

Editado por

Eliana X.L. de Andrade

Universidade Estadual Paulista - UNESP

São José do Rio Preto, SP, Brasil

Rubens Sampaio

Pontifícia Universidade Católica do Rio de Janeiro -

Rio de Janeiro, RJ, Brasil

Geraldo N. Silva

Universidade Estadual Paulista - UNESP

São José do Rio Preto, SP, Brasil

A Sociedade Brasileira de Matemática Aplicada e Computacional - SBMAC publica, desde as primeiras edições do evento, monografias dos cursos que são ministrados nos CNMAC.

Para a comemoração dos 25 anos da SBMAC, que ocorreu durante o XXVI CNMAC em 2003, foi criada a série **Notas em Matemática Aplicada** para publicar as monografias dos minicursos ministrados nos CNMAC, o que permaneceu até o XXXIII CNMAC em 2010.

A partir de 2011, a série passa a publicar, também, livros nas áreas de interesse da SBMAC. Os autores que submeterem textos à série Notas em Matemática Aplicada devem estar cientes de que poderão ser convidados a ministrarem minicursos nos eventos patrocinados pela SBMAC, em especial nos CNMAC, sobre assunto a que se refere o texto.

O livro deve ser preparado em **Latex (compatível com o Miktex versão 2.7)**, as figuras em **eps** e deve ter entre **80 e 150 páginas**. O texto deve ser redigido de forma clara, acompanhado de uma excelente revisão bibliográfica e de **exercícios de verificação de aprendizagem** ao final de cada capítulo.

Veja todos os títulos publicados nesta série na página
<http://www.sbmac.org.br/notas.php>

TÉCNICAS DE MODELAGEM DE PROCESSOS EPIDÊMICOS E EVOLUCIONÁRIOS

Domingos Alves
quiron@unisantos.com.br

Laboratório de Computação Científica Aplicada a Saúde Coletiva - LCCASC
Programa de Mestrado em Saúde Coletiva
Universidade Católica de Santos - UNISANTOS
LCCASC - UNISANTOS
Departamento de Informática em Saúde - DIS
Universidade Paulista de Medicina - UNIFESP

Henrique Fabricio Gagliardi
henrique.gagliardi@gmail.com

Laboratório de Computação Científica Aplicada a Saúde Coletiva - LCCASC
Universidade Católica de Santos - UNISANTOS
LCCASC - UNISANTOS
Programa de Mestrado em Informática
Universidade Católica de Santos - UNISANTOS



Sociedade Brasileira de Matemática Aplicada e Computacional

São Carlos - SP, Brasil
2012

Coordenação Editorial: Sandra Augusta Santos

Coordenação Editorial da Série: Eliana Xavier Linhares de Andrade

Editora: SBMAC

Capa: Matheus Botossi Trindade

Patrocínio: SBMAC

Copyright ©2012 by Domingos Alves e Henrique Fabricio Gagliardi. Direitos reservados, 2012 pela SBMAC. A publicação nesta série não impede o autor de publicar parte ou a totalidade da obra por outra editora, em qualquer meio, desde que faça citação à edição original.

Catálogo elaborado pela Biblioteca do IBILCE/UNESP
Bibliotecária: Maria Luiza Fernandes Jardim Froner

Alves, Domingos

Técnicas de Modelagem de Processos Epidêmicos e Evolucionários

- São Carlos, SP : SBMAC, 2012 92 p ; 20,5cm

- (Notas em Matemática Aplicada; v. 26)

e-ISBN 978-85-86883-90-3

1. Simulação Computacional. 2. Epidemias. 3. Quase-espécies Virais.

I. Alves, Domingos. II. Gagliardi, Henrique Fabricio. II. Título.

III. Série.

CDD - 51

Esta é uma republicação em formato de e-book do livro original do mesmo título publicado em 2006 nesta mesma série pela SBMAC.

Agradecimentos

Queremos registrar aqui nossos agradecimentos a todos os alunos colaboradores do Laboratório de Computação Científica Aplicada a Saúde Coletiva (LCCASC - UNISANTOS) que tem acompanhado as pesquisas nessa área e que de uma maneira direta ou indireta contribuíram para o conteúdo deste trabalho.

Aos Professores José Fernando Fontanari do IFSC-USP e Antonio Caliri da FCFRP - USP pelo permanente apoio e incentivo, sendo que muitos dos resultados apresentados aqui tem a participação deles.

Aos organizadores do XXIX CNMAC e aos membros da Sociedade Brasileira de Matemática Aplicada e Computacional - SBMAC.

À FAPESP pelo auxílio concedido durante a realização deste trabalho (proc. número: 2002/03564-8).

Conteúdo

Prefácio	9
1 Elementos de Modelagem de Sistemas Biológicos Complexos	11
1.1 Modelos compartimentais e equações diferenciais	13
1.2 Modelagem estocástica	17
1.3 Autômatos celulares	20
1.4 Grafos e redes complexas	24
1.5 Algoritmos evolucionários	27
2 Modelagem de Processos Epidêmicos	33
2.1 Modelando a heterogeneidade espacial e das interações	34
2.2 Epidemias na rede usando autômatos celulares	37
2.3 Redes complexas em epidemiologia	41
3 Modelagem da Evolução dos Patógenos	45
3.1 Evolução darwiniana in silico	45
3.2 Paradoxos evolucionários e a evolução de uma doença	49
3.3 A caixa de Pandora e a teoria de quase-espécies virais	53
3.4 O paradigma da modelagem da AIDS	59
3.5 A batalha entre o sistema imune e os patógenos	64
3.6 O conceito de memória imunológica revisitado	67
4 Velhos Problemas e Novas Fronteiras	73
4.1 Modelos de controle farmacológico	74
4.2 Como fazer uma modelagem afinal?	80
Bibliografia	83

Prefácio

Neste curso fazemos uma revisão crítica e apresentamos o desenvolvimentos de várias técnicas matemáticas e computacionais de modelagem de processos epidêmicos e evolucionários, de maneira a tentar fazer uma ponte entre os modelos macroscópicos clássicos de espalhamento e a dinâmica populacional de uma doença e uma abordagem ao nível microscópico, onde sejam considerados tanto a dinâmica de crescimento do agente infeccioso, bem como os aspectos evolucionários associados a cada tipo de patógeno sob a influência das pressões seletivas do sistema imunológico do hospedeiro. Para isso, estaremos utilizando técnicas analíticas padrões dessas áreas, bem como técnicas de simulação de uma maneira geral, para formular e testar problemas matematicamente complexos. De qualquer maneira, para fazer isso teremos que ser extremamente restritivos quanto as técnicas e os problemas possíveis de modelagem já que estaremos preocupados principalmente em oferecer uma visão geral e prática do problema da modelagem desses processos. É importante destacar que a motivação principal do curso é estimular o aluno de graduação ou pós-graduação a pensar e desenvolver modelos mais realísticos nas áreas de interface entre, epidemiologia e evolução molecular, voltado a tópicos de relevância experimental de problemas em (micro)biologia, epidemiologia e saúde pública. Nesse sentido, antes de oferecermos aqui uma visão fechada sobre um fenômeno específico, propositadamente estaremos oferecendo algumas visões (aparentemente dispersas) de como modelar, procurando levantar o raciocínio por de trás de cada técnica, e tentando estabelecer as suas limitações. Finalmente, destacamos ainda, o caráter inovador de tal curso onde enfatizamos a criação de laboratórios virtuais para validação de alguns modelos trabalhados no curso de maneira a colocar certos desafios de ordem computacional para a evolução do raciocínio por detrás da modelagem dos processos tratados aqui.

Santos, 20 de maio de 2006.

Domingos Alves
Henrique Fabricio Gagliardi

Capítulo 1

Elementos de Modelagem de Sistemas Biológicos Complexos

A modelagem surge como resultado da investigação científica em qualquer campo e é notoriamente importante no mundo biológico. Um fato que é evidente à primeira vista é que o estudo formal da maioria dos sistemas biológicos não se pode fazer diretamente. Para ter acesso a um estudo teórico de um sistema real necessita-se de um processo gradual de abstração, quer dizer, *é preciso abstrair do sistema natural as características mais importantes*, à luz do objetivo que se deseja alcançar, e ir efetuando sua formulação conceitual para poder realizar o estudo do meio estritamente físico.

Em um primeiro grau de aproximação, as vias por onde se concretiza a modelagem biológica, poderiam ser divididas em dois grupos: a via experimental e a via da simulação. A primeira representa a construção e prova do que se conhece como sistemas modelos experimentais. Essa via tem sido amplamente utilizada, principalmente em biologia molecular. A segunda via contempla os processos catalogados como simulação e é precisamente nesses processos em que vamos centrar nossa atenção. Assim, dentre o mosaico metodológico dentro da biologia a modelagem matemática e computacional (ou simulação de sistemas biológicos) aparece como uma alternativa metodológica que consiste primariamente em entender e manipular para propósitos preditivos os principais *mecanismos* que regem tais fenômenos. É importante frisar, todavia, que o processo de modelagem não é um processo linear, mas recursivo e interativo na maioria dos casos. De uma maneira geral, por um lado temos a realidade física, ou sistema real, que proporciona uma informação que poderíamos chamar de experimental: alguns dados, um comportamento observável, etc. Por outro lado, temos o processo gradual de abstração pelo qual mediante a elaboração de hipóteses, pressupostos, aproximações ou inclusive mediante o emprego de teorias, chegamos à construção de um modelo, reflexo dessa realidade física. Esse modelo posteriormente é desenvolvido, analisado e estudado para obter um comportamento ou uns dados chamados teóricos. Este comportamento é con-

frontado com a realidade experimental, fechando-se o círculo. Desta comparação obtemos informações sobre a qualidade do modelo adotado e da exatidão das hipóteses realizadas.

É importante observar que, a importância atual da modelagem e da simulação não poderia se entender sem a existência de meios e métodos computacionais, cada vez mais potentes, que fazem possível atacar problemas que por sua complexidade até agora eram impossíveis de abordar. Não se trata somente da potência de cálculo mas também da capacidade de representação formal de processos e sistemas de grande complexidade, que não podem ser descritos *exatamente* pela matemática tradicional.

Em particular, a integração de métodos teóricos e computacionais tem desempenhado um papel fundamental no desenvolvimento de modelos epidemiológicos e da dinâmica de populações dos agentes de doenças infecciosas, e, com efeito, existe uma vasta e sofisticada literatura relacionada com a teoria de processos epidêmicos em uma população e a dinâmica dos fenômenos epidêmicos e endêmicos de uma doença. Modelos para estudar epidemias são importantes, pois através deles podemos testar o impacto causado pelos parâmetros inerentes àquele tipo de epidemia. Entretanto, não existe uma solução única para modelar a dinâmica de uma epidemia. A solução ideal para cada caso deve ser buscada tentando responder o porque, onde e quando do fenômeno epidêmico, através da integração das escalas temporal e espacial articulada com um modelo matemático-computacional definido para descrever a epidemia e prever sua evolução. Em adição, atualmente é reconhecido que o fenômeno saúde-doença em comunidades humanas somente pode ser entendido considerando a complexidade e a dinâmica das inter-relações entre muitos fatores operando simultaneamente em múltiplas escalas espaço-temporais e organizacionais. De fato, um indivíduo saudável ou doente sofre os efeitos ininterruptos da evolução micro-biológica dos patógenos, do ambiente antropogênico, de ecossistemas estressados e muitos outros crimes resultantes de desigualdades sócio-econômicas.

Ora, o espalhamento de infecções em uma população envolve o efeito recíproco de duas classes importantes de processos dinâmicos. Um deles é o mecanismo de transmissão, o qual pode se dar através de um processo social, como a AIDS, ou via agentes secundários, como no caso da dengue. O outro componente é a evolução do patógeno invasor dentro do indivíduo, modulada pela pressão seletiva do sistema imune do hospedeiro. A modelagem clássica (*macroscópica*) em epidemiologia tem se preocupado basicamente com a dinâmica populacional e processos demográficos envolvidos no espalhamento de uma doença [1, 2]. Por outro lado, existem muitos estudos recentes da dinâmica do crescimento de um patógeno dentro do indivíduo (*microscópica*) [3, 4, 5], com uma ênfase especial no HIV, vírus causador da AIDS. Particularmente, durante uma epidemia, mutações acumulam-se nos genomas dos patógenos e particularmente em viroses, devido à baixa fidelidade durante a replicação, sendo que algumas dessas mutações confere diferenças fenotípicas, de maneira a permitir que os vírus possam infectar tipos diferentes de células, escapar da resposta imune do hospedeiro, ou serem transmitidos por vias diferentes. A informação epidemiológica, portanto, pode ser obtida através de seqüências geneticamente correlacionadas isoladas de indivíduos infectados.

De uma maneira geral, os modelos podem ser subdivididos segundo vários eixos classificatórios e aqui estaremos tomando como referência de classificação principal a distinção entre modelos determinísticos e estocásticos. Basicamente os primeiros são aqueles que utilizam equações de diferença, diferencial ou integro-diferencial, para descrever a evolução temporal e/ou espacial do fenômeno epidêmico e evolucionário. Ou seja, dadas as condições iniciais as soluções são únicas. Os modelos estocásticos lidam com probabilidades de transição de estados e as conclusões são obtidas pela média sobre varias realizações. No caso das simulações que teremos que realizar sobre esses modelos estará pressuposto então a utilização de geradores de números aleatórios.

No que segue, neste capítulo estaremos discutindo algumas formas de modelar simples dos fenômenos epidêmico e evolucionário, sem a preocupação ainda de entender os processos biológicos por detrás deles. A idéia aqui é distrair-se com os métodos de modelagem, procurando entender as variações e relações entre os métodos. Como efeito colateral esperamos causar um primeiro entendimento de cada fenômeno.

1.1 Modelos compartimentais e equações diferenciais

A modelagem matemática de doenças infecciosas é antiga e remonta os trabalhos de Daniel Bernoulli na análise da varíola em 1760 [1]. De lá aos nossos tempos seu ferramental foi utilizado para explicitar conceitos como imunidade de rebanho utilizado em políticas de vacinação. Levou a *insights* sobre resistência a drogas, taxa de espalhamento de infecções, efeitos de diferentes tipos de tratamentos, etc. Para começarmos a tratar desse problema vamos discutir nesta seção um exemplo simples da formulação compartimental dinâmica. Nesse tipo de abordagem a estratégia é classificar a população em categorias (ou compartimentos) disjuntas de densidades de indivíduos [2].

Dentro dessa abordagem determinística o modelo SIR é um dos mais comuns na literatura, já que com ele é possível introduzir os principais conceitos do processo epidêmico de maneira simples, servindo de base, como veremos, para a formulação de desenhos e procedimentos nos estudos mais gerais em modelagem de epidemias. Segundo esse modelo toda a população pode ser dividida da seguinte forma: (S) a categoria dos susceptíveis, ou seja, dos indivíduos (ou agentes) que não estão infectados, mas podem ser infectados; (I) a categoria dos infectados, ou seja, daqueles indivíduos que são capazes de transmitir a doença aos susceptíveis; (R) a categoria dos recuperados (ou removidos/imunes), ou seja, daqueles que já foram infectados pela doença e morreram (ou se isolaram) ou ficaram permanentemente imunes à doença. Assim, em um instante de tempo t uma população é caracterizada pelas densidades $S(t)$, $I(t)$ e $R(t)$ onde

$$N(t) = S(t) + I(t) + R(t) \tag{1.1.1}$$

é o número total de indivíduos da população a cada instante (tomado como con-

stante no que segue). Usando essa lógica compartimental, se quisermos um maior realismo da descrição do processo epidêmico devemos aumentar o número de compartimentos (ou estados) a que os indivíduos dessa população estão classificados. É importante chamar a atenção aqui que essa classificação não significa em uma divisão da população. Nessa classe de modelos não há espaço e qualquer indivíduo pode estar em qualquer lugar a cada instante, somente os atributos são diferentes para cada indivíduo. Isto quer dizer que este tipo de modelo assume que os indivíduos de qualquer classe estão *homogeneamente misturados* (aleatoriamente misturados), de maneira que qualquer indivíduo de uma classe pode estar em contato com qualquer outro indivíduo de outra classe em qualquer instante.

