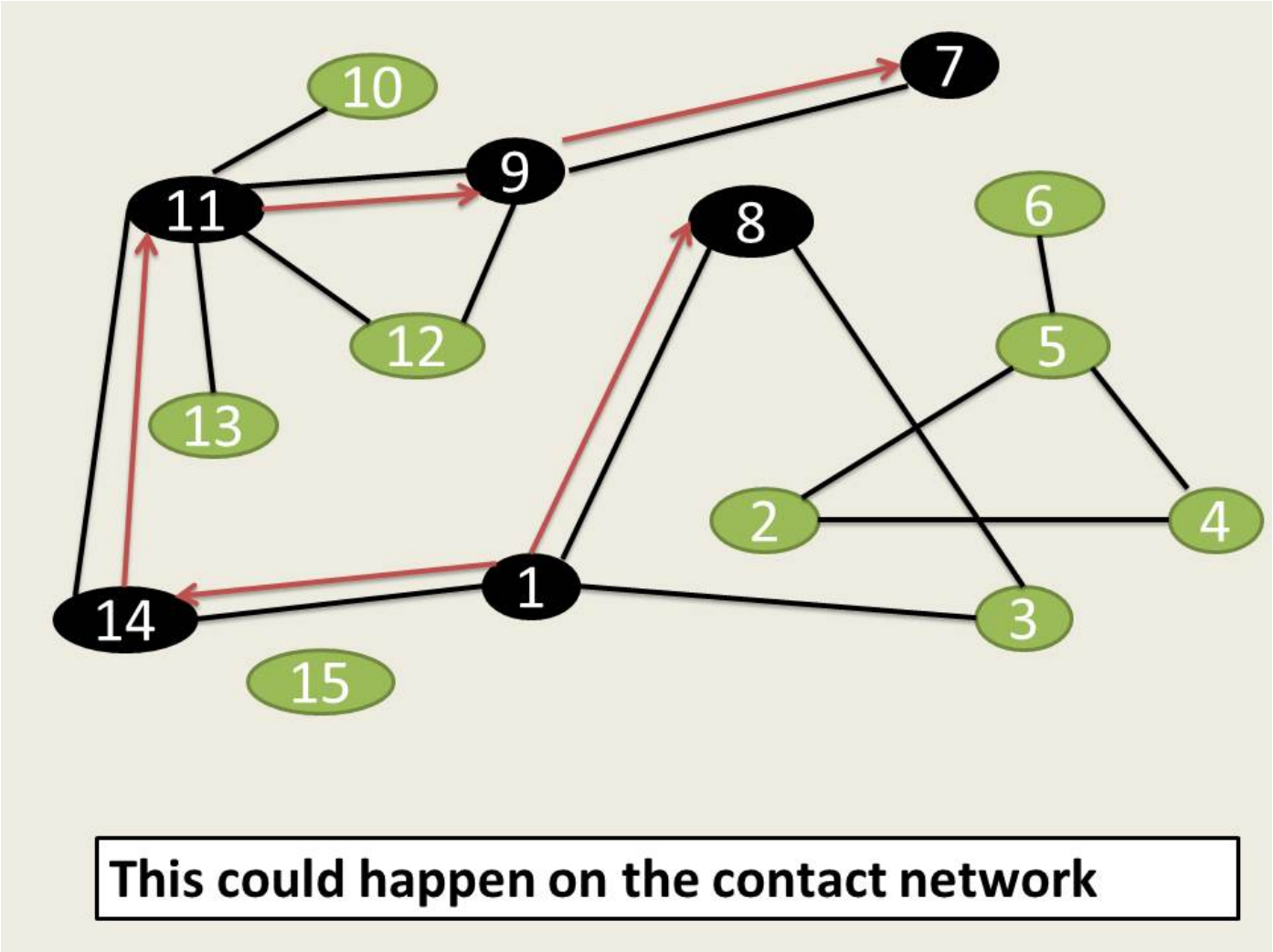
# Could Example 1 occur on this contact network?



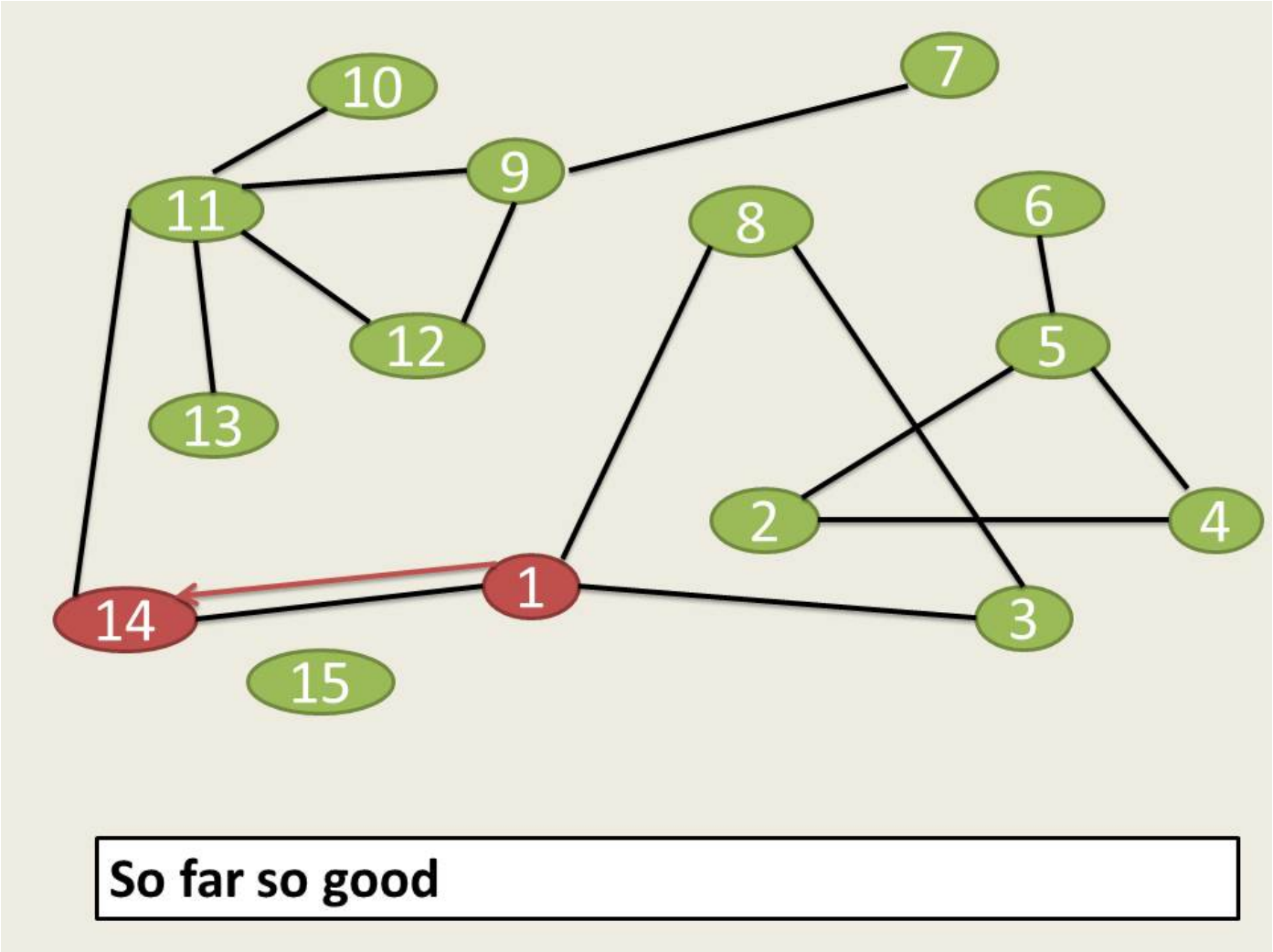
# Could Example 1 occur on this contact network?



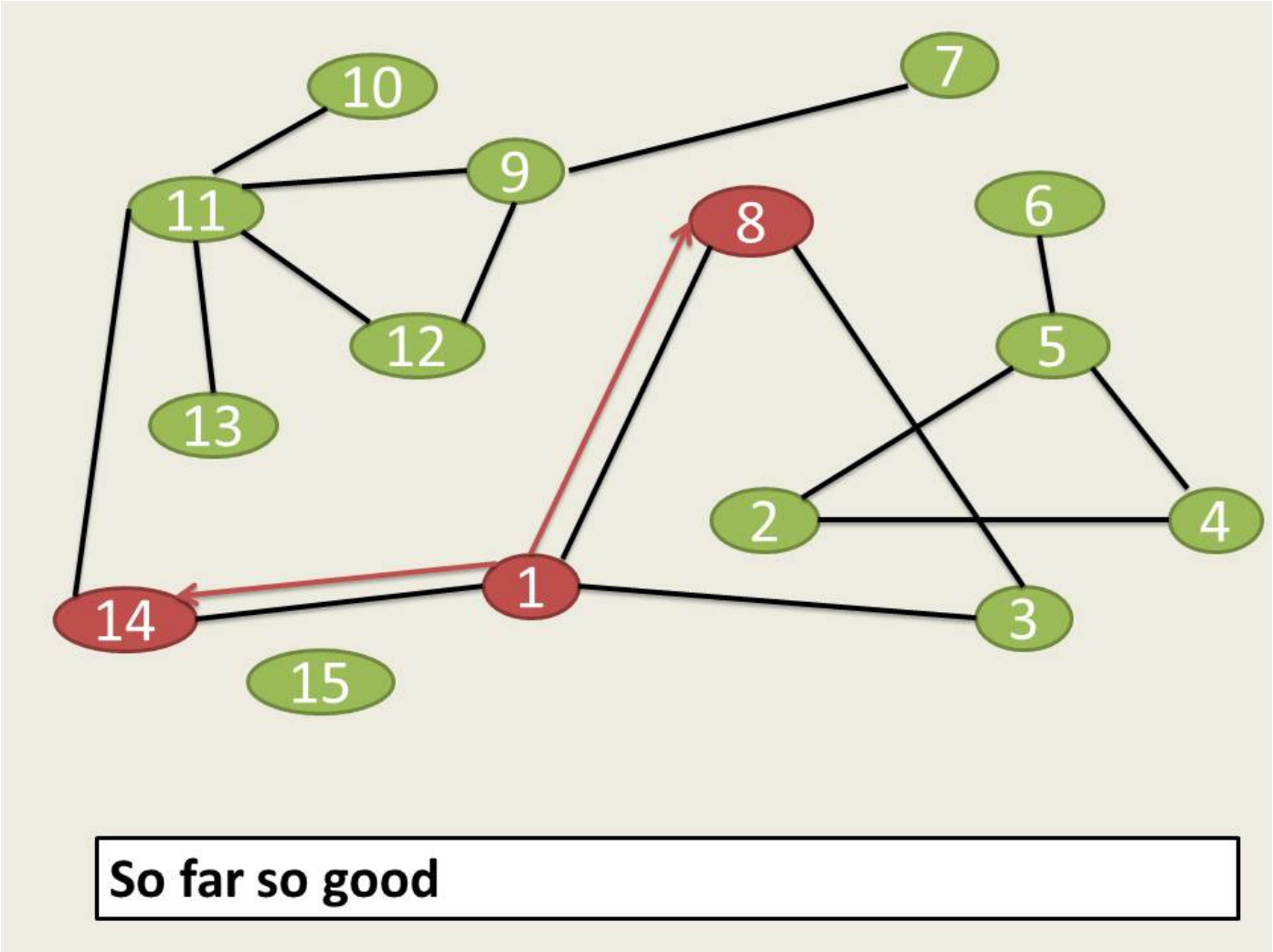
# Could Example 1 occur on this contact network?