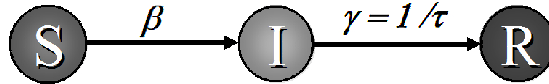


Figura 1.1: O modelo SIR

Com isso em mente podemos entender os mecanismos do modelo através da transição entre esses estados que por sua vez define a dinâmica do modelo (representada esquematicamente na Figura 1.1), fazendo-se três hipóteses básicas: (1) O número de agentes na classe de infectados aumenta a uma taxa que é proporcional ao número de agentes na classe infectante e o número de agentes na classe susceptível. Isso pode ser modelado por $\beta S(t)I(t)$, onde β é uma constante chamada de taxa de infecção. Obviamente essa situação representa a taxa na qual os susceptíveis diminuem em seu compartimento. Esta hipótese é herdada do estudo de cinética química ao considerar que uma reação química se dá quando dois componentes se chocam de maneira efetiva, conhecida como *lei de ação das massas*. (2) A taxa na qual os indivíduos infectados são transferidos para a classe dos removidos (ou recuperados) é proporcional ao número de infectados. Esse fato pode ser modelado por $\gamma I(t)$, onde $\gamma = 1/\tau$ também é uma constante, é a taxa de recuperação, modelada como o inverso do tempo em que um indivíduo infectado permanece nesse estado até se recuperar.

(3) Consideramos nesse modelo um período muito curto de incubação, (desprezível) o que faz com que os indivíduos susceptíveis passem diretamente para o estado infectado.

Esses mecanismos são descritos formalmente pelas seguintes equações diferenciais:

$$\frac{dS(t)}{dt} = -\beta S(t)I(t) \quad (1.1.2)$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t) \quad (1.1.3)$$

$$\frac{dR(t)}{dt} = \gamma I(t) \quad (1.1.4)$$

Combinando as duas primeiras equações:

$$I(t) = I(0) + (S(0) - S(t)) + \sigma \ln \frac{S(t)}{S(0)} \quad (1.1.5)$$

onde fizemos $\sigma = \gamma/\beta$ por simplicidade. Na Figura 1.2, mostramos a solução numérica dessas equações para uma população hipotética de tamanho $N = 10^6$. É interessante notar que a solução para $\gamma = 0$ significa não considerar no modelo o compartimento dos recuperados R , ou seja, estudar um modelo mais simples com dois compartimentos SI. Nesse caso, todos os indivíduos passam para o estado infectado. À medida que vamos diminuindo o tempo em que o indivíduo permanece infectado, diminuímos o dano na população.

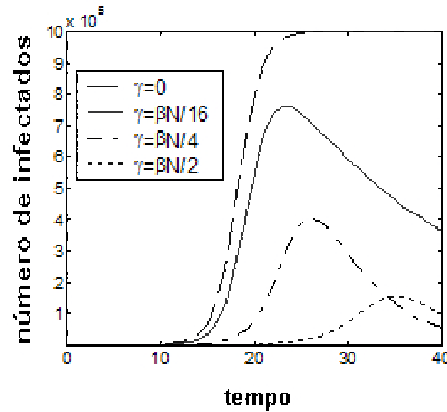


Figura 1.2: Curvas epidêmicas obtidas ao se variar os parâmetros do modelo SIR definido pelo conjunto de equações diferenciais acima.

Vamos aproveitar este modelo ainda para estabelecer quais as condições para uma epidemia se estabelecer. De maneira simples, manipulando as equações 1.1.2 e 1.1.3, podemos escrever um critério para que o número de indivíduos infectados não cresça à medida que o tempo passa:

$$\frac{dI(t)}{dt} = \left(-1 + \sigma \frac{1}{S(t)} \right) \frac{dS(t)}{dt} < 0 \quad (1.1.6)$$

Essa relação é satisfeita quando $S(t) < \rho$, em particular, a epidemia não se estabelece quando $S(0) < \rho$. Esse critério pode ser melhor entendido introduzindo o número de reprodução básica R_0 [2], um parâmetro adimensional que fornece o número de casos secundários produzidos por um infectante típico durante o período infeccioso. Em doenças de transmissão direta, esta grandeza é igual ao número médio de contatos adequados de um indivíduo infectante durante o período infeccioso.

Assim, para o modelo que estamos tratando aqui, se fizermos:

$$R_0 = \beta\tau = \frac{\beta}{\gamma} = \frac{1}{\sigma} \quad (1.1.7)$$

Essa é uma medida complexa e ambígua, muito cara na modelagem de epidemias principalmente quando se parte do pressuposto compartimental, já que resume várias propriedades importantes do processo infeccioso, mas de qualquer maneira depende do modelo adotado (ver [6] para uma discussão detalhada da evolução desse conceito). Ainda, denotando-se um valor crítico de susceptíveis S_T satisfazendo esse critério e S como o número de susceptíveis da população, podemos escrever:

$$R_0 = \frac{S}{S_T} \implies R_0 > 1 \quad (1.1.8)$$

ou seja, R_0 é dado nesse modelo pela razão entre o número de susceptíveis na população e a densidade limiar necessária para a persistência da doença. Em termos dos parâmetros do modelo a equação 1.1.7 fornece um critério como condição necessária para a persistência da epidemia. Esse resultado pode ser derivado de uma maneira mais geral e é conhecido como teorema do limiar epidêmico [7]:

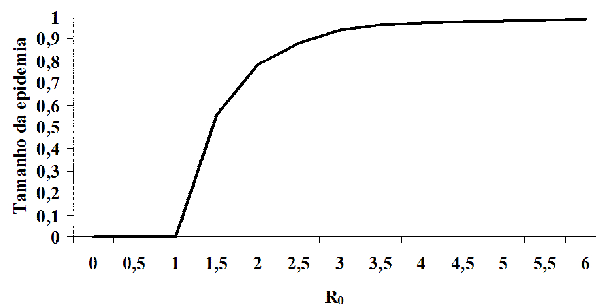


Figura 1.3: Fração da população infectada durante uma epidemia em função do número de reprodução básica R_0 .

O gráfico da Figura 1.3 ilustra as propriedades dessa grandeza, onde podemos observar o efeito de se aumentar R_0 . É importante destacar ali o *limiar de transmissibilidade* definido pelo valor $R_0 = 1$, uma vez que a persistência da infecção é possível quando $R_0 > 1$.

Como já comentado anteriormente, a partir desse modelo, outros modelos podem ser desenvolvidos com o intuito de estudar os mais variados tipos de doenças como o sarampo, dengue ou doenças venéreas, sendo necessário em alguns a inserção de outros compartimentos, como é o caso do modelo SEIR compartimental, baseado no modelo SIR compartimental, por exemplo, onde o período de incubação passa a ser considerado gerando a classe E dos indivíduos expostos, caracterizada por armazenar a densidade de indivíduos que são infectados, porém ainda não infecciosos no sistema. No entanto, apesar das importantes contribuições desses modelos

compartimentais para o desenvolvimento de teorias epidemiológicas, eles apresentam limitações, pois, por mais que neles sejam agregadas informações, sempre serão regidos por alguns pressupostos básicos que limitam sua abrangência [8].

Com efeito, os modelos baseados em equações diferenciais, assumem que a população é totalmente misturada e, portanto desprezam fatores espaciais no espalhamento da epidemia. Tal abordagem faz mais sentido se considerarmos uma população pequena e fechada, ao passo que o mesmo se torna bastante irreal se adotarmos uma população grande e com pouca mobilidade, visto que nesse último caso deveria haver maior probabilidade da transmissão ocorrer entre os vizinhos de cada indivíduo. Além disso, esse tipo de modelagem não leva em consideração o fato de que a população é composta de indivíduos interagentes, ao considerar as populações como entidades contínuas, o que produz resultados não realísticos para pequenas populações. Em adição, nos modelos compartimentais todos os indivíduos dentro de um compartimento são tidos como semelhantes entre si. Sabe-se, entretanto, que há muitos fatores biológicos (idade, predisposição genética, dentre outros) e sociais (hábitos, costumes e círculo de amizades) que acabam influenciando para que um indivíduo adoça. Tais fatores podem ser, de certa forma, ainda inseridos nesse tipo de modelagem a partir da criação de sub-compartimentos menores que serviriam para classificá-los, contudo isso se demonstra inviável quando os fatores são muitos.

É interessante salientar que, tais pressupostos acabam limitando vários fatores importantes para o estudo de moléstias e que poderiam conduzir a uma nova visão sobre seu processo de espalhamento por considerar tudo de forma homogênea no sistema (interações e indivíduos), desconsiderando a individualidade de cada elemento. Portanto, de uma maneira geral para uma melhor previsão e para definir melhores estratégias no combate ao espalhamento de uma epidemia é necessário desenvolver novas metodologias matemáticas e computacionais.

1.2 Modelagem estocástica

Dentre as limitações inerentes a formulação determinística, por sua própria natureza, este tipo de modelagem não leva em consideração efeitos de flutuações estocásticas de uma maneira geral. Esse tipo de complexidade tem sido incorporada por modelos epidemiológicos estocásticos (ver [1] para uma revisão aprofundada) que por sua vez são baseados no mesmo raciocínio compartimental, e assim construídos com variáveis probabilísticas e suas relações de incertezas entre suscetíveis, Infectados e Recuperados.

Para entender a transição entre essas duas abordagens podemos interpretar os modelos determinísticos, como o discutido na seção anterior, como aproximações válidas para grandes populações conhecidas da física como aproximação de campo médio, ou seja, as soluções desses modelos como sendo a média dos mecanismos epidêmicos que evoluíssem estocasticamente. Ora, quando tomamos a formulação determinística e dizemos que um evento (uma transição de estado, por exemplo) ocorre a uma taxa λ tipicamente devemos interpretar que o tempo τ entre sucessivas ocorrências tem uma distribuição exponencial de probabilidade com parâmetro λ ,

isto é, $P(\tau \leq t) = 1 - \exp(-\lambda t)$. Com isso em mente, vale destacar que de qualquer maneira a aproximação determinística das flutuações estocásticas é reconhecida-mente inapropriada para pequenas populações [1].

Vamos rever brevemente alguns aspectos da modelagem estocástica do modelo SIR estabelecido na seção anterior. Para isso vamos manter a hipótese de que N indivíduos na população estão homogeneamente misturados e por simplicidade supor que a epidemia começa no tempo $t = 0$. Em adição, devemos interpretar os estados $S(t)$, $I(t)$ e $R(t)$ como variáveis aleatórias tomando o cuidado de fazer $s(t) = S(t)/N$, $i(t) = I(t)/N$ e $r(t) = R(t)/N$ de maneira que essas variáveis podem ser aproximadas por números reais no limite em que $N \rightarrow \infty$. Assim, podemos escrever que a probabilidade de que ocorra uma nova infecção em um pequeno intervalo de tempo Δt é $\beta SI\Delta t + o(\Delta t)$. Quando essa transição ocorre, S decresce por uma unidade e I aumenta de uma unidade. Todavia, nesse intervalo de tempo ainda temos a possibilidade de que um indivíduo na população se recupere da infecção. A probabilidade de que isso aconteça é $\gamma I\Delta t + o(\Delta t)$ e nesse caso I decresce de uma unidade enquanto S permanece inalterado. Podemos escrever agora a probabilidade $p_{si}(t)$ para que no instante t existam s susceptíveis não infectados e i infectados em circulação. Para fazer isso, tomamos como hipótese que o intervalo de tempo para a infecção e uma eventual recuperação da infecção tem uma distribuição exponencial negativa. Usando argumentos sobre a relação entre probabilidade de estados adjacentes podemos escrever o seguinte conjunto de equações diferencial-diferença para as transições:

$$\frac{dp_{si}}{d\tau} = w_c(S + 1, I - 1)p_{s+1, i-1} - [w_c(S, I) + w_r(I)]p_{si} + w_r(I + 1)p_{s, i-1} \quad (1.2.9)$$

onde $w_c(S, I) = \beta SI/N$ é a probabilidade de transição de contágio por unidade de tempo e $w_r(I) = \gamma I$ é a probabilidade de transição de recuperação por unidade de tempo com a condição inicial de que: $p_{si}(0) = 1$. É conveniente aqui mudar a escala de tempo que vamos trabalhar para $\tau = \beta t$ de maneira que a probabilidade de que $S(\tau)$ diminuir de uma unidade em $\Delta\tau$ é $S(N - S)\Delta\tau$ aproximadamente, de maneira que essa expressão está intimamente relacionada ao número exato de novas infecções no intervalo de tempo $\Delta\tau$ na versão determinística do modelo SIR.

Embora a expressão 1.2.9 seja aparentemente simples em sua forma, a sua solução explícita é bastante complicada de se manipular. Nessa equação, conhecida como equação mestra estão contidas todas as informações do sistema modelado e apesar de se dispor de vários métodos analíticos para resolvê-la (ver [1] para os detalhes da solução heróica desse sistema) nem sempre é possível obter uma solução exata ou mesmo aproximada para esse sistema. Em qualquer caso, sempre se pode tentar uma simulação do processo estocástico, ainda mais se temos em mãos o método adequado para tal. Para entendermos o sentido da palavra adequado, colocado aqui, precisamos reconhecer que o modelo SIR estocástico representado pela equação 1.2.9 é um processo estocástico markoviano do tipo nascimento e morte (particularmente de estado discreto e tempo contínuo), sendo que por construção ele herdou o comportamento espacialmente homogêneo do modelo SIR determinístico. Esses requisitos

são importantes porque podemos aplicar um método de simulação estocástica exato comumente utilizado para estudar reações químicas acopladas e que foi desenvolvido por D. T. Gillespie [9, 10]. Gillespie demonstrou que a probabilidade de que em um tempo τ não ocorra nenhuma transição de estado, ou seja, o sistema permaneça no estado i é dada por:

$$P(\tau) = e^{-[w_c+w_r]\tau} = e^{-[\beta si+\gamma i]\tau} = e^{-[si+\sigma i]\tau} \quad (1.2.10)$$

com esse expressão podemos observar que quando t tende a zero a probabilidade de o sistema ficar no mesmo estado tende a unidade. Se isolarmos τ nessa expressão obtemos ainda:

$$\tau = \frac{1}{si + \sigma i} \ln \frac{1}{P(\tau)} = \frac{1}{si + \sigma i} \ln r_1 \quad (1.2.11)$$

e assim se associarmos um número aleatório r_1 , compreendido entre 0 e 1, à probabilidade $P(\tau)$, podemos calcular τ . Isto quer dizer que podemos gerar um número aleatório utilizando um gerador de números aleatórios, para calcularmos o tempo até ocorrer uma transição (ou um susceptível tornando-se infectado ou um infectado tornando-se recuperado). Temos que advertir que como efeito colateral desse tipo de simulação, esse tempo τ calculado dessa maneira tem as mesmas dimensões das taxas de transição.

A segunda parte do método consiste em averiguar qual evento vai ocorrer dentre os dois possíveis depois de transcorrido o tempo τ . Para fazer isso geramos um segundo número aleatório r_2 , também compreendido entre 0 e 1. O esquema segue averiguando-se, se esse número aleatório estiver no intervalo:

$$0 \leq r_2 \leq \frac{si}{si + \sigma i}$$

deve ocorrer a transição $i \rightarrow i + 1$. Ao contrário se estiver no intervalo:

$$\frac{si}{si + \sigma i} \leq r_2 \leq 1$$

deve ocorrer a transição $i \rightarrow i - 1$. Obviamente será mais provável que suceda aquele processo cuja probabilidade de transição por unidade de tempo seja maior. A partir desse passo de simulação temos de repetir o procedimento tantas vezes seja necessário, obtendo-se um caminho aleatório (uma realização de nosso processo estocástico) seguido pela variável aleatória que estamos considerando, no caso, o número de infectados em um tempo t . Como já comentado para extrair resultados sobre valores médios e flutuações será necessário repetir a simulação muitas vezes. Quanto maior for o número de simulações, tanto melhor é o ajuste da média e variância, etc, com os valores dedutíveis analiticamente (quando isso for possível) da equação mestra.

É importante neste ponto entendermos um pouco melhor o que estamos fazendo com esse método de simulação. Cada realização de nosso processo estocástico, ou seja, cada rodada de nossa simulação durante um tempo t em que levantamos a curva epidêmica para um conjunto de parâmetros particular significa que estamos

fazendo um experimento computacional sobre uma população hipotética em que controlamos as possíveis causas da epidemia. Ao tomarmos as médias, vemos o efeito de várias epidemias similares em suas causas e efeitos em várias populações. As causas e efeitos são modeladas aqui pelas taxas de transmissão e de recuperação características, por exemplo, de uma doença (como o sarampo), em uma população fechada de tamanho N sem considerar fatores demográficos como nascimentos e mortes.

De uma maneira geral esses modelos *clássicos* de epidemia como os determinísticos e os estocásticos falham em muitos aspectos para modelar adequadamente. Particularmente, no caso estocástico geramos um resultado em geral diferente do determinístico sendo que a curva epidêmica ou se preferirmos o tamanho da epidemia é em geral menor do que no análogo determinístico e os intervalos de tempo em que a epidemia ocorre também são diferentes nos dois casos. Como veremos em uma versão estocástica mais abrangente do modelo SIR, em alguns experimentos simplesmente a epidemia não se estabelece e ficamos com esses resultados em média menor.