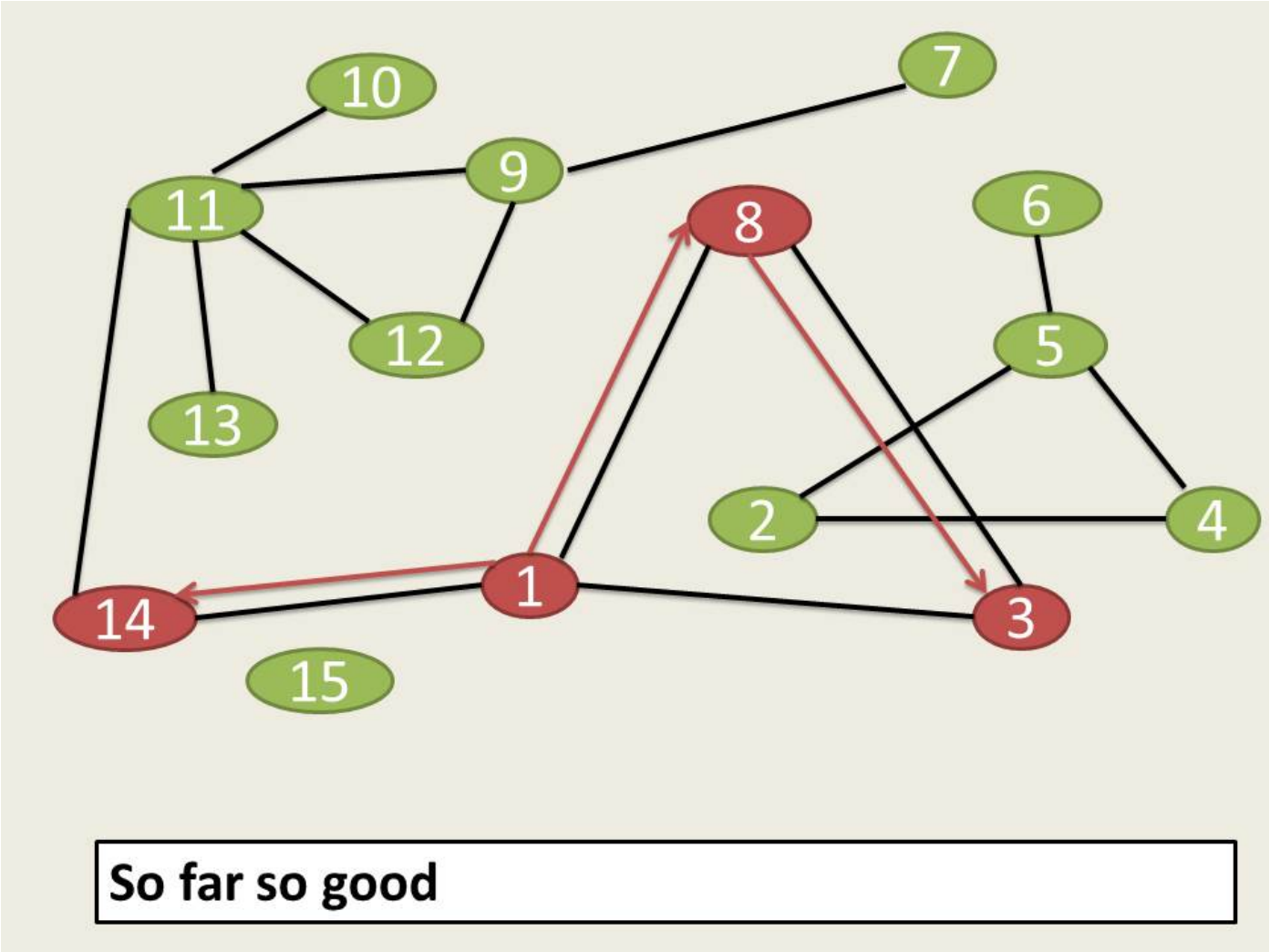
# Could Example 2 occur on this contact network?



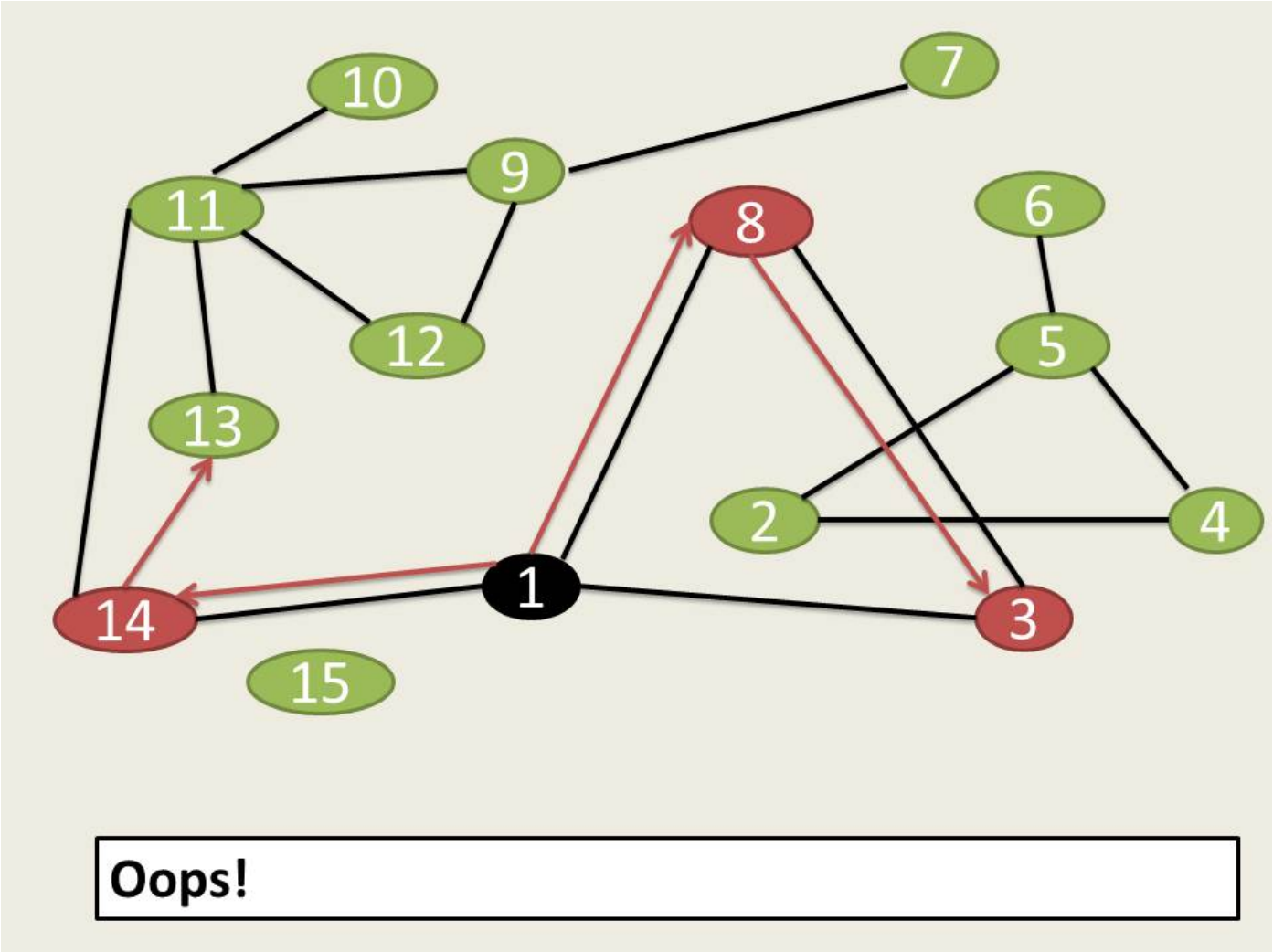
# Could Example 2 occur on this contact network?