1.3 Autômatos celulares

Vamos agora tentar levar em consideração no modelo SIR a heterogeneidade espacial das relações populacionais, ou seja, vamos explicitar a dinâmica espacial do processo infeccioso, já que como vimos no SIR baseado na premissa compartimental não existe espaço. Para fazer isso de uma maneira simples, temos de alterar o pressuposto compartimental e formular *modelos baseados no indivíduo*, onde a unidade básica a ser modelada é o indivíduo e não mais uma população. Dessa maneira, agora devemos trabalhar com cada indivíduo explicitamente considerando sua individualidade mesmo que, em uma primeira análise, essa individualidade seja atribuída apenas pela distinção entre localizações espaciais, e que a transmissão entre eles ocorra de forma local e não mais global (como o pressuposto compartimental).

A base de quase todo modelo baseado no indivíduo é a modelagem via autômatos celulares, introduzidos originalmente por John von Neumann e Stanislaw Ulam em 1940 [11, 12], e que também são uma ferramenta de grande auxílio para o estudo de sistemas que possuem fenômenos coletivos, como quebra de simetria, caos, ordenamento de fases e turbulência, tendo muitas aplicações na simulação de sistemas biológicos (processos de reprodução), fenômenos físicos (difusão) e sistemas sociais (formação de comunidades), dentre outras, por apresentarem a formação de padrões espaciais e temporais.

Os autômatos celulares podem ser definidos como sistemas dinâmicos discretos cujo comportamento é especificado em termos de relações (ou regras) locais. Assim, em um autômato celular é necessário que o espaço seja criado para que este possa conter os indivíduos que formarão a população que se pretende estudar e, para isso, usualmente utilizamos uma rede regular, na qual cada sítio ou célula assume um conjunto discreto de estados.

Para se construir um autômato celular é necessário estabelecer quais são as regras de transição que deverão ser seguidas, pois, uma pequena variação nestas regras pode alterar a solução do problema estudado. Tais regras dependem de uma série de escolhas que devemos fazer, que são: (1) *Dimensões da rede*: unidimensional, bidimensional, e tridimensional. (2) *A geometria da rede*: deve-se especificar qual será a forma que a rede deve ter: quadrada, hexagonal ou triangular, dentre outras. (3) *A vizinhança*: para qualquer célula de uma rede, há uma vizinhança que é formada por um conjunto de células que podem ou não afetar seu estado e que pode ser definida a partir da dimensão do autômato e a geometria de suas ligações. Os dois critérios mais freqüentemente utilizados para uma rede bidimensional são a vizinhança de von Neumann, em que são consideradas como vizinhas às células que compartilhem as arestas de determinado indivíduo (no caso da rede quadrada são somados quatro vizinhos); e a vizinhança de Moore, em que são acrescentados os vizinhos que têm vértices em comum com a célula que esta sendo considerada (elevando o número de vizinhos para oito no caso da rede quadrada) [13]. (4) *As condições de contorno*: como o tamanho da rede deve ser finito, uma célula quando se encontra na borda do autômato pode ter problemas em *saber* quem são seus vizinhos que por sua vez podem representar uma porcentagem muito grande da população quando esta for reduzida. Por exemplo, considerando uma rede de 10×10 células, 40% destes se encontram nas bordas. Assim, surge a necessidade de se estabelecer condições de contorno neste autômato, que podem ser periódicas, reflexivas ou fixas. As condições periódicas de contorno são muito utilizadas por proporcionarem uma situação equivalente a de simulações com redes infinitas (em uma rede unidimensional esse procedimento gera uma fita circular e em uma rede bidimensional o resultado é uma forma toroidal).

Como comentado anteriormente, os estados que uma determinada célula pode assumir é descrito por um número discreto e finito de estados, cada qual sendo definido por um número ou propriedade, conforme necessário. As regras de transição no autômato são as regras que fazem com que um estado possa mudar em determinada célula que representa um indivíduo da população. Essas transições dependem do estado atual desta célula e também dos estados em que seus vizinhos se encontram. Essas regras podem ser: determinísticas, quando há a dependência quanto à configuração do sistema (a partir da vizinhança); totalísticas, quando estas regras não dependem da configuração destas células vizinhas (dependem somente do número de vizinhos em um estado específico) e regras probabilísticas, na qual não existe uma configuração específica para cada configuração da vizinhança, mas possibilita vários resultados associados a uma probabilidade. No que segue, vamos pensar no modelo SIR, utilizando um autômato celular probabilístico considerando uma rede bidimensional, quadrada, uma vizinhança de Moore e condições periódicas de contorno.

Para isso, vamos utilizar os três estados do modelo SIR e o mesmo esquema de transição feito na abordagem compartimental, de maneira que esses indivíduos, são distribuídos em células de uma rede $\mathbf{M} = \{m_{ij}\}$, com i e j variando de 1 a L (de maneira que $N = L \times L$). Assim, cada indivíduo de uma célula recebe três atributos específicos: (1) a posição geográfica ou posição na rede (i, j) ; (2) um dos

possíveis estados clínicos, estados específicos da doença para cada indivíduo, s , i e r ; e finalmente (3) um período médio de tempo τ no qual o indivíduo infectado permanece nesse estado, podendo transmitir a doença para os outros indivíduos na população. Observe que mantemos a hipótese de uma população fechada fazendo-se $\sum s + \sum i + \sum r = N$, com N constante. A dinâmica de transmissão da infecção nessa população é descrita, passo a passo, através de regras de interação entre os indivíduos, e assume que cada nova configuração do sistema depende apenas do estado prévio de cada indivíduo que mudam segundo as seguintes regras: (1) Qualquer indivíduo pode ser infectado com uma probabilidade local p_L . (2) Os infectados tornam-se recuperados deterministicamente (tornam-se imunes) depois do período infeccioso τ , que como uma simplificação é considerado constante para todos os indivíduos infectados. (3) Uma vez recuperados, os indivíduos participam apenas passivamente do espalhamento da infecção por um período de imunidade maior que o processo completo da epidemia. Durante um passo de tempo desse algoritmo, essas regras são aplicadas sincronicamente a todos os sítios da rede. A probabilidade $p_L = p_L(i, j)$, entendida como uma probabilidade por unidade de tempo, pode ser interpretada como uma probabilidade individual (contato da vizinhança próxima) a partir da qual a epidemia vai se espalhar pela rede). Assim, se imaginarmos um indivíduo infectado na rede, ele estará em contato com n indivíduos susceptíveis, seus primeiros vizinhos ($0 \leq n \leq 8$ no caso da vizinhança de Moore que estamos considerando). Modelamos que esse indivíduo infectado pode infectar um desses vizinhos susceptível com probabilidade de contato λ ($\lambda \in [0, 1]$). Daí, a probabilidade de ele não infectar esse vizinho é $(1 - \lambda)$ e a probabilidade de ele não infectar nenhum dos n vizinhos susceptíveis é $(1 - \lambda)^n$. Portanto, a probabilidade p_L de a infecção ser transmitida aos n primeiros vizinhos de um indivíduo infectado na rede pode ser escrita como:

$$p_L = 1 - (1 - \lambda)^n \quad (1.3.12)$$

Esta simples aproximação permite o estudo detalhado da dinâmica do comportamento do modelo numa rede completa a partir do controle do parâmetro λ . Assim, para obter uma simulação do processo epidêmico nesse modelo baseado no indivíduo, basta repetir o seguinte algoritmo: (1) para cada indivíduo da rede verificamos seu estado: se ele for susceptível verificamos se existem infectados na vizinhança para calcular a probabilidade de ele tornar-se infectado ou permanecer susceptível; se ele for infectado observamos quantos passos de tempo ele esta nesse estado, se for maior que τ ele se recupera da infecção deterministicamente (podemos alternativamente tomar τ de uma distribuição de probabilidades como argumentado na seção anterior). Se o estado desse indivíduo for recuperado ele permanece nesse estado indefinidamente. (2) Anotadas todas as possíveis transições ao percorrer a rede fazemos uma atualização sincrônica de todos os estados (voltamos ao primeiro passo). A Figura 1.4 mostra uma simulação utilizando-se desse procedimento para $\lambda = 0,3$ (ou seja, apenas 30% dos contatos em média entre infectados e susceptíveis são efetivos para a transmissão da infecção), tomada em quatro instantes de tempo distintos. Ali é possível observar a *onda* epidêmica de infectados caminhando por toda a população.

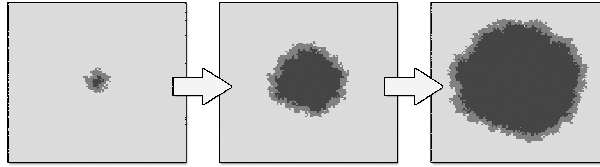


Figura 1.4: Modelo SIR discreto utilizando-se o algoritmo descrito para o autômato celular. Representa uma epidemia que se espalha devido aos contatos locais (físicos) em 4 instantes de tempo distintos: 10, 30 e 50 dias respectivamente.

É interessante notar, que quando o espalhamento da infecção é determinístico no autômato, ou seja $\lambda = 1$, os 8 vizinhos mais próximos de um infectado acabam se infectando e nesse caso teríamos a *velocidade da onda epidêmica* com seu valor máximo. À medida que diminuimos λ essa *velocidade* diminui até um valor crítico em que a epidemia não se espalha em média nos experimentos (ou populações) em que aplicamos com esse algoritmo. Com efeito, é fácil verificar que o número de reprodução básica nesse caso pode ser escrito como, $R_0 = 8\lambda\tau$ (a despeito da real utilidade dessa grandeza) e tipicamente leva a conclusão de que o espalhamento local necessita de um R_0 maior do que o calculado no modelo compartimental para gerar uma epidemia. Um raciocínio mais simples e realístico seria observar que se a probabilidade de transmissão é pequena, é bem possível que o indivíduo infectado se recupere antes de transmitir a infecção para alguém.

De qualquer maneira, o algoritmo descrito (como veremos) pode ser facilmente generalizado para outros modelos epidêmicos e pode ser facilmente modificado para incorporar de forma mais realística a dinâmica da conexão entre indivíduos realizada por suas atividades diárias. O que é necessário reter por enquanto é que em uma rede extremamente regular de contatos (como a montada para esse modelo específico), o número e a *intensidade* de contatos experimentados por cada membro da população é igual e local (interação entre vizinhos na rede). Obviamente, esse tipo de modelo não é biologicamente realístico e de fato o progresso de uma infecção é lento nesse caso. Isto ocorre, por causa dos altos níveis de correlações introduzidas pelas interações locais, de maneira que a maioria dos contatos sejam efetivamente perdidos. Em um outro extremo, podemos pensar em uma população distribuída em uma rede e que qualquer indivíduo pode interagir com qualquer outro, com a mesma probabilidade (como no caso da formulação compartimental onde efetivamente não há espaço). Esse pressuposto de modelo de agregação (do tipo campo médio) [2] é bastante irreal quando uma população grande e de pouca mobilidade está sendo modelada. Aqui, a chance de transmissão entre vizinhos deverá ser maior que entre pessoas distantes.

Observamos, portanto que esses extremos de modelagem não produzem os resultados observados no desenvolvimento de uma epidemia ou por produzirem epidemias muito rápidas ou muito lentas. Mais ainda, o fator chave que falta modelar nessas duas aproximações é a *mobilidade*, ou seja, devemos pensar em introduzir uma heterogeneidade espacial nas interações dessa população introduzindo uma idéia

adicional de que os indivíduos podem ser distinguidos pelo grau de mobilidade na rede.

1.4 Grafos e redes complexas

Muitos dos desafios teóricos enfrentados por epidemiologistas são problemas de escala. Como podemos construir um modelo que, partindo dos processos acontecendo no nível individual seja capaz de gerar as conseqüências daquele processo ao nível populacional? Particularmente, a modelagem da heterogeneidade espacial é um tópico importante em epidemiologia, que envolve esse tipo de problema. Nesse caso, as populações são compostas de indivíduos que interagem somente com um número limitado de outros indivíduos. Em contrapartida, todo membro da população pode entrar em contato com qualquer outro *via* uma rede de contatos.

Existem vários tipos de redes de contato que podem ser utilizadas por epidemiologia para modelar transmissão. Na Figura 1.5 relacionamos alguns tipos de redes que são comumente usadas em epidemiologia para modelar as interações que podem ocorrer em modelos baseados em indivíduos. Nessas redes, um indivíduo interage com muitos outros desta população, que por sua vez estão conectados a uma rede maior. Com efeito, o espalhamento de uma epidemia ocorre a partir das conexões entre eles e um indivíduo estando infectado pode transmitir a doença por meio destas conexões e assim ela vai se alastrando pela população. A Figura 1.5-a), por exemplo, representa uma rede regular de indivíduos, na qual o número de contatos entre eles é sempre igual (que é o caso modelado na seção anterior).

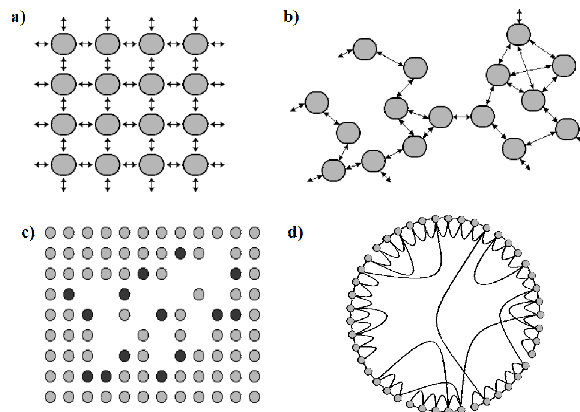


Figura 1.5: Vários tipos de redes de contato que podem ser utilizados em epidemiologia para modelar transmissão.

A Figura 1.5-b) apresenta uma rede cujos contatos são obtidos por meio de arranjos aleatórios das conexões entre indivíduos, o que gera uma heterogeneidade espacial relacionada a transmissão. Ali cada indivíduo pode ser infectado por ou

transmitir a infecção para os indivíduos com os quais ele está conectado. O número de conexões médias determina a dinâmica de transmissão do sistema. Na Figura 1.5-c) temos a representação esquemática de uma rede que vem sendo utilizada como um modelo para a epidemia de sarampo [14]. Trata-se de uma rede quadrada onde há a possibilidade de cada sítio poder estar vazio, conter um indivíduo saudável (círculo cinza) ou infectado (círculo preto). A cada passo de tempo a doença se propaga localmente, onde os elementos saudáveis que fazem parte da vizinhança de um indivíduo infectado são infectados com uma probabilidade estabelecida. Os indivíduos infectados então se tornam imunes, são eliminados da rede e seus sítios ficam vazios.

Um exemplo interessante e mais geral de modelagem espacial de epidemias é o modelo de *mundo pequeno*, recentemente proposto por Watts e Strogatz [15]. Particularmente, esse tipo de modelagem da rede apresentada na Figura 1.5-d) (para o caso particular de uma rede unidimensional) busca mimetizar a heterogeneidade de interações entre indivíduos (interações sociais), deixando estes interagir com os seus vizinhos mais próximos e ocasionalmente com indivíduos distantes entre si através de um protocolo de re-conexão entre estes contatos, caracterizando um cenário em que ocorre um aumento na velocidade da transmissão da epidemia atribuindo-lhes mobilidade, reduzindo assim, a distância média entre qualquer dois indivíduos. Basicamente, esse algoritmo gera um modelo abstrato, onde a população é inicialmente disposta em uma rede regular e unidimensional, utilizando condições de contorno periódicas (em formato circular, no qual o primeiro elemento desta rede relaciona-se com o último e vice-versa) de N indivíduos, cujas relações são representadas por $2K$ arestas definidas (número de coordenação K para os vizinhos adotados em cada sítio). Em seguida é verificado no sentido anti-horário (ou horário) da rede, se cada conexão pode ser modificada com uma probabilidade p e, em caso afirmativo, um elemento qualquer é escolhido aleatoriamente para substituí-la, desde que não haja auto-conexões (um indivíduo se conectar a ele mesmo) ou múltiplas conexões (mais de uma conexão para um mesmo indivíduo). A partir de um simples parâmetro p , é possível então ajustar o grau de desordem da rede, conforme esquematizado na Figura 1.6.

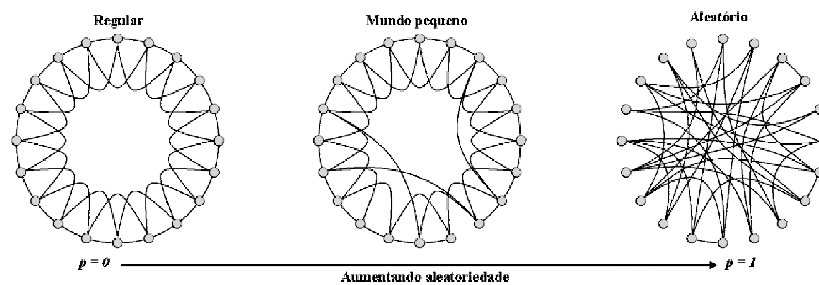


Figura 1.6: Desenho esquemático do modelo de mundo pequeno. Muitos contatos são entre indivíduos adjacentes, mas ocasionalmente contatos de longa distância podem ocorrer com o aumento da probabilidade p .

Assim, cada indivíduo dessa rede passa a ter uma mobilidade adicional, entrando em contato com outros indivíduos da rede (mais distantes que seus vizinhos mais próximos) através de re-conexões feitas a partir do valor ajustado para p . No caso da transmissão de uma infecção, essas re-conexões garantem uma heterogeneidade do sistema com relação às interações (introduzem-se interações não-locais), pois o indivíduo pode alcançar uma região da rede que precisaria de vários passos para ser atingida e transmitir a doença, encurtando assim as distâncias entre os indivíduos, dando mobilidade e estabelecendo um cenário mais susceptível a uma transmissão mais rápida da infecção na população, caracterizando assim o fenômeno de *mundo pequeno* entre eles. Com esse algoritmo, pode-se interpolar entre uma estrutura regular (quando $p = 0$) e uma estrutura aleatória (quando $p = 1$).

Estabelecido o valor do parâmetro p , ou seja o *mundo* no qual a epidemia vai se disseminar, podemos estabelecer os mesmos estados possíveis para os indivíduos dessa população como no modelo SIR, mantendo a partir daí o mesmo esquema de transmissão inter-individual como a probabilidade p_L e o mesmo tempo durante o qual um indivíduo infectado permanece nesse estado τ . Obviamente que o esquema de vizinhança neste caso muda e o número de vizinhos que um determinado indivíduo pode ter depende de qual *mundo* o protocolo de re-conexão gerou. Uma vez gerada a rede a partir desse protocolo fazemos os nossos experimentos com várias populações distribuídas nessa *cidade* abstrata onde a epidemia se espalha. A Figura 1.7 mostra o resultado desses experimentos ao se variar o parâmetro de desordem p , ou como estamos argumentando até aqui para várias redes (ou mundos) no qual uma epidemia pode se espalhar.

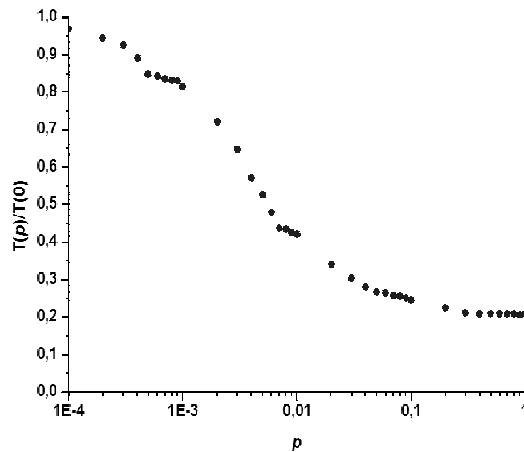


Figura 1.7: Tempo $T(p)$ requerido para que uma epidemia infecte toda a população, em função do parâmetro de desordem p . Nessa figura observamos a curva do tempo para a disseminação de uma epidemia baseada no modelo SIR (dividindo todos os tempos pelo maior tempo obtido, quando $p = 0$) para cada valor distinto p , calculando médias em 20 realizações para uma vizinhança de $k = 4$ indivíduos em redes unidimensionais.

Nessa figura, por simplicidade, tomamos a probabilidade de infecção $\lambda = 1$ e assim podemos nos ater exclusivamente aos efeitos do parâmetro de desordem. Para isso anotamos o tempo de duração da epidemia relativo a esse mesmo tempo quando a rede é regular ($p = 0$).

É interessante notar que, mesmo para valores pequenos de p , esse tempo diminui bruscamente. Isso significa que, ao aumentar o valor de p , menor vai se tornando o número de elementos intermediários necessários para que a doença transmitida por um infectado atinja determinado susceptível, o que caracteriza o fenômeno de mundo pequeno nesse tipo de modelo. Além disso, tal variação possibilita que a rede toda seja atingida pela doença mais rapidamente, dado que o espalhamento vai adquirindo proporções maiores à medida que a ocorrência de atalhos entre indivíduos distantes entre si (gerados pelo protocolo de re-conexão), são mais freqüentes.

1.5 Algoritmos evolucionários

Até aqui vimos como alterar um modelo simples de epidemia de maneira a inserir paulatinamente níveis de complexidade. Em todas as instâncias de modelagem que trabalhamos com o modelo SIR estivemos observando as propriedades macroscópicas da transmissão de uma infecção entre indivíduos de uma população. A questão que nos ocupamos aqui nesta seção é saber o que ocorre dentro de cada indivíduo durante uma infecção e, melhor ainda (caso seja possível), como essa visão microscópica pode influenciar na transmissão da infecção na população. Para fazer isso, vamos restringir nossa análise aos aspectos evolucionários inerentes a invasão de um patógeno em um hospedeiro humano. Assim, se excluirmos acidentes ou criação sobrenatural, resta-nos a tarefa de delinear as leis naturais que governam a dinâmica dessas entidades invasoras, isto é, em sistemas que são capazes de se auto-reproduzir e auto-sustentar, ocasionalmente com variações na descendência. O que procuramos são princípios de uma *evolução biológica* (mais familiar no sentido de Darwin) onde, de uma maneira muito simplificada, indivíduos geneticamente distintos competem por recursos limitados sendo que aqueles que o fazem melhor produzem mais filhos, tornando-se os maiores *acionistas genéticos* da próxima geração. Somado a isso, mutações aleatórias misturam os gens para criar filhos com novas combinações genéticas e em toda geração, a *peneira* da seleção natural passa entre os competidores, eliminando combinações que não funcionam e adaptando esses organismos ao seu meio, o ser humano. Paralelamente, devemos ficar atentos para os mecanismos de como estas leis operam; por exemplo, Darwin assumiu que os genes dos pais misturam-se em seus filhos. Isso era problemático porque as variantes dos genes mais aptos deveriam diluir-se na *sopa de genes* depois do cruzamento. O Mendelismo resolveu esse paradoxo construindo um mecanismo para descendência, onde *alelos* individuais permanecem intactos e não se misturam mesmo depois do cruzamento. Este exemplo mostra que o conhecimento dos mecanismos pelos quais variantes são geradas, teve profundo efeito sobre a teoria evolucionária.

A maioria dos sistemas auto-replicativos reais que envolvem patógenos estão formados por moléculas de DNA que apresentam pouco polimorfismo estrutural de-

vido aos diversos mecanismos de reparo que esses sistemas possuem para minimizar erros de cópia na reprodução [16]. Para ser um bom modelo experimental de algo que tem de ser ao mesmo tempo objeto e sujeito do processo evolutivo, é necessário que o sistema seja encontrado na natureza com uma vasta variabilidade de estruturas, requisito esse encontrado no RNA de uma grande variedade de populações virais. Nesse sentido, os primeiros resultados importantes para evolução molecular foram descritos por Spiegelman [17] utilizando um tipo de vírus de RNA que infecta a bactéria *Escherichia coli* (um bacteriófago), o Q β . O tipo dominante (seqüência mestra) desse fago tem um RNA com mais de 4000 nucleotídeos e foi utilizado como *semente* nos experimentos. Eles utilizam o método tradicional de passagens em série e a pressão de seleção era controlada durante o experimento. Assim, depois de várias rodadas de replicação, foi observado que as moléculas de RNA sobreviventes eram formadas por somente algumas centenas de nucleotídeos. Para explicarmos esses resultados é importante primeiramente descrever a técnica experimental utilizada, que aparece esquematizada na Figura 1.8. Este método de passagens consiste de uma série de tubos de ensaio, todos eles contendo monômeros ricos em energia em igual concentração e a enzima replicase que catalisa a replicação. Ao primeiro tubo coloca-se o RNA viral, e depois de um período determinado no qual deu-se a replicação do RNA, uma alíquota muito pequena do conteúdo do primeiro tubo de ensaio é transferida ao segundo, e assim sucessivamente (muito parecido com as alíquotas de vírus transmitidas durante uma infecção entre hospedeiros, ao se dar um espirro, por exemplo).

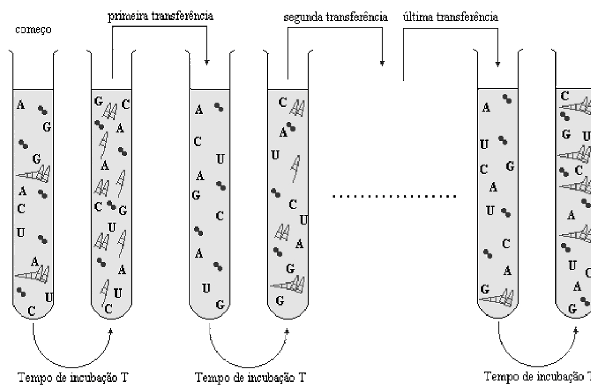


Figura 1.8: Desenho esquemático do método de passagens em série. O tubo inicial contém os quatro tipos de moléculas necessárias para a síntese do RNA, além da enzima replicase do fago Q β e uma molécula de RNA utilizada como semente. Esta macro-molécula é copiada muitas vezes com alguns erros. Depois de uma incubação por um tempo T , uma pequena amostra da solução (uma gota) é transferida a um novo tubo. Os processos de incubação e transferência são repetidos muitas vezes e as mudanças evolucionárias na população de RNA podem então ser observadas.

Durante o período de incubação aparecem variantes de RNA devido aos erros

de cópia na replicação, podendo-se observar então um processo de evolução. Além disso, o processo de passagens de um tubo a outro introduz no sistema uma forma sutil de *competição*, já que em cada operação de passagem os RNA de maior concentração, correspondendo àqueles com maior *fitness*, terão maior probabilidade de serem transferidos. O único critério nesse processo para definir o *fitness* das variantes produzidas é a velocidade de replicação e, desde que as moléculas menores replicam-se mais rápido, estas serão selecionadas mais vezes. A razão pela qual estas moléculas não diminuem em tamanho indefinidamente deve-se ao fato de que elas necessitam de um tamanho mínimo para manter a atividade de replicação como substratos das replicases na solução. A prova mais evidente de que são produzidas cópias cada vez *mais adaptadas* é a observação de que a velocidade de síntese das novas cópias cresce rapidamente nas transferências sucessivas. De fato, originalmente nos experimentos de Spiegelman, foi necessário diminuir o tempo T entre transferências, conforme se avançava na série. Ao cabo de aproximadamente 70 transferências o sistema pareceu estabilizar-se, e não se observou variação aparente no *fitness* das moléculas produzidas. O RNA resultante deste processo de seleção e evolução tinha perdido toda a sua capacidade infectiva, enquanto aumentou muito sua capacidade replicativa, única pressão de seleção existente nestes experimentos.

Vamos construir um modelo de micro-evolução com base no procedimento experimental de amostragens utilizado por Spiegelman, e guardar todas as analogias com os mecanismos por de trás da transmissão de uma infecção. Para realizar isso, discutiremos os passos de um algoritmo que contém as principais características desse procedimento, mas que pode ser estendido a sistemas mais complexos e a condições mais interessantes na prática como a replicação dos vírus *in vivo*. Portanto, vamos utilizar este experimento como sistema modelo, fazendo-se primeiramente algumas simplificações necessárias, no sentido de não nos distrairmos com detalhes de menor importância. Assim, consideraremos populações de patógenos com *tamanho constante* (N) sendo que cada indivíduo é representado por seu genoma, uma seqüência, com o *tamanho de cada seqüência* também *constante* (L). Mais ainda, consideraremos que os caracteres herdáveis de cada indivíduo estão codificados em uma seqüência binária de 0's e 1's ($\kappa = 2$), composta de L símbolos (teríamos $\kappa = 4$ para os ácidos nucléicos (G, A, C, U/T) ou $\kappa = 2$ se os agruparmos nos dois tipos de bases existentes (purinas e pirimidinas), ou ainda, $\kappa = 20$ para as proteínas). Cada seqüência é caracterizada pelo número de 1's que ela contém, sem levar em conta a posição desses elementos dentro da *seqüência*. Isto faz com que existam $L + 1$ tipos (ou classes) diferentes de *seqüências* possíveis (*espaço de seqüências*) os quais são identificados aqui pela variável inteira $P = 0, 1, 2, \dots, L$. Esta consideração é razoável, desde que as características que distinguem os genomas são suas taxas de replicação que, em muitas análises, têm sido escolhidas de modo a depender de P somente [18].

Com essas simplificações, e estabelecida a dinâmica a que nosso sistema está sujeito, podemos *mimetizar* a estratégia experimental descrita anteriormente considerando uma *sopa de seqüências* contendo os $L + 1$ diferentes tipos de seqüências nas proporções Π_P ($P = 0, 1, \dots, L$), da qual retiramos com reposição N indivíduos aleatoriamente. Assim, o estado $\mathbf{n}(t) = (n_0, \dots, n_L)$ da população está relacionado

com as freqüências Π_P através da distribuição de probabilidade multinomial

$$\mathcal{P}_{\Pi}(\mathbf{n}) = \frac{N!}{n_0! n_1! \dots n_L!} \Pi_0^{n_0} \Pi_1^{n_1} \dots \Pi_L^{n_L}. \quad (1.5.13)$$

Dessa maneira, em cada geração as seqüências dos descendentes são uma amostra da *sopa de seqüências* dos pais, similantemente ao experimento em que transferimos pequenas alíquotas entre os tubos de ensaio. Isto faz com que em cada geração a população seja uma combinação aleatória das seqüências constituintes. Embora esse procedimento destrua as correlações entre as seqüências, ele não causa qualquer perda significativa da informação genética desde que, nos mecanismos de replicação estudados aqui, a aptidão de uma seqüência para produzir descendentes em um dado meio depende somente do número de 1's em sua representação e que, em média, não é afetado pelo procedimento. Além disso, cabe ressaltar que a reprodução como descrita acima (mais propriamente, um processo de amplificação), diferentemente dos modelos de evolução existentes na literatura, leva em consideração os efeitos de população finita e, como veremos, descreve bastante bem o fenômeno da *deriva genética*, já que, a cada amostragem, a freqüência dos novos indivíduos deve ser um pouco diferente da contida na *sopa de seqüências* original, o que levará à fixação de um dos indivíduos depois de muitas gerações (ou passagens). Podemos, agora, inserir, nesse procedimento, as mudanças na composição da população \mathbf{n} . Seguindo a prescrição comumente usada na implementação de algoritmos evolucionários, em particular os chamados algoritmos genéticos [19], consideraremos primeiro o efeito da seleção natural e depois o efeito das mutações. A Figura 1.9 a seguir mostra esquematicamente esse algoritmo de amostragens [16] que define uma passagem no experimento de Spiegelman.

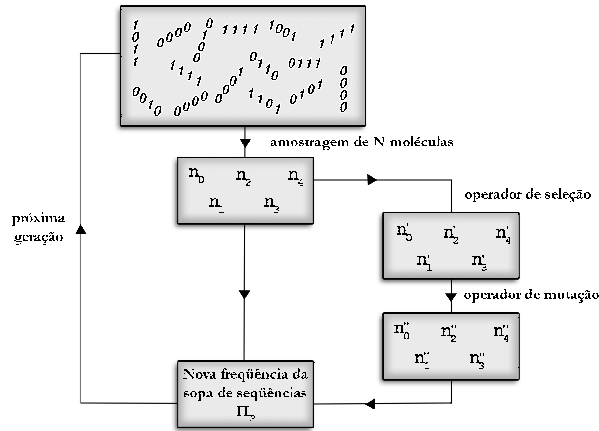


Figura 1.9: A sopa de seqüências inicial contém 2^4 tipos diferentes de moléculas nas proporções Π_P ($P=0, \dots, 4$). Depois da amostragem com reposição de N moléculas, cria-se um conjunto $\{n_P\}$ onde n_P é o número de moléculas com P 1's. A proporção de moléculas na sopa, na próxima geração, é diferente da inicial devido às flutuações introduzidas pela amostragem. Podemos inserir nesse procedimento mecanismos evolucionários, como seleção e mutação, que irão produzir mudanças no conjunto $\{n_P\}$.