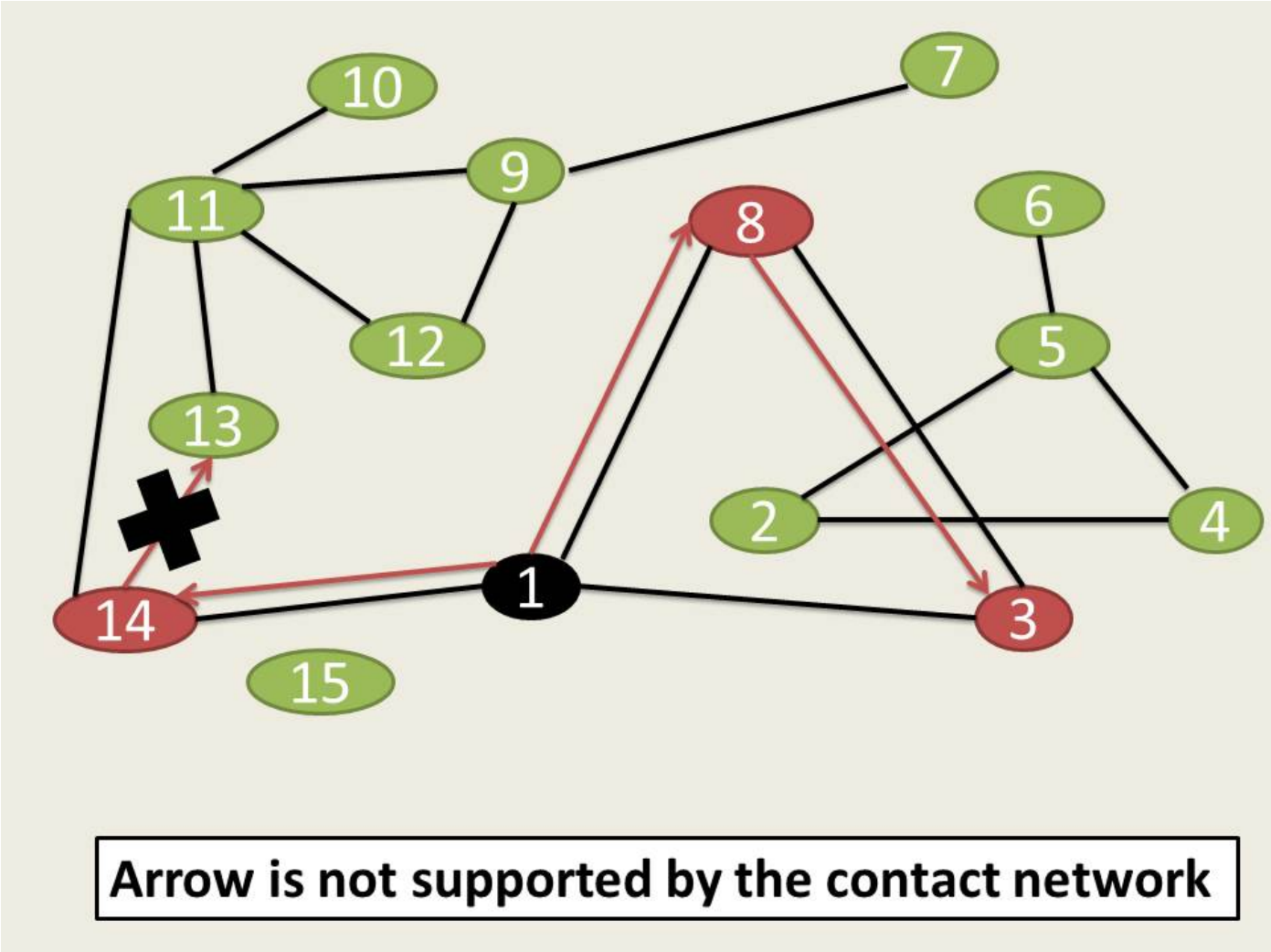
# Could Example 2 occur on this contact network?