Esse procedimento faz com que o tempo seja uma variável discreta que mede o número de gerações (gerações sem superposição). A rigor, aqui uma geração é igual ao tempo de incubação do procedimento experimental e, portanto, engloba várias gerações. Assim, o estado da população no tempo t pode ser descrito indicando-se o número de indivíduos ou *número de ocupação* para cada um dos $L + 1$ pontos do *espaço de seqüências*, denotado por $n_P(t)$. Portanto, dado o estado da população na geração t como sendo o vetor $\mathbf{n}(t) = \{n_P(t)\}$ (de modo que $\sum_P n_P(t) = N$), o processo evolucionário ocorrendo nas próximas gerações pode ser pensado como um *processo estocástico* em três estágios, cujas regras dinâmicas são: *reprodução*, não propriamente pertencente ao processo evolucionário, mas sua premissa; *seleção natural*, modelada pela taxa de replicação de cada indivíduo A_P ; e *mutação*, modelada pela taxa de erro por dígito $\mu = 1 - q$. Observe-se que a inclusão de mecanismos evolucionários, como seleção e mutação, é feita de maneira independente da amostragem; além disso, o algoritmo não depende do particular *agente* (a replicase, por exemplo) que define a característica (ou *fenótipo*) particular da seqüência que será selecionada. Após a amostragem dos N indivíduos, suporemos que o número de descendentes com que uma seqüência do tipo P contribui para a próxima geração seja proporcional a sua taxa de replicação relativa

$$W_P(\mathbf{n}) = \frac{n_P A_P}{\sum_R n_R A_R}. \quad (1.5.14)$$

Obviamente, os indivíduos com maior taxa de reprodução $W_P(\mathbf{n})$ terão maior probabilidade de aparecer nas próximas gerações, de maneira que a composição da população depois de selecionados os N novos indivíduos passa a ser descrita pelo vetor aleatório $\mathbf{n}'(t) = (n'_0, \dots, n'_L)$. Este procedimento de seleção não é trivial para populações finitas, desde que ele introduz correlações na população (essas correlações são responsáveis pela estrutura da árvore genealógica da população). Diferentemente da literatura ligada a algoritmos genéticos [19] ou a genética de populações [21], que tenta manipular essas correlações, nós simplesmente as eliminamos, usando uma amostragem adicional depois do procedimento de seleção. Podemos também considerar as mudanças em \mathbf{n}' devido às mutações, através das quais as seqüências herdadas por todos os indivíduos na população sofrem transições aleatórias, de maneira que *cada elemento* de uma *seqüência* é modificado com uma dada probabilidade μ , independentemente dos outros elementos (consideramos somente mutações pontuais). Com esse mecanismo simples, o estado da população depois da mutação é descrito pelo vetor $\mathbf{n}''(t) = (n''_0, \dots, n''_L)$.

Para derivarmos uma equação de recorrência simples a partir desse modelo, nós consideramos um procedimento semelhante à aproximação *annealed* da mecânica estatística de sistemas desordenados: ao invés de tomarmos a média sobre a população somente depois que o estado estacionário é alcançado, essa média é realizada a cada geração. Como no procedimento do modelo descrito acima, o resultado obtido $\bar{\Pi}_P(t)$, é então usado para construir a nova população. Naturalmente, essa aproximação não leva em conta as flutuações de $\Pi_P(t)$ para gerações diferentes. Assim a

freqüência média de seqüências do tipo P na geração $t + 1$ é escrita como

$$\bar{\Pi}_P(t + 1) = \frac{1}{N} \sum_{\mathbf{n}} \sum_{\mathbf{n}'} \sum_{\{n''_{PR}\}} \sum_R n''_{PR} \mathcal{P}_m(\{n''_{PR}\} | \mathbf{n}') \mathcal{P}_s(\mathbf{n}' | \mathbf{n}) \mathcal{P}(\mathbf{n}) \quad (1.5.15)$$

Portanto, nesse modelo a freqüência de moléculas do tipo P na próxima geração $\bar{\Pi}_P(t + 1)$, após o sistema ter sofrido as transições $\mathbf{n} \longrightarrow \mathbf{n}' \longrightarrow \mathbf{n}''$, é dada simplesmente por $(1/N) \sum_R n''_{PR}$. Esta freqüência é então usada como a nova freqüência da sopa de seqüências que será utilizada para gerar a nova população de N moléculas de comprimento L de acordo com a distribuição (1.5.13). Assim, durante o tempo de incubação do experimento, estão ocorrendo as transições que definem uma geração neste modelo. Este procedimento é então repetido a cada geração, simulando as passagens entre os tubos de ensaio. No capítulo 3 vamos estudar mais a fundo esses processos evolucionários tentando estabelecer suas conseqüências para a evolução de patógenos dentro de um hospedeiro que por sua vez participa da transmissão de uma epidemia.

Capítulo 2

Modelagem de Processos Epidêmicos

A despeito da simplicidade do modelo SIR utilizado em quase toda extensão do capítulo anterior, ele não só pode ser usado como um sistema modelo em que trabalhamos hipóteses como também explica razoavelmente bem o padrão observado para alguns tipos de infecções observadas no mundo real. Por exemplo, as curvas da Figura 1.2 daquele capítulo se ajustam satisfatoriamente (mediante a escolhas correta dos parâmetros do modelo) a disseminação do sarampo em uma população quando o número de casos é anotado semanalmente ou mesmo no caso da disseminação da peste bubônica quando os dados da infecção são integralizados diariamente. Como já havíamos comentado ainda no capítulo anterior, podemos buscar maior realismo da modelagem ao incorporar outros esquemas compartimentais, ou porque mudamos o esquema de setas que esquematizam a transição entre um conjunto de estados que a população deve passar ou porque simplesmente acrescentamos novos estados possíveis para a população modelada. Assim, por exemplo, podemos trabalhar com modelo SIRS admitindo perda de imunidade, quando um indivíduo recuperado permanece nesse estado por um período finito de tempo e voltando posteriormente ao estado susceptível, formando assim um ciclo. Tal modelo ainda bastante simples de se trabalhar é freqüentemente usado para o estudo de doenças recorrentes, onde uma pessoa pode ser contaminada mais de uma vez como é o caso da gripe, por exemplo. Da mesma forma, podemos adicionar um período de latência discriminando assim o período em que a doença se encontra no corpo de um hospedeiro, mas ainda não se tornou contagiosa. Para isso é comumente utilizado o modelo SEIR ou com recorrência o modelo SEIRS com um comportamento similar ao descrito para os modelos SIR e SIRS exceto pelo estado Exposto, no qual o indivíduo permanece infectado, mas não infectante por um período.

Vale salientar mais uma vez que tais categorias podem ser utilizadas em várias técnicas de modelagem distintas, como é o caso dos modelos compartimentais baseados em equações diferenciais parciais (também utilizadas para analisar a compo-

nente espacial nesses modelos como um processo de difusão) ou modelos baseados no indivíduo. Outros aspectos ainda podem ser modelados utilizando esse ferramental teórico como, por exemplo, considerar a demografia da população e trabalhar com uma população aberta (deixar o número total de indivíduos variar) no sentido de que queremos estudar os efeitos da taxa de mortalidade causados pela doença na população (fator que pode ser essencial ao se modelar doenças como a peste em populações da idade média ou mesmo das várias ondas de gripe que acometeram as populações do começo do século XX), assim, como a taxa de natalidade da população. Existe também uma vasta literatura [2] de modelos que consideram a estrutura etária da população, criando modelos mais complexos para estudar epidemias, bem como é possível implementar e estudar esquemas de vacinação aproveitando o conceito de limiar epidêmico para definir qual parcela da população devemos vacinar (quando uma vacina for disponível) de maneira que uma epidemia não se instale.

Neste capítulo iremos discutir vários graus de complexidade que podemos incorporar a modelagem do processo epidêmico, deixando vários desses aspectos de lado. Particularmente, estaremos aprofundando a discussão na influência de se incorporar o espaço como categoria de análise do espalhamento de uma epidemia, mas de qualquer maneira, estaremos sendo tendenciosos e restritivos deixando de explicitar outras formas modelar.

2.1 Modelando a heterogeneidade espacial e das interações

Vamos voltar à questão abordada nas seções 4 e 5 do capítulo anterior, quando estávamos preocupados em modelar os aspectos espaciais que envolvem a disseminação de uma epidemia. Para isso, nessas seções utilizamos dois estágios da modelagem: o primeiro criando um modelo baseado no indivíduo cuja ferramenta de modelagem foi o autômato celular e posteriormente introduzimos um grau de desordem nessas interações individuais para ajustar a velocidade de espalhamento da epidemia fazendo com que as regras de transição desse autômato fosse aplicado sobre um grafo mais geral que a rede regular. Como neste capítulo queremos modelar o processo epidêmico utilizando o mesmo ferramental teórico do capítulo 1, vamos fazer isso discutindo uma versão mais geral do autômato do capítulo anterior.

Esse outro tipo de abordagem, portanto, envolve um modelo estocástico geral de espalhamento de epidemias [22], e sua principal diferença com relação aos autômatos celulares tradicionalmente utilizados na epidemiologia é a utilização de dois tipos de interação entre seus indivíduos: as interações locais, que ocorrem devido à influência exercida pelos vizinhos infectados sobre os indivíduos susceptíveis do sistema, e as interações globais, nas quais todos os indivíduos têm igual probabilidade de entrar em contato entre si e, desta forma, possibilitar que tal disseminação ocorra por outras regiões da rede, distantes da onda local de infectados. Isso é feito de maneira simples, mudando um das regras de transição do autômato ao modelar que a probabilidade de qualquer susceptível tornar-se infectado é modelada como a *superposição* desses dois tipos de interações por meio da equação:

$$p_S = \Gamma p_G + \Lambda p_L \quad (2.1.1)$$

onde Λ e Γ são parâmetros utilizados para se ajustar as interações de curta (formação de clusters) e longa distância (do tipo campo médio) [2] entre indivíduos de uma mesma população (de maneira que a condição $\Lambda + \Gamma = 1$ seja satisfeita).

A influência global p_G é modelada pelo número total de indivíduos infectados que estão na rede $I(t)$, isto é, devido à presença e mobilidade de qualquer indivíduo infectado na população, descrita pela equação:

$$p_G = \frac{\rho}{N} \sum_{\{k,l\}} \delta_{i,\sigma(k,l)} \quad (2.1.2)$$

onde $0 \leq \rho \leq 1$ é um parâmetro do modelo que limita o valor máximo de p_G e está relacionado com a mobilidade intrínseca da população e $\delta_{i,\sigma(k,l)}$ é a função delta de Kronecher. Mais simplesmente a soma na equação 2.1.2 conta o número instantâneo de indivíduos infectados na população.

A influência local p_L pode ser calculada a partir da vizinhança de um determinado indivíduo, que pode conter um ou mais infectados. Ou seja, é exatamente a mesma equação 1.3.12 para p_L dada no capítulo anterior.

Com essa simples alteração e seguindo as prescrições de um autômato temos que a dinâmica de transmissão da infecção na população de indivíduos é governada pelas mesmas regras de interação ou contato, sendo que agora, qualquer indivíduo susceptível pode tornar-se infectado com uma probabilidade p_S (definida pela equação 2.1.1). Assim, a partir de tal modificação é estabelecido um link entre os dois cenários trabalhados até aqui, podendo o sistema, a partir dos parâmetros Λ e Γ , comportar-se como um modelo baseado no indivíduo tradicional (para $\Gamma = 0$), ou um modelo no qual efetivamente não há espaço (para $\Gamma = 1$) ou ainda um cenário intermediário interessante, no qual ocorre simultaneamente as duas abordagens (para $0 < \Gamma < 1$), o que possibilita atribuir ao sistema um comportamento similar, como veremos, ao gerado no protocolo de re-conexão do modelo de mundo pequeno original [15] sem, no entanto, a necessidade de se discriminar tais conexões.

Um resultado bastante instrutivo desta forma de modelar epidemias é observado ao retomar a o modelo SIR e nos atermos ao conceito de limiar epidêmico a fim de que se possa analisar o poder preditivo de R_0 sob o impacto da estocasticidade das interações, introduzido pelo protocolo desse autômato mais geral. Para fazer isso utilizando o autômato discutido, de uma maneira geral podemos caracterizar o estabelecimento de uma infecção em uma população em uma configuração populacional na rede como sendo

$$R(I(t)) = \left\langle \sum_S (\Gamma p_G + \Lambda p_L) \right\rangle_\tau \quad (2.1.3)$$

onde, aqui assumimos que em $t = 0$ existem $I > 1$ diferentes indivíduos infecciosos. A notação nessa equação indica depois que somarmos todos os danos prováveis sobre todos os susceptíveis dispostos na rede tomamos a média temporal durante o

período em que cada indivíduo infectado permanece nesse estado τ . Daí, podemos calcular o número total médio de contatos que apenas um indivíduo infeccioso faz durante o período de infecção:

$$R_0 = \Gamma \frac{\rho\tau S(0)}{N} + \Lambda\delta\lambda\tau \quad (2.1.4)$$

Na Figura 2.1, mostramos a probabilidade epidêmica, $P_E = n_e/n$, definida como sendo a fração de simulações ou populações experimentais, que produzem mais que 1 indivíduo infeccioso durante o período de infecção, onde: n_e é o número de experimentos de simulação em que um indivíduo infectado inicialmente na rede tenha infectado pelo menos um suscetível em um total de n experimentos; em função do número de reprodução básica R_0 , calculado segundo esse modelo alternativo. Aqui, por simplicidade os pré-fatores de ajuste são fixados em $\Gamma = \Lambda = 0,5$ e os parâmetros ρ e λ estão no intervalo $[0, 0,2]$. O espalhamento dos pontos no gráfico (principalmente para valores grandes de R_0) é devido ao fato de que no espaço de parâmetros (ρ, λ) existem diferentes combinações de ρ e λ que resultam (aproximadamente) no mesmo valor de R_0 .

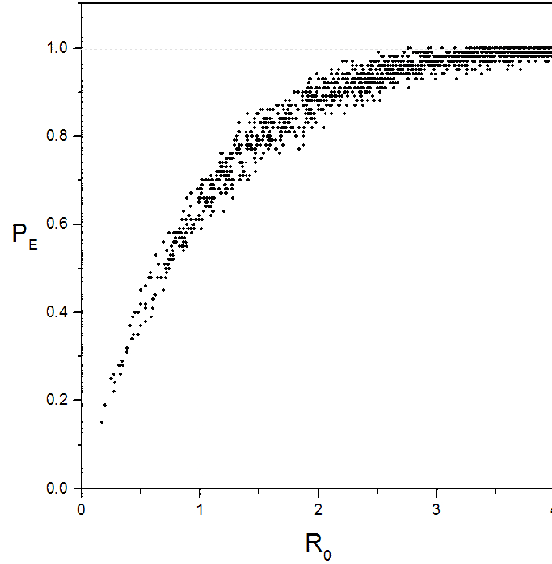


Figura 2.1: Probabilidade de epidemia P_E (definida como sendo a fração de simulações ou populações experimentais, que produzem mais que 1 indivíduo infeccioso durante o período de infecção) em função do número de reprodução básica médio R_0 .

A Figura 2.1 é instrumental para apreciarmos algumas propriedades do modelo introduzido nessa seção, mesmo quando limitado ao tentar explicar o conceito de limiar epidêmico. A quantidade P_E tende a saturar em $P_E \approx 1$ quando os valores de R_0 são suficientemente grandes ($R_0 \geq 3$), de modo que o espalhamento da

epidemia na população é sempre observado (em 100% das populações simuladas), ou seja, somente para valores grandes de R_0 , pode-se ter certeza de que a epidemia se espalhou. Por outro lado, mesmo para $R_0 < 1$ ainda é possível observar o estabelecimento de epidemias em algumas populações. Assim, de uma perspectiva do controle de epidemias, reduzir o R_0 efetivo abaixo de 1 (através de programas de vacinação, por exemplo) pode ser um problema potencial de estratégia, já que essa estratégia irá falhar em aproximadamente 40% dos eventos.

Dessa maneira, segundo o modelo alternativo colocado aqui, apesar do critério clássico do limiar de epidemia, é improvável reconhecer (usando somente dados de censo em populações) a iminência de qualquer desastre epidêmico, se o sistema está perto da região de limiar. Um critério mais preciso (embora frustrante) é entender que para qualquer valor de R_0 que um programa de vacinação possa forçar, sempre existe uma probabilidade (mesmo que pequena) da epidemia invadir uma população.

No que segue neste capítulo vamos explorar outras características dessa formulação alternativa e mais geral.

2.2 Epidemias na rede usando autômatos celulares

Uma aplicação mais ousada dessa formulação alternativa é a tentativa de se modelar a disseminação da Dengue em uma população, o que acarretará em introduzir nesse modelo o efeito da população de vetores (o mosquito *Aedes aegypti*) interagindo com a população de humanos. Antes de fazer isso entretanto vamos lembrar brevemente que a Dengue é uma doença comum em regiões tropicais do mundo causada por um arbovírus e transmitida para populações humanas (os hospedeiros), de todas as idades e de ambos os sexos, somente pela fêmea adulta do mosquito *Aedes aegypti* (o vetor). Este ciclo epidemiológico sustenta-se em áreas urbanas e após a picada a um indivíduo infectante, o mosquito nunca se recupera da infecção deixando de infectar novos hospedeiros somente com sua morte, sem nada sofrer com a presença do vírus em seu organismo.

Para descrever a dinâmica de transmissão do vírus da Dengue dentro da comunidade de hospedeiros, devido à interação com a população de mosquitos, classificamos os indivíduos inicialmente com uma estrutura similar aos modelos clássicos em epidemiologia de compartimentalização de uma população em classes de indivíduos, de acordo com o *status* de doença de cada um. Assim, a Dengue pode ser modelada admitindo-se que a população de hospedeiros com N_H indivíduos, em qualquer instante de tempo pode ser dividida nas seguintes classes: os indivíduos susceptíveis que podem ser infectados com o vírus ($S_H(t)$), os indivíduos que foram expostos ao vírus ao serem picados, mas que ainda estão incubando o vírus e, portanto não infectam ($E_H(t)$), os infectantes ($I_H(t)$) e os indivíduos recuperados da infecção ao adquirirem imunidade ($R_H(t)$). Uma simplificação adotada aqui é a de considerar que a população total de hospedeiros

$$N_H = S_H + E_H + I_H + R_H \quad (2.2.5)$$

é mantida constante. Esta é uma hipótese realística desde que queremos modelar a

infecção da população por um único sorotipo e que, portanto, não há mortalidade adicional causada pela doença (nesse caso a doença é benigna, e os indivíduos sintomáticos não são levados à morte [23, 24]). No caso da população de *mosquitos adultos* os N_V indivíduos também podem ser divididos em classes como: susceptíveis ($S_V(t)$) que não contém o vírus da dengue sendo inofensivo para a população de hospedeiro, infectados porém não infectantes ($E_V(t)$) chamados de expostos e infectantes ($I_V(t)$). Particularmente, consideramos que os mosquitos, uma vez na classe dos infectantes permanecem nesse estado até a sua morte, já que o vírus não interfere no ciclo de vida do vetor [23]. Diferentemente, da consideração feita sobre a população de hospedeiros, a população total de mosquitos

$$N_V = S_V + E_V + I_V \quad (2.2.6)$$

não é constante no tempo, pois o tempo de vida do mosquito varia de alguns dias a meses, muito menor do que a vida média de um ser humano. Do que vimos então a população de humanos é modelada por um esquema de transição de estado do tipo SEIR e a de mosquito por um esquema do tipo SEI. Nesse esquema a transmissão da doença começa no exato momento em que o mosquito infectado (I_V) pica um indivíduo que não tenha adquirido a doença e nem possua o vírus no seu corpo, indivíduo susceptível (S_H). Este hospedeiro passará para o estado infectado (I_H), quando pode ocorrer a transmissão do vírus para o vetor (mosquito) caso ocorra um encontro entre este indivíduo e um vetor que esteja no estado de susceptível (S_V). O período de infectado (I_H) na população humana dura cerca de 3 a 7 dias e é chamado de período infeccioso (viremia). Após a chegada do indivíduo para o estado de recuperado (R_H) aparentemente apresenta imunidade permanente ao vírus. A população de vetores comporta-se de maneira similar, todo vetor que atinge a fase adulta é susceptível (S_V) até que entre em contato com algum indivíduo que esteja no estado infeccioso (I_H). Ao efetuar a picada o vetor passa para o estado conhecido de exposto (E_V), este período corresponde ao tempo que o vírus leva para alojar-se nas glândulas salivares e efetivar a transmissão, este é o período de incubação extrínseco que varia de 7 a 10 dias. Ao tornar-se efetivamente um transmissor da doença ele permanece infectado até a sua morte no estado de infectante (I_V).

Como estamos trabalhando efetivamente com redes que localizam os indivíduos no espaço, esta interação entre as duas populações será dada entre dois autômatos distintos, um autômato para a população de hospedeiro e outro para a população de vetores, como pode ser visto esquematicamente na Figura 2.2. Com efeito, consideramos esta uma nova maneira de modelar em que além de se definir as regras de transição entre os estados em autômatos celulares para cada população interagente (de humanos e de mosquitos) definimos regras de interação entre esses dois autômatos.

Nesse sentido temos de pensar também na definição de um novo esquema de vizinhança, que agora deve ser estabelecida entre os dois autômatos (ver destaque na Figura 2.2). Assim, essas duas redes devem possuir um elo de ligação. Ou seja, a posição geográfica do elemento na rede é o que determinará o seu elemento

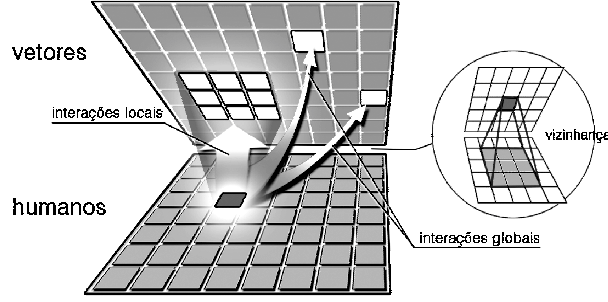


Figura 2.2: Esquema da interação entre dois autômatos representando a população de humanos e a de mosquitos, na modelagem da transmissão da Dengue. No destaque mostramos o novo esquema de vizinhança que tem de ser adotado nesse tipo de modelo.

correspondente na outra rede. Para exemplificar o que acabamos de discutir, se o elemento estiver numa posição [35,80] na rede dos vetores o seu correspondente será a área em volta do elemento [35,80] na rede dos hospedeiros. Portanto, essa área será determinada pelo tamanho da vizinhança, correspondente ao raio determinado a partir do centro que será o indivíduo [35,80], todos os indivíduos que estiverem dentro desse raio serão considerados para poder fazer o indivíduo alterar o seu estado.

Portanto, com esse novo protocolo de interação entre autômatos podemos manter para cada um as regras de transição de estados estabelecidas na seção anterior. Dessa maneira ficamos com as seguintes probabilidades para cada população: para a população de humanos especificamos que,

$$p_L = 1 - (1 - \lambda)^{n_{Im}} \text{ e } p_G = \frac{\rho_m N_{mi}(t)}{N_m} \quad (2.2.7)$$

e de maneira similar para a população de mosquitos ficamos com

$$p_L = 1 - (1 - \lambda)^{n_{Ih}} \text{ e } p_G = \frac{\rho_h N_{hi}(t)}{N_h} \quad (2.2.8)$$

onde n_{Im} é o número de vetores infectantes do autômato dos mosquitos que interagem como vizinhos de um indivíduo susceptível no autômato de humanos e n_{Ih} é o seu análogo representando o número de humanos infectados que podem ser picados por um mosquito susceptível. Nesse caso a interpretação da probabilidade λ é fundamental já que ela passa a ser a probabilidade de picada. É interessante notar que nesse modelo as probabilidades globais em cada população modelam as mobilidades que ambas podem ter na rede, em princípio distintas entre si. A Figura 2.3 mostra a simulação desse modelo em dois instantes distintos de tempo assumindo que inicialmente a rede tinha um humano infectado. Assim na parte a da figura podemos ver o espalhamento da epidemia nos primeiros instantes em que é detectado o aparecimento de um aglomerado de casos em torno desse primeiro caso.

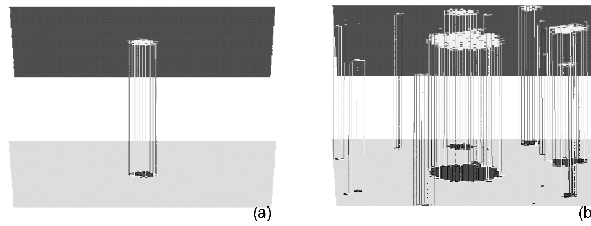


Figura 2.3: Simulação do modelo baseado na interação entre dois autômatos, em dois instantes distintos de tempo.

A parte b da figura mostra a epidemia já instalada na população e o aparecimento de vários aglomerados (clusters) na rede em lugares distintos dessa cidade virtual. Esse fato é o efeito da mobilidade das duas populações modeladas pelas probabilidades globais. Com esse ferramental é possível testar várias outras propriedades dos mecanismos mais gerais do espalhamento da Dengue. Realmente com esse laboratório virtual podemos fazer vários estudos adicionais em todo espaço de parâmetros [25]. Por exemplo, ao testar o efeito de fixar uma porcentagem da população inicial de mosquitos no estado de infectante o que acarreta em que a infecção se propaga preferencialmente de forma local, apenas os contatos mais próximos acabam por se infectar, e como os vetores estão distribuídos aleatoriamente, esta disposição dos vetores favorece o desenvolvimento da epidemia localmente. De uma maneira geral, é possível observar também que o tamanho da rede (da nossa cidade virtual) acaba influenciando bastante no dano causado pela probabilidade de picadas, já que a rede aumentando o número de indivíduos acaba aumentando também o número de possíveis casos de doença e conseqüentemente mais gente se tornaria infectada e de qualquer maneira a mobilidade da população de humanos acaba ganhando mais força à medida que mais gente se torna agente infeccioso no sistema (é importante deixar claro, todavia que nesse modelo não são adotadas medidas de controle para o avanço das ondas infecciosas). Mesmo assim, nesse caso podemos aferir os resultados frente a tamanhos de municípios diferentes, no mínimo, além de poder testar outros parâmetros espaciais importantes na dinâmica de transmissão da Dengue.

Uma análise total desse modelo foge completamente a motivação principal destas notas e do curso como um todo, já que o que queremos fundamentalmente é mostrar variantes no modo de modelar e entender o processo epidêmico. De qualquer maneira, esse modelo é instrumental tanto para a análise de uma epidemia da Dengue como também ao servir como um laboratório virtual para experimentar propriedades novas desse modelo baseado em autômatos celulares.

2.3 Redes complexas em epidemiologia

Uma das características fundamentais do modelo alternativo baseado no autômato celular probabilístico discutido na seção 2.1 e aplicado à epidemia da Dengue na seção anterior, é incorporar a mobilidade da população a partir da probabilidade p_G , similar ao protocolo do modelo de mundo pequeno de Watts e Strogatz [15], discutido na seção 4 do capítulo anterior. Com efeito, se retomarmos a formulação alternativa da seção 2.1 para o modelo SIR e refizermos os experimentos realizados com o protocolo de mundo pequeno obtemos exatamente o mesmo resultado que o observado na Figura 1.7 do capítulo 1. Isto quer dizer que de uma maneira geral o modelo alternativo baseado em autômatos celulares reproduz os mesmos resultados do protocolo utilizado para definir o fenômeno de mundo pequeno, inclusive nos permitindo observar esse mesmo fenômeno para situações mais complexas como o caso da Dengue discutido na seção anterior (o aumento da mobilidade da população de humanos e mosquitos aumenta a velocidade de espalhamento da epidemia). De uma maneira geral observamos, em ambos os modelos, que a estrutura da rede de contatos influencia a velocidade da transmissão da doença na população. O ponto alarmante e menos óbvio (que não aparece na modelagem clássica de um mundo sem espaço e nos modelos baseados em indivíduos modelados com interações entre primeiros vizinhos) é como que já para valores pequenos do parâmetro que mimetiza mobilidade, é suficiente para fazer o *mundo pequeno*.

Vamos avançar um pouco mais nessa discussão do efeito de considerar uma mobilidade da população em detrimento de aumentar a complexidade da rede de contatos ao se trabalhar com um parâmetro de desordem caracterizando esse grafo.

Particularmente, os modelos clássicos de epidemias, ao se considerar perda de imunidade, ou seja, modelos do tipo SIRS e SEIRS, apresentam um comportamento oscilatório. Tais oscilações podem gerar uma sincronização espontânea de uma fração significativa de elementos no sistema, ou seja, durante o desenvolvimento de uma epidemia com essa característica de perda de imunidade (com retorno dos indivíduos voltando ao estado susceptível), os elementos apresentam uma sincronização entre si quanto ao tempo em que se encontram infectados ou recuperados, fazendo com que transitem ao mesmo tempo para outros estados. Para entender melhor esse fenômeno Abramson e Kuperman [40] fizeram um estudo detalhado do modelo SIRS usando o protocolo de mundo pequeno para uma rede unidimensional, em particular introduzindo uma medida para o grau de sincronismo que o sistema apresenta dado por:

$$\sigma(t) = \frac{1}{N} \sum_{j=1}^N e^{i\phi_j(\tau)} \quad (2.3.9)$$

onde $\phi_j = 2\pi(\tau_j - 1)/\tau_0$ é uma fase geométrica que corresponde a τ sendo que em suas simulações eles levam em conta apenas a parte determinística dos ciclos. Assim, quando o valor de sincronismo σ é baixo encontramos poucos indivíduos em um mesmo estado na rede, no caso, susceptíveis, infectados e recuperados. Para valores grandes de σ , no entanto, muitos elementos se encontram simultaneamente

no mesmo estado (assim, a sincronização apresentaria um valor $\sigma = 1$ se todos os elementos se encontrassem no mesmo estado ao mesmo instante, sendo tal fenômeno fortemente ligado à estrutura da rede simulada). Eles fizeram várias simulações para entender como essa medida varia com o parâmetro de desordem p e o número de coordenação K (parâmetros que como vimos caracterizam essa rede de mundo pequeno). Para fazer comparações com esses resultados, fizemos várias simulações utilizando o autômato celular probabilístico em uma rede unidimensional. Não vamos discutir aqui os resultados mais gerais dessas duas maneiras de modelar, mas tomar um conjunto de experimentos para efeito de comparação.

Assim, nos gráficos da Figura 2.4 são mostrados vários resultados de simulações usando o autômato em redes de $N = 10^4$ indivíduos onde anotamos a fração de indivíduos infectados em cada instante de tempo. Aqui consideramos, $\tau_I = 4$ dias, $\tau_R = 9$ dias e uma taxa inicial dos indivíduos infectados $n_{inf}(0) = 0.1$. Também nessas simulações desse modelo alternativo, produzimos gráficos para três graus de mobilidade distintos: para $\Gamma = 0.01$ (gráfico do topo), $\Gamma = 0.2$ (gráfico do meio) e $\Gamma = 0.9$ (gráfico mais abaixo).

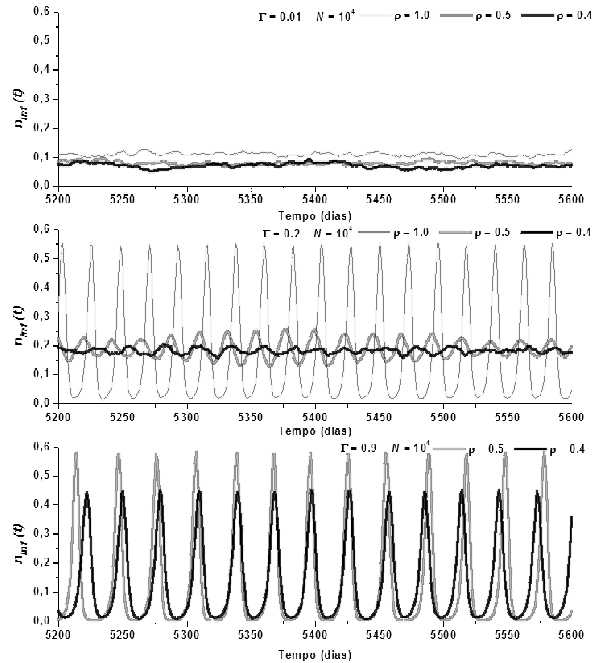


Figura 2.4: Fração de infectados em função do tempo (em dias), apresentada em três séries onde variamos o parâmetro global conforme apresentado nas legendas dos gráficos utilizando um autômato celular probabilístico do tipo SIRS.