# Could Example 2 occur on this contact network?

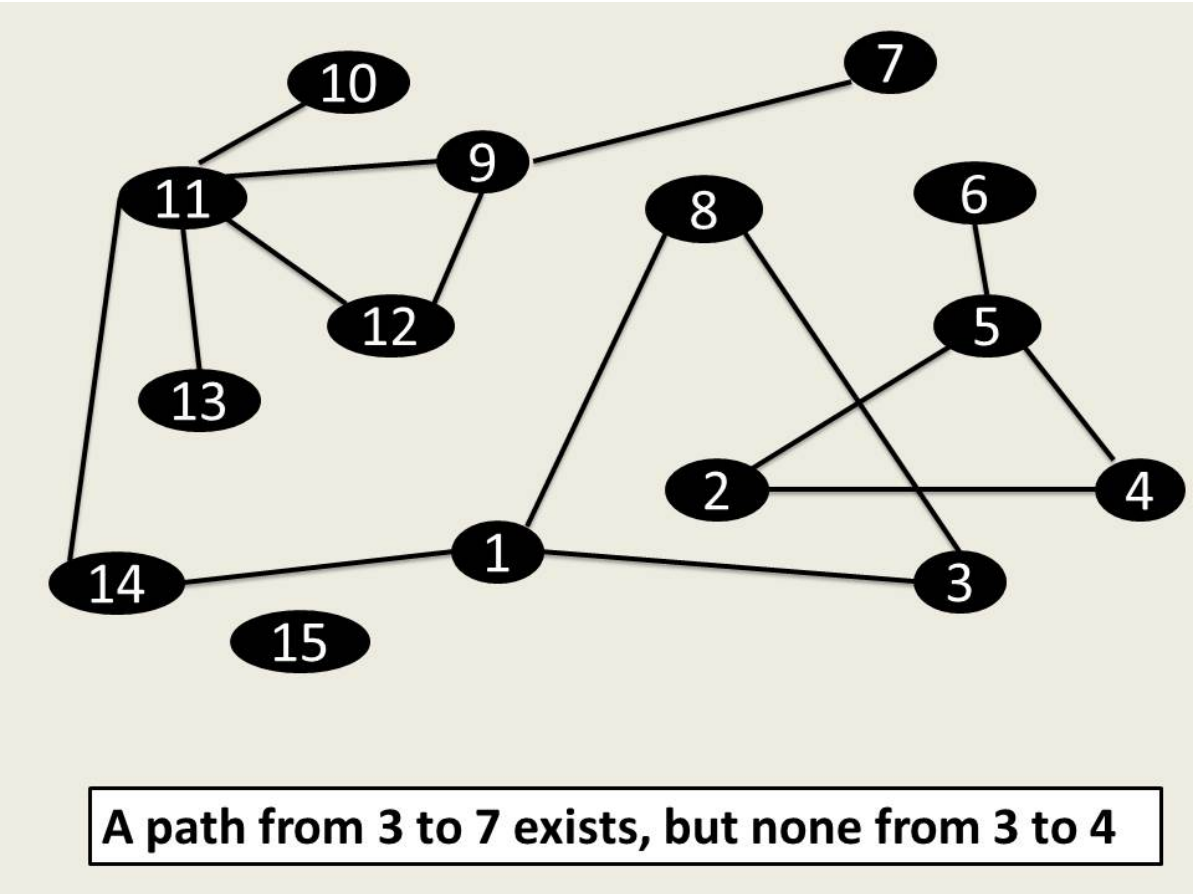


# Example 2 could not occur on this contact network



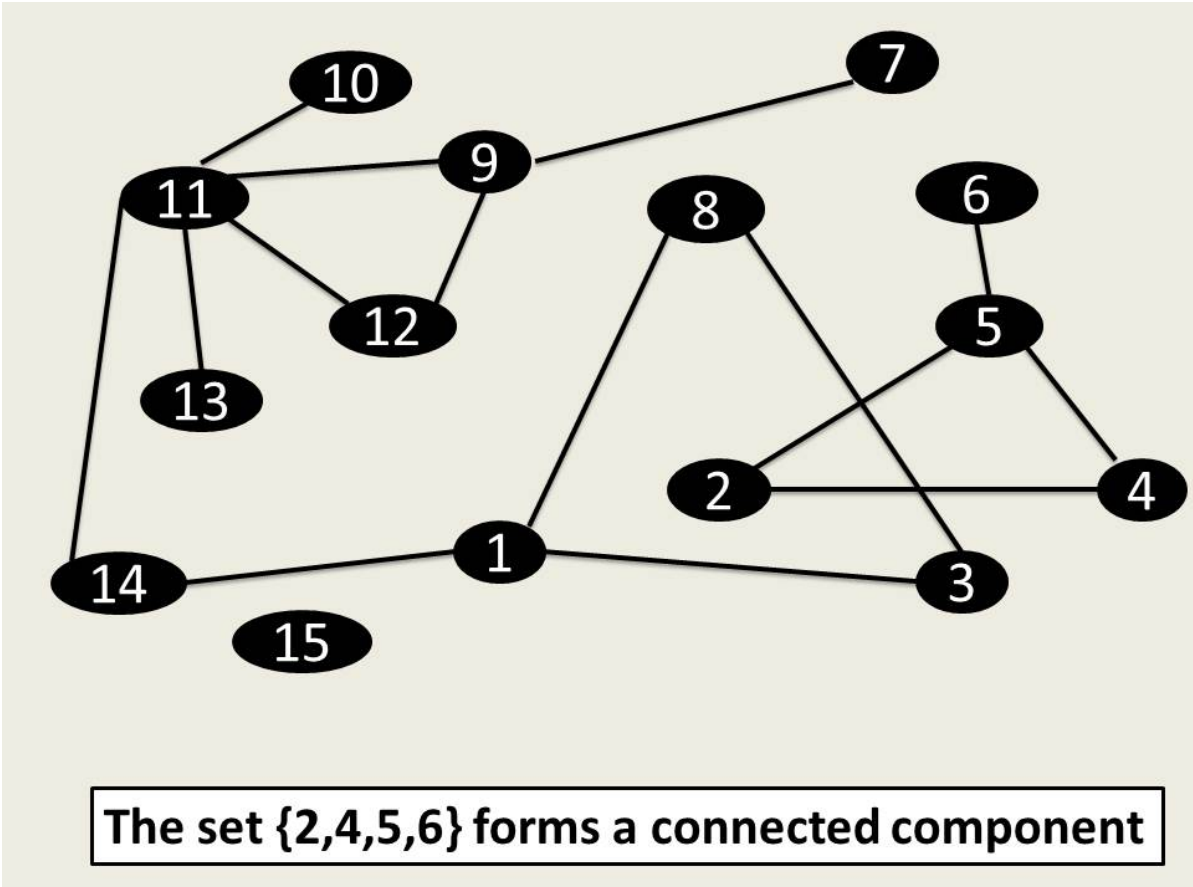
# Some graph jargon: paths

A path from vertex $v_1$ to vertex $v_n$ is a sequence $v_1 v_2 \dots v_n$ of pairwise distinct vertices such that each $\{v_i, v_{i+1}\} \in E$. If $v_1$ is infected and $v_2, \dots, v_n$ are all susceptible, then $v_n$ may eventually become infected.



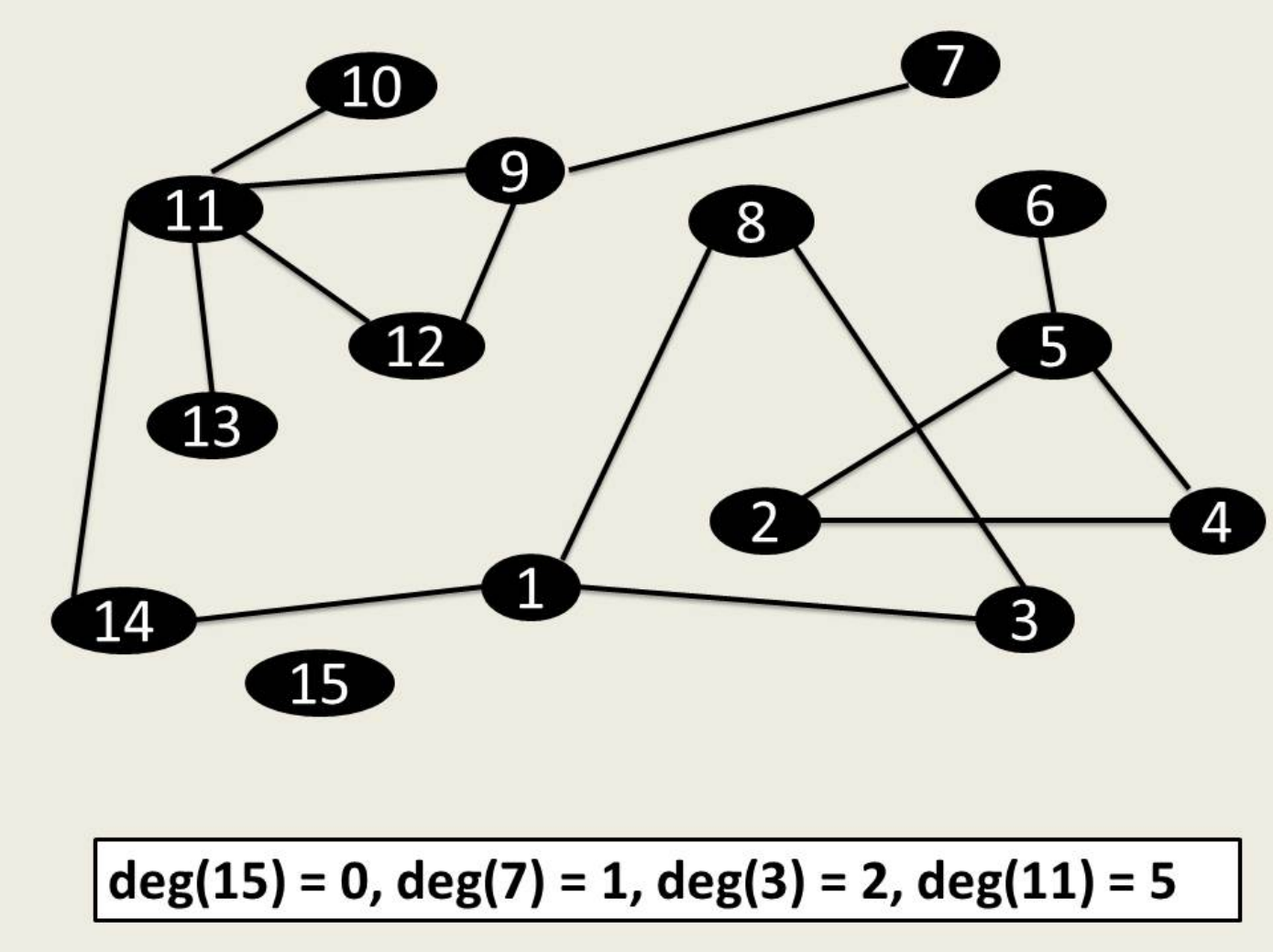
# Some graph jargon: connected components

The connected component of a vertex $v$ is the set of vertices that comprises $v$ together with all $w$ such that there exists a directed path from $v$ to $w$. If $v$ and $w$ belong to different connected components, then infection of $v$ in an otherwise susceptible population cannot lead to infection of $w$.



# Some graph jargon: degrees

The degree of a vertex $v$ is the number of adjacent vertices, that is, vertices $w$ with $\{v, w\} \in E$.



# Stochastic process models for disease transmission on networks

Ingredients:

- Specification of the type of model (*SI*, *SIR*, or *SIS*).
- A graph $G = (V, E)$ with $N$ vertices that represents the contact network.
- A parameter $\alpha$ that represents the removal or recovery rate.
- A parameter $\beta$ that specifies the rate at which a given susceptible host acquires infections from a given adjacent infectious host.

The process will then be modeled exactly as before, with

$$\beta_{i,j} = \begin{cases} \beta & \text{if } \{i,j\} \in E, \\ 0 & \text{if } \{i,j\} \notin E. \end{cases} \tag{1}$$

The uniform mixing assumption corresponds to the case where $G$ is the complete graph that contains all possible edges.

# What is to be gained from network models of disease transmission?

While the kind of network models we defined here are more realistic than models based on the uniform mixing assumption, they still rely on a lot of simplifying assumptions. But they can give us some valuable insights:

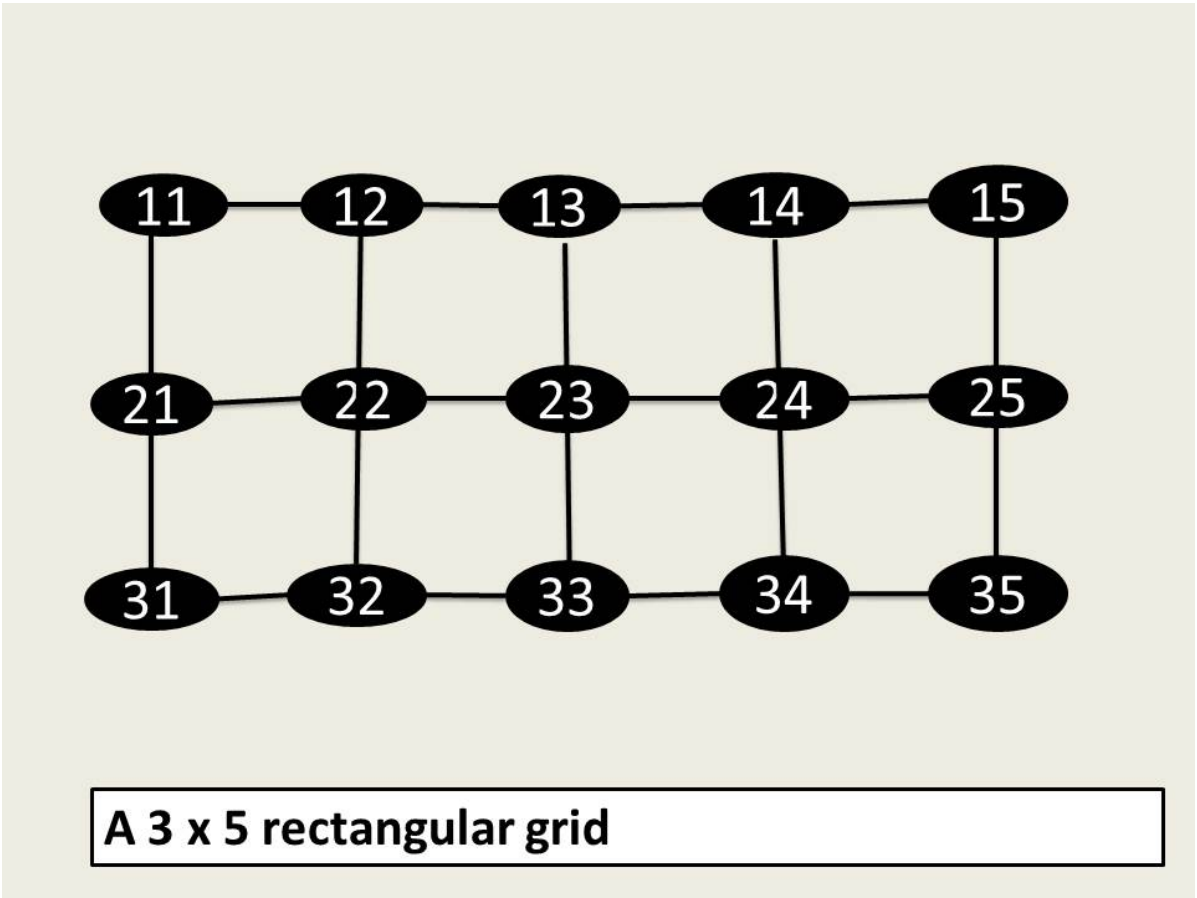
- They can make predictions about the probability of an epidemic that are not available from ODE models.
- They may allow us to discern cases when the uniform mixing assumption is inadequate.
- They may point to features of the contact network that significantly influence the outcome of an epidemic. This gives some guidance about what kind of data we need to collect in order to be able to make reasonably accurate predictions.
- They can inform the design of effective control measures when the uniform mixing assumption is inadequate.

# But how to model the network?

A major problem is that we usually have only very limited knowledge of the actual contact network. There are basically two ways of building mathematically meaningful models of the underlying networks.