Em cada gráfico apresentamos ainda três séries distintas variando o parâmetro ρ conforme apresentado nas legendas. Com efeito, as linhas em preto represen-

tam exatamente os valores obtidos por Abramson e Kuperman [40] ao utilizarem o protocolo de mundo pequeno reproduzidos agora utilizando o modelo baseado em autômatos celulares (é interessante notar nesses gráficos a idéia de sincronismo discutida anteriormente, ao observar as características das *ondas* geradas por essas séries). Ou seja, existe um conjunto de parâmetros no modelo baseado em autômatos que reproduzem exatamente os resultados do protocolo de mundo pequeno (como já havíamos observado antes para o modelo SIR). Dessa maneira, podemos notar que para um intervalo de 400 dias, acompanhando as linhas pretas passamos de um cenário endêmico para outro oscilatório, conforme o valor de Γ é acrescido. No caso, temos um cenário no qual a rede é formada por indivíduos dispostos de forma quase regular (para $\Gamma = 0.01$) e outro no qual a epidemia já apresenta uma característica periódica, gerando ondas de grande amplitude auto-sustentáveis (para $\Gamma = 0.9$). Para $\Gamma = 0.2$, contudo, nota-se um cenário intermediário no qual o comportamento do sistema passa a se alternar de endêmico para oscilatório e vice versa, havendo uma certa flutuação devido à própria configuração da rede, que se apresenta como uma transição de um cenário cujas ligações são regulares para outro em que elas são quase totalmente aleatórias. O comportamento então observado para $\Gamma = 0.9$, por exemplo, ocorre devido à dinâmica provocada pelos atalhos estabelecidos entre os vários elementos da rede, que torna possível a ocorrência de um maior sincronismo na transição de seus estados.

Contudo, o aspecto mais interessante do gráfico apresentado se manifesta ao pensarmos que todos esses cenários foram possíveis de serem obtidos para um único valor de ρ em nosso modelo alternativo, sendo, portanto, apenas um subconjunto de cenários possíveis. Isso pode ser notado ao compararmos tais resultados na Figura 2.4 com os outros obtidos para o mesmo valor de Γ , contudo ajustando a intensidade dessas relações a partir do parâmetro ρ na mesma figura. Na figura central, em que $\Gamma = 0.2$, conseguimos obter para um mesmo valor de mobilidade global Γ os três cenários que eram possíveis de serem obtidos no modelo de mundo pequeno ao varrer quase todo o conjunto de valores possíveis de p , simplesmente ajustando-se aqui o parâmetro ρ desse modelo alternativo, mostrando uma modelagem bem mais abrangente do que a baseada em grafos de Watts e Strogatz [15], pois possibilita a configuração de diferentes tipos de cenários para um mesmo grau de mobilidade (apenas ajustando sua *intensidade*, dada no modelo pelo parâmetro ρ).

É curioso notar que para $\rho = 1.0$ no último gráfico, a simulação não conseguiu atingir o número de dias do intervalo adotado (verificar a ausência dessa série no gráfico). Isto ocorre porque nesse nível de mobilidade a epidemia ocorre tão rapidamente que em um determinado momento a mesma se extingue por não encontrar elementos suscetíveis ao seu alcance e as oscilações não se mantêm dado o alto grau de sincronismo que o sistema consegue atingir para tal configuração de parâmetros.

Capítulo 3

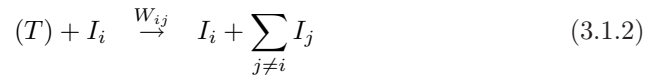
Modelagem da Evolução dos Patógenos

Na seção 5 do Capítulo 1 introduzimos os princípios de evolução darwiniana que regem a invasão e adaptação um patógeno dentro de um hospedeiro. Isso foi feito de maneira simples a partir de um algoritmo evolucionário desenvolvido para mimetizar um procedimento experimental de passagem em série. Esses princípios podem ser entendidos de uma maneira mais formal através da teoria de quase-espécies que foi derivada por M. Eigen e P.Schuster [26, 27], originalmente no contexto de evolução pre-biótica, para descrever a dinâmica de macro-moléculas replicantes (ácidos nucleicos) sob a influência dos mecanismos de seleção e mutação, inerentes da evolução darwiniana; mas em um sentido mais geral ela descreve qualquer população de organismos que se auto-reproduzem. Esse modelo, baseado em cinética química, é descrito através de um conjunto de equações diferenciais ordinárias para as concentrações dos diferentes tipos de moléculas que compõem a população de patógenos. Mesmo tendo sido proposta a mais de 20 anos, a teoria das quase-espécies vem tendo uma enorme importância nos últimos anos, devido à possibilidade de se interpretar populações de vírus e de algumas bactérias como uma quase-espécie [5, 28]. Assim, atualmente o conceito de *quase-espécies virais* forma a base teórica para o estudo da evolução de vírus caracterizados por altas taxas de mutação, como: influenza A, estomatite vesicular, polio-1, além de toda a família de vírus responsáveis pela hepatite e o mais conhecido: o HIV causador da AIDS [5, 29]. Na sequência desse capítulo estaremos também estudando mais detalhadamente as consequências dessa formulação.

3.1 Evolução darwiniana in silico

A formulação original de Eigen foca em um sistema modelo bem definido: o reator de fluxo. Basicamente ele consiste em um aparato no qual macro-moléculas biológicas são continuamente construídas a partir de monômeros ricos em energia T

(trifosfatos), que são utilizados para síntese macro-molecular e decaem depois de um certo tempo em monômeros deficientes em energia M (monofosfatos). Além disso, é assumido que o sistema pode trocar energia e matéria com o ambiente, regulando assim o suprimento de energia, a quantidade de monômeros deficientes de energia e o número total de macro-moléculas. Cada macro-molécula replicadora é modelada como uma seqüência de L dígitos $I_i = (s_1^i, s_2^i, \dots, s_L^i)$, com as variáveis s_j^i assumindo κ valores diferentes (como antes, teríamos $\kappa = 4$ para os ácidos nucleicos (G, A, C, U/T) ou $\kappa = 2$ se os agruparmos nos dois tipos de bases existentes (purinas e pirimidinas), ou ainda, $\kappa = 20$ para as proteínas), cada um representando um tipo diferente de monômero usado para formar a molécula (ou um genoma). Desse modo, para uma seqüência concreta da classe i , a série de eventos que podem ocorrer dentro do reator de fluxo podem ser modelados através de passos simples de reações químicas:



Nessas reações, os monofosfatos M (monômeros deficientes em energia) e os trifosfatos T (monômeros ricos em energia) são colocados entre parêntesis, desde que suas concentrações não são consideradas como variáveis. A reação (3.1.1) representa a auto-replicação fiel da molécula I_i e a reação (3.1.2) a auto-replicação com erro permitindo a formação de moléculas I_j devido à mutação de I_i . A matriz de replicação W leva em conta a estrutura primária das moléculas, sendo seus elementos dados por:

$$W_{ii} = A_i q^L, \quad (3.1.5)$$

$$W_{ij} = \frac{A_j}{(\kappa - 1)^{d(i,j)}} q^{L-d(i,j)} (1 - q)^{d(i,j)}, \quad i \neq j, \quad (3.1.6)$$

onde A_i é a taxa de replicação das moléculas do tipo i , $d(i, j)$ é a distância de Hamming entre as moléculas i e j , e $0 \leq q \leq 1$ é o parâmetro que mede a fidelidade de replicação de cada monômero, que é suposta a mesma para todos monômeros. Auto-replicação e mutação então, são representadas por duas reações de autocatálise, homogênea e heterogênea respectivamente. A reação (3.1.3) corresponde a degradação da molécula I_i , sendo a constante D_i a taxa de desagregação das moléculas do tipo i , e a reação (3.1.4) representa a saída para fora do sistema da molécula I_i por difusão sendo Φ_0 o coeficiente de difusão global que é suposto o mesmo para todas as moléculas.

As variáveis relevantes desse sistema dinâmico são as concentrações de cada seqüência individual $[I_i]$ ou, melhor ainda, as concentrações relativas das espécies

diferentes

$$x_i(t) = \frac{[I_i]}{\sum_{i=1}^L [I_i]}. \quad (3.1.7)$$

Assim, a evolução temporal da concentração x_i das moléculas do tipo $i = 1, 2, \dots, \kappa^L$ para esse sistema obedece a equação diferencial,

$$\frac{dx_i}{dt} = \sum_j W_{ij} x_j - [D_i + \Phi_0] x_i. \quad (3.1.8)$$

Esta formulação sintetiza o modelo de quase-espécies por meio de todo o conjunto de equações diferenciais ordinárias para as concentrações dos diferentes tipos de moléculas que compõem o reator. Este formalismo, entretanto, é válido somente no limite aonde o número total de moléculas N vai ao infinito.

Como foi dito anteriormente para que apareça um processo de seleção é necessário impor algum tipo de restrição ao sistema, que pode aparecer na natureza de muitas formas distintas. Para apreciarmos o efeito da competição em nosso sistema (que pode levar ou não à seleção), vamos imaginar a situação ideal em que a auto-replicação das moléculas se dá sem nenhuma possibilidade de erro. Nesse caso todos os elementos fora da diagonal principal da matriz de replicação W se igualam a zero e a equação (3.1.8) reduz-se a forma simples

$$\frac{dx_i}{dt} = [W_{ii} - D_i - \Phi_0] x_i. \quad (3.1.9)$$

É fácil ver que esta equação para a molécula da classe i tem uma solução exponencial trivial no caso em que não impomos nenhum vínculo, desde que podemos manter Φ_0 constante. Desse modo, se existirem k espécies de moléculas distintas dentro do reator, todas aquelas com taxa de replicação W_{kk} maior que $D_k + \Phi_0$, terão sua concentração aumentada exponencialmente sem limites a um tempo infinito. Ao contrário, as moléculas com taxa de replicação menor que $D_k + \Phi_0$, diminuirão sua concentração até desaparecerem da população.

Este comportamento está esquematizado na parte (a) da Figura 3.1, para o caso de cinco moléculas binárias ($\kappa = 2$, $s_k = [0, 1]$). Com relação ainda ao caso da replicação sem erro podemos impor agora um vínculo a que a população de moléculas dentro do reator esta sujeita. Aqui só discutiremos o caso em que a *concentração total de moléculas* dentro do reator $\sum_i x_i = N$ é mantida *constante*. Isto faz com que o termo Φ_0 varie no tempo, introduzindo uma não linearidade em (3.1.8) e em (3.1.9) que é determinada pela condição $\sum_i dx_i/dt = 0$. Portanto, o fluxo global de moléculas dentro do reator terá de ser controlado dependendo da replicação das moléculas de maneira que

$$\Phi_0 = \frac{\sum_i (W_{ii} - D_i)}{N}. \quad (3.1.10)$$

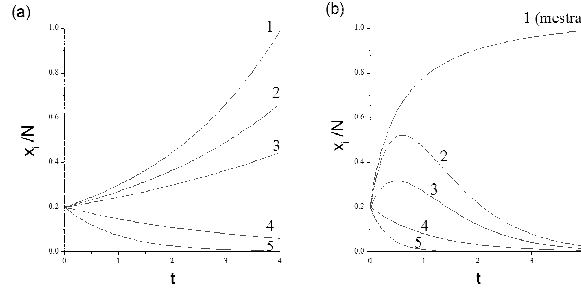


Figura 3.1: Variação da frequência de 5 tipos de moléculas binárias no reator de fluxo, sem possibilidade de erro no processo de auto-replicação. (a) segregação: dependendo dos valores de W_i , D_i e Φ_0 tem lugar um crescimento exponencial das moléculas. (b) seleção: uma vez submetidas a uma restrição de população constante, somente aquela com maior produtividade, $W_{ii} - D_i$, sobreviverá. No caso (a) mantivemos $\Phi_0 = 1.0$ e constante e em ambos os casos fizemos: $D_i = 0.5$, $W_{11} = 1.9$, $W_{22} = 1.8$, $W_{33} = 1.7$, $W_{44} = 1.2$ e $W_{55} = 0.5$.

A parte (b) da mesma figura esquematiza o comportamento temporal da mesma população do exemplo mostrado na parte (a) quando impomos a restrição de população constante ao sistema. Podemos fazer assim, algumas considerações qualitativas, que intuitivamente ajudem a compreender o comportamento do sistema. Se interpretarmos o termo $W_{ii} - D_i$ como a produtividade ou replicação líquida de cada molécula (ou genoma), o termo Φ_0 será a produtividade média da população, como pode ser visto da equação (3.1.10). De fato, diferentemente do caso sem restrição em que Φ_0 era mantido constante, ao se impor que a população total fique constante, temos que ir aumentando o fluxo global cada vez mais. Isto por sua vez faz com que, à medida que o tempo passe, um número maior de tipos de moléculas seja segregado da população de acordo com (3.1.9). Ao final dessa competição, somente a molécula com maior produtividade que a média será *selecionada* frente a todas as outras da população. O processo de seleção é, portanto, um efeito do meio através de um vínculo imposto à população. No estado estacionário desta competição, chamamos de *seqüência mestra* (I_m) aquela que é selecionada quando a população de moléculas é submetida a algum tipo de restrição e o estado estacionário desse sistema é denominado *equilíbrio de seleção*.

Vamos agora discutir o caso, representado pela equação (3.1.8), em que é levada em consideração a possibilidade de erros no processo de auto-replicação. As soluções no equilíbrio de seleção para essa equação podem ser encontradas, em termos dos auto-valores e auto-vetores da matriz de replicação W [30, 31]. Se definirmos o vetor $\mathbf{x} = (x_1, x_2, \dots, x_L)$, cujas componentes representam a concentração de cada espécie molecular dentro de toda população, podemos escrever

$$W\mathbf{x} = \lambda\mathbf{x}. \quad (3.1.11)$$

Assim, definimos uma quase-espécie precisamente em termos matemáticos, como

sendo o auto-vetor dominante $\mathbf{x} = (x_1, x_2, \dots, x_L)$ ao qual pertence o maior auto-valor λ_{\max} da matriz de replicação W . Este auto-vetor descreve a estrutura exata da população de quase-espécies: cada mutante I_i esta presente na quase-espécie com uma frequência x_i (nós podemos normalizar a concentração total de modo que $\sum_i x_i = 1$). O maior auto-valor é exatamente a taxa de replicação média da quase-espécie, $\lambda_{\max} = \sum_i A_i x_i$. Dessa maneira, a frequência de um dado mutante dentro da quase-espécie não depende apenas de sua taxa de replicação, mas também da probabilidade com que ele é produzido por erro na replicação de outras espécies moleculares e sua frequência na população. Portanto, um dos resultados cruciais da teoria é que, no equilíbrio, a seleção não leva em geral a uma população homogênea de um tipo de *indivíduo* molecular mais apto, mas sim a um *ensemble* de variantes geneticamente distintas, mas bastante próximas. Para distinguir do conceito clássico de espécie já existente em biologia, este *ensemble* é então chamado de *quase-espécie*. Segundo essa teoria, portanto, a adaptação é uma propriedade não de uma seqüência, mas dessa distribuição de mutantes centrada em torno da *seqüência mestra* (a com maior *fitness*, ou valor reprodutivo, do ensemble). A mutação $\mu = (1 - q)$ nesse modelo é o parâmetro que caracteriza a largura da distribuição, ou seja, o quanto à *quase-espécie* esta espalhada no espaço de genomas. Esse novo conceito de quase-espécie tem importantes implicações, desde que, evolução é normalmente pensada como a interação entre mutação e seleção, este último fator sendo aquele que favorece mutantes com vantagens reprodutivas que tenham sido gerados puramente ao acaso (realmente, é considerado um erro pensar em mutações sendo guiadas de outra maneira que não ao acaso). Uma quase-espécie, contudo pode guiar mutações: isto não significa que há qualquer correlação entre o ato (intrinsecamente probabilístico) da mutação e a vantagem seletiva de um mutante, mas que seleção opera sobre a estrutura de toda a quase-espécie, que por sua vez é adaptada ao seu relevo replicativo ou *fitness landscape* (este termo foi originariamente introduzido por Sewall Wright [32]). Portanto, evolução pode ser guiada na direção dos *picos* deste relevo, isto acontecendo porque os mutantes mais bem sucedidos (que podem estar perto dos picos do relevo) irão produzir mais filhos que os mutantes (que podem estar longe dos picos). Assim, esta *otimização evolucionária* pode ser vista [26, 27] como uma *escalada da montanha reprodutiva* das quase-espécies, que ocorre por meio de certos caminhos no espaço de seqüências. Há rotas alternativas, mas seu número é tão restrito que se tem a impressão de um fenômeno de direcionamento automático para a fixação dos mutantes produzidos com maior valor seletivo. As rotas evolucionárias são, neste grau, dirigidas deterministicamente.