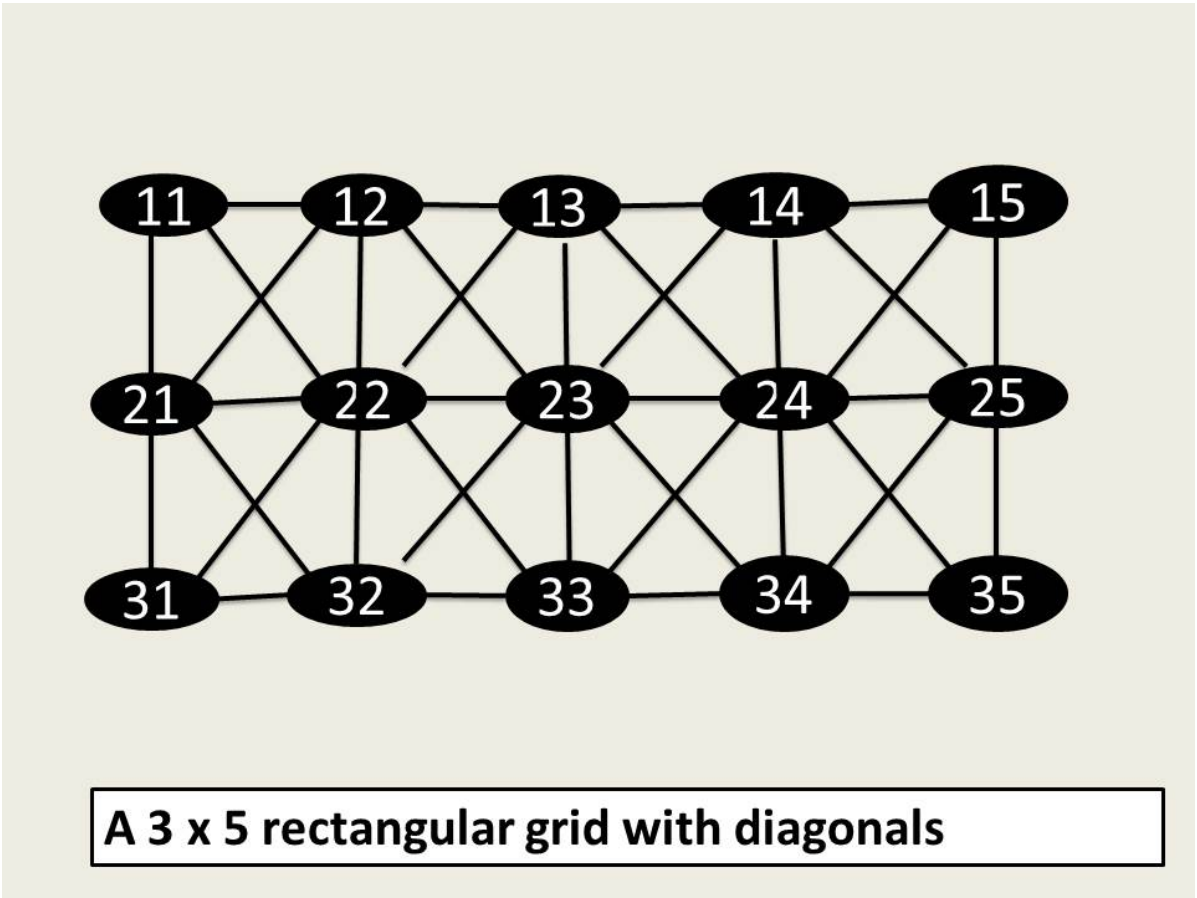
- In some cases the network may have a very special structure that can be determined from data.
- Alternatively, we can assume that the network is randomly drawn from a probability distribution with certain parameters. The values of these parameters should be chosen in such a way that they favor networks with properties that conform to whatever data we have about the actual network.

# Special case 1: $L \times M$ rectangular grids




Think of a banana plantation where the pathogen can move only by a distance of at most 1.

# Special case 2: Rectangular grids with diagonal edges




Think of a banana plantation where the pathogen can move only by a distance of at most 1.5.

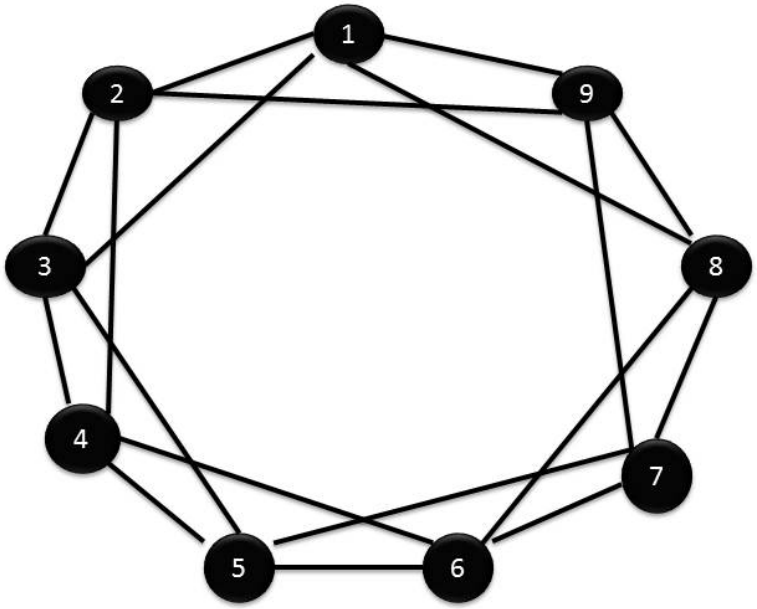
# The Sisters of the Round Table

Consider the monastic order of the Sisters of the Round Table. The sisters spend most of their lives in their individual cells, where they devote themselves to prayer and meditation. The only time they have contact with each other is during meals that they take seated in a fixed order around a giant round table.

Within this community, diseases can be can spread only during mealtime. The probability of transmission will be largest between sisters who sit next to each other, and then decrease with the distance at the table. When constructing a network model, we need to make a decision on the cutoff. Making a reasonable choice here is part of the art of modeling; there are no fixed rules.

Let us assume that there is a significant probability of transmission from sister $i$ to sister $j$ if at most one sister sits in between.

# Special case 3: The contact network of the Sisters of the Round Table



# Towards more realistic network models

I called all three of the above examples "special cases" because they assumed a very rigid structure of the contact network. In most real-world situations the situation is much more messy.

In fact, even in a strictly monastic setting, contact networks will rarely have such a rigid structure. The Sisters will not necessarily head straight to the table from their cells. More likely, along the way they will tend to exchange a few kind words with their next-cell neighbors who may be seated across the table.

**How could we incorporate these more informal contacts into our network model without knowing who occupies adjacent cells?**

By adding a few randomly chosen edges to the network of the previous slide.

**How could we produce a set of randomly chosen edges?**

# Erdős-Rényi random graphs

The most basic construction of graphs with randomly chosen edges, or random graphs, was first systematically studied by the Hungarian mathematicians Paul Erdős and Alfred Rényi. A version of it can be described as follows:

Specify the set of $N$ nodes. For each potential edge $e = \{i, j\}$, flip a (biased) coin that comes up heads with probability $p$ and include $e$ in $E$ iff the coin does come up heads.

The parameters of this model are $p$ and $N$. The mean degree in this model will be equal to $p(N-1)$.

By the Central Limit Theorem, the degree distribution should be roughly normal for large enough $N$, thus strongly peaked around $p(N-1)$.

There are other types of random graphs whose degree distributions are not close to normal, such as so-called scale-free networks.