3.2 Paradoxos evolucionários e a evolução de uma doença

À medida que a taxa de mutação aumenta, o modelo prevê que a composição da população deixa de estar distribuída como uma *quase-espécie* e passa a um regime, no qual a distribuição dos κ^L tipos de moléculas é *uniforme* (todas as seqüências aparecem em proporções iguais). Em outras palavras, como comentamos antes,

se a replicação ocorresse livre de erros, nenhum mutante apareceria e a evolução cessaria; contudo a evolução também deve ser impossível se o erro na replicação for muito alto (se houver poucos mutantes na população, isso pode melhorar a adaptação, porém se houver muitos, eles irão levar a deterioração da população). A transição entre esses dois regimes é outro importante resultado da teoria, conhecido por *limiar* ou *catástrofe de erro de replicação*. É interessante ver como o fenômeno do *limiar de erro* manifesta-se para a equação (3.1.8) que caracteriza o modelo, permitindo-se que muitos tipos de seqüências estejam dentro do reator. Para isso, precisamos definir primeiramente o mais simples, e talvez o mais utilizado [26, 27, 33, 34] relevo de replicação em que aparece esse fenômeno, que é conhecido como *relevo de replicação de um pico*, e pode ser visto como uma aproximação local de um relevo rugoso, mas com picos distantes. Matematicamente atribuímos a taxa de replicação $A_m = a > 1$ à seqüência mestra, e $A_d < a$ para as seqüências remanescentes. A Figura 3.2 mostra a distribuição de quase-espécies para seqüências binárias de tamanho $L = 30$, depois que o equilíbrio de seleção foi alcançado, em diferentes valores de taxa de mutação $\mu = (1 - q)$. Nesse caso, o número possível de seqüências no reator de fluxo é $N_{es} = 2^{30}$.

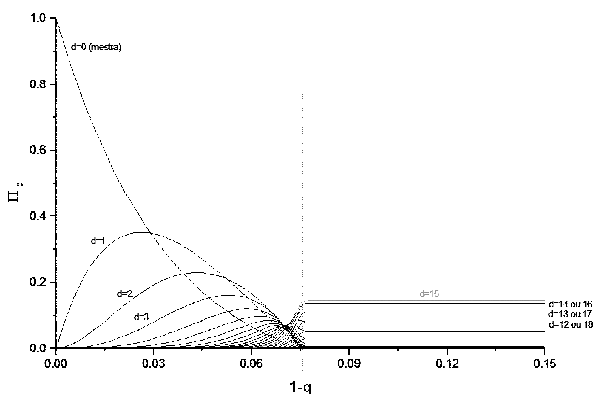


Figura 3.2: Frequência de seqüências em função da taxa de mutação $\mu = (1 - q)$. d é o número de diferenças entre uma seqüência particular e a mestra que tem o maior fitness. Então a curva para $d = 2$ é a proporção de seqüências diferindo da mestra por exatamente duas mutações. Na figura as seqüências binárias tinham 30 monômeros de comprimento, onde tomamos $a = 10$. Destacamos os primeiros mutantes da mestra, que já aparecem na população para valores pequenos de μ , bem como alguns mutantes mais distantes que só começam a aparecer para valores de mutação perto do limiar de erro. Em torno do limiar de erro todas as seqüências são igualmente prováveis. Desde que o número de seqüências com $d = 14$ e 16, por exemplo, são o mesmo, a frequência das duas classes são iguais.

Portanto, para uma melhor visualização, as seqüências são agrupadas em classes de mutantes (como feito na seção 5 do capítulo 1): todas as seqüências d mutações da mestra são membros da classe d . Fazemos todas as d classes com taxa de replicação $A_d = 1$ e a seqüência mestra ($d = 0$) com $A_0 = a$. Este procedimento de

estratificação reduz o número de equações da representação da cinética química, caracterizada por (3.1.8), à um conjunto de $L + 1$ equações diferenciais de primeira ordem acopadas [35], representando cada uma das classes de moléculas dentro do reator. No caso do relevo de replicação de um pico em questão estas equações são do tipo

$$\frac{d\Pi_d}{dt} = \sum_{R=0}^{L-1} M_{dR}\tilde{\Pi}_R + a\Pi_0 M_{P0} - \Pi_d [1 + \Pi_0(a - 1)], \quad (3.2.12)$$

onde Π_d agora, denota a concentração de molécula da classe $d = 0, \dots, L$, com $\sum_d \Pi_d = 1$. A figura então mostra a frequência de cada uma das seqüências mutantes e a da mestra, em função de μ , no equilíbrio de seleção. Cada uma das curvas é obtida fazendo-se $d\Pi_d/dt = 0$ em (3.2.12). Ainda nessa figura destacamos a linha de transição separando o regime caracterizado por uma *quase-espécie* e o regime uniforme: o *limiar de erro*. Para derivarmos uma expressão para o fenômeno de *limiar de erro* de uma maneira bastante simples [41, 36], vamos assumir que a população consiste de duas subpopulações: uma contendo a seqüência mestra (I_m), ou a mais apta dentre todas; e outra formada pela *cauda de erro* da quase-espécies, ou seja, todos os mutantes da distribuição. Estas últimas são substituídas por uma seqüência média (I_{ce}). Podemos também desprezar a pequena probabilidade de que uma seqüência mutante replique a uma mestra (*back mutations*) desde que é muito mais provável que elas mutem à outras seqüências mutantes. Dessa maneira, a probabilidade de que a mestra replique sem erro é q^L . Com essas hipóteses simplificadoras a equação (3.1.8) é reescrita para as duas sub-populações como:

$$\frac{dx_m}{dt} = A_m q^L x_m - x_m (A_m x_m + A_{ce} x_{ce}) \quad (3.2.13)$$

$$\frac{dx_{ce}}{dt} = A_{ce} x_{ce} + A_m (1 - q^L) x_m - x_{ce} (A_m x_m + A_{ce} x_{ce}), \quad (3.2.14)$$

onde fizemos ainda, por simplicidade, todas as constantes $D_i = 0$ e tomamos $x_m + x_{ce} = 1$, como o vínculo de população constante para haver seleção: uma seqüência só pode crescer as expensas das outras. As parcelas entre parênteses, em ambas equações, correspondem a Φ_0 . Estamos interessados na coexistência da mestra e os mutantes. A alternativa é que somente os mutantes estão presentes no equilíbrio: a seleção não consegue evitar o decaimento mutacional da mestra. Portanto, nós fazemos ambas equações iguais a zero, e perguntamos se neste equilíbrio pode acontecer $x_m \neq 0$. Isto requer, no caso de o relevo de replicação ser de um pico, que

$$q^L > \frac{A_{ce}}{A_m} = \frac{1}{a}. \quad (3.2.15)$$

O que quer dizer que q tem que ser maior que um valor crítico $q_c = (A_{ce}/A_m)^{1/L}$ para que a seqüência mestra não seja perdida da população. Este valor crítico é então interpretado como o *limiar de erro* de replicação. Isto leva a uma relação importante entre a precisão da replicação q e o comprimento da seqüência L

$$L < \frac{\ln a}{(1 - q)}, \quad (3.2.16)$$

que é obtida tomando-se o logaritmo de (3.2.15) e fazendo-se a aproximação $\ln(q) \approx q - 1$. Esta equação significa que a quantidade de informação que pode ser seletivamente mantida (L) é limitada pela fidelidade de cópia por dígito (q). A relação descrita pela equação (3.2.16) é mostrada na Figura 3.3.

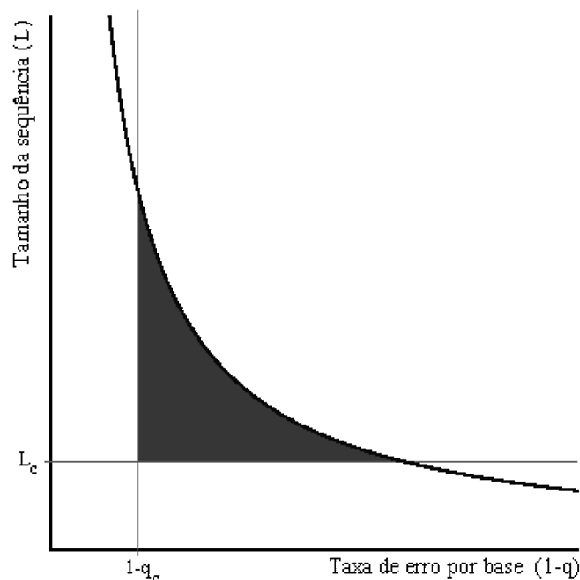


Figura 3.3: Tamanho de seqüência permitido, L em função da taxa de mutação $\mu = (1 - q)$ (área escura). A hipérbole é construída para $\ln a = 1$, onde a é a vantagem seletiva da seqüência mestra sobre seus mutantes. Há outros vínculos: o tamanho da seqüência não pode exceder o limiar L_c para poder codificar informação suficiente e a taxa de erro não pode ser reduzida abaixo de um valor crítico $1 - q_c$ sem um custo excessivo de tempo e energia na replicação.

Quanto maior a fidelidade na cópia, maior é a seqüência mestra que pode ser seletivamente mantida. Se a cópia mestra cresce perto do *limiar de erro* ela deve rapidamente deteriorar-se. Abaixo do *limiar de erro*, entretanto, a população deve consistir da seqüência mestra circundada por uma *nuvem* de seus mutantes mais próximos: a quase-espécie. Assim, outra das conseqüências do fenômeno da catástrofe de erro de replicação é conhecida como *paradoxo de Eigen*: o tamanho de uma molécula sujeita a esse mecanismo de replicação não pode exceder a um certo comprimento máximo L_c , fixado o grau de precisão ou confiabilidade do mecanismo de cópia; por outro lado, o sistema de moléculas não pode aperfeiçoar seu mecanismo de cópia enquanto seu comprimento for limitado, pois para instruir uma enzima, por exemplo, a fazer cópias com alta fidelidade necessita-se de um conteúdo mínimo de informação, maior do que o permitido pelo limiar de erro. Esta versão moderna do *paradoxo do ovo e da galinha* levou Eigen e Schuster [37] a proporem um modelo incorporando um mecanismo fechado de colaboração ou catálise entre as moléculas.

las auto-replicas, o qual chamaram de *hiperciclo*, assim evitando a *catástrofe de erro* (a descrição desse modelo foge do contexto deste trabalho). Nós exploramos nesta seção um sistema simples de moléculas auto-replicas e descobrimos que, se o sistema é fechado, no sentido de que uma cópia errada de uma molécula é uma das outras espécies presentes, o resultado será uma coleção estável das várias espécies, ou como foi chamada: uma *quase-espécie*. Contudo, o modelo tem muitos inconvenientes: além de determinístico é contínuo e de uma maneira geral nenhuma dessas propriedades é particularmente interessante para um sistema tal como uma sopa molecular, que consiste de um número relativamente pequeno de moléculas em movimento aleatório. Um modelo mais realístico deve ter duas características: ser discreto e estocástico.

Para ilustrar o quão errado o modelo baseado em equações diferenciais pode ser, suponha que nós tivéssemos um sistema de N moléculas, cada uma de uma espécie diferente. Considere um caso ideal onde $A_i = D_i$ e $q^L = 1$, de maneira que as taxas de nascimentos e de mortes sejam iguais e não há nenhum erro na reprodução. Nesse caso, todos os elementos da matriz de replicação W são zero, e o modelo prediz que nada irá mudar no sistema. Mais ainda, reações químicas ordinárias envolvem *algumas* cópias de espécies moleculares, cada uma das quais presentes em *número grande* de cópias (6.02×10^{23}). No caso da evolução acontece ao contrário, o número de polinucleotídeos possíveis N_{es} é muito maior que o número de moléculas presentes em qualquer experimento realístico que se queira realizar (por exemplo $2^{100} \approx 10^{30}$) e, portanto a adoção desse formalismo tem que ser considerada cuidadosamente sempre que o limite determinístico é usado.

3.3 A caixa de Pandora e a teoria de quase-espécies virais

Em um delicioso artigo de divulgação publicado em 1993 [28], Eigen revê o seu modelo no contexto de evolução viral nos remetendo a uma analogia da mitologia grega bastante intrigante: "Quando a curiosa Pandora finalmente abriu uma caixa proibida, deixou sair todas as misérias e demônios conhecidos no mundo. Uma dessas pragas eram sem dúvida os vírus". Com efeito, a despeito de qualquer mito, o conceito de *quase-espécies virais* é o paradigma teórico da dinâmica evolucionária de vários tipos de populações de vírus, que atualmente são o centro das atenções dos estudos experimentais, bem como das estratégias de tratamento de doenças causadas por esses organismos.

Realmente, os vírus podem ser vistos como *programas genéticos* com uma única mensagem para a célula hospedeira: - *reproduza-me!*...Todavia, outras característi-

cas são contundentes nesses *quase-organismos*. As populações de vírus, em geral, são formadas por uma distribuição de mutantes largamente dispersa no espaço de genomas, ao invés de uma população homogênea de seqüências do tipo selvagem (a seqüência mestra), mostrando que esse tipo de *vida* é possível com uma alta taxa de erro de replicação, quando comparadas com organismos autônomos.

Suas taxas de mutação são adaptadas ao seu comprimento e desde que adaptação depende da distribuição de mutantes, vírus diferentes podem mostrar estruturas de população bastante diferentes, mesmo sendo similar no comprimento. Realmente, as taxas de erro de replicação tem sido determinadas experimentalmente para vários tipos de vírus, sendo que todos os resultados mostram correlação entre taxa de mutação e comprimento da seqüência.

Como pode ser observado então, para os vírus de uma maneira geral, a substituição do termo *espécie* pelo de *quase-espécie* não é simplesmente semântica: uma *espécie viral* é uma população que se *autoperpetua* composta de entidades diversas mas relacionadas, atuando como um todo. Entretanto, tanto a reprodução, quanto à interpretação de como a mutação opera para gerar as seqüências, de uma virose específica, terão que ser tomadas cuidadosamente, nos levando à algumas modificações no modelo de quase-espécies original.

Assim, realizamos esta análise [18] retomando a modelagem efetuada na seção 5 do capítulo 1. Ali, consideramos que os caracteres herdáveis (*alelos*) de cada indivíduo estão codificados em uma seqüência binária ($\kappa = 2$) de 0's e 1's e como antes, composta de L símbolos (*locus*). Cada genoma é caracterizado pelo número de 1's que ele contém, sem levar em conta a posição desses *alelos* dentro do *genoma*. Isto faz com que existam $L + 1$ tipos (ou classes) diferentes de *genomas* possíveis (*espaço de genomas*) os quais são identificados pela variável $P = 0, 1, 2, \dots, L$. Finalmente, assumimos que todos os indivíduos da população na geração t dão lugar a seus filhos na geração $t + 1$ (gerações sem superposição).

Logo, o estado da população no tempo t pode ser descrito, indicando para cada um dos $L + 1$ pontos do *espaço de genomas*, definindo o estado da população na geração t como o vetor $\mathbf{n}(t) = \{n_P(t)\}$, (de modo que, $\sum_P n_P(t) = N$). Daí, o processo evolucionário ocorrendo nas próximas gerações foi representado como um *processo estocástico* em três estágios, cujas regras dinâmicas são: *reprodução*, *seleção natural*, e *mutação*.

Na Figura 3.4, mostramos uma comparação entre esses modos de modelar ao mostrar o espaço de seqüências gerado por todas as possíveis mutações e o espaço de seqüências reduzido em classes de acordo com a distância de Hamming d e com o número de 1's que cada seqüência contém. No exemplo, o comprimento das moléculas é fixado em $L = 4$, o que nos leva aos 5 tipos de seqüências possíveis,

mostrados como exemplo no capítulo anterior.

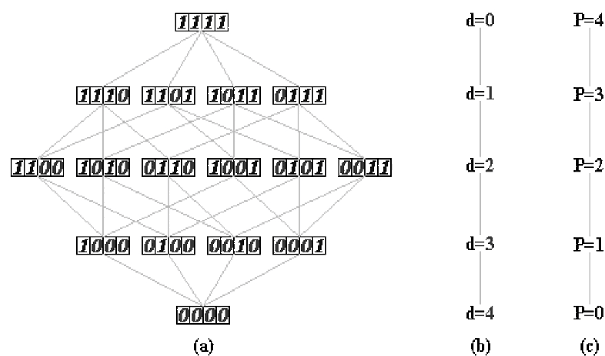


Figura 3.4: Comparação entre os espaços de seqüências. (a) Espaço gerado no modelo original de quase-espécies onde cada seqüência, das 2^4 possíveis, é obtida a partir de mutações pontuais das seqüências mais próximas a mestra $\{1, 1, 1, 1\}$. (b) Espaço reduzido onde cada conjunto de degenerescências da posição dos monômeros na seqüência é representado pela distância de Hamming à seqüência mestra (diferença do número de zeros em relação a mestra). (c) Mesmo que (b) para o número de 1's que cada seqüência contém.

Vamos a seguir derivar, da equação (1.5.15), uma equação de recorrência para a freqüência média de moléculas do tipo P na população. Naturalmente, isso só será possível se fizermos algum tipo de aproximação na dinâmica do modelo. Primeiramente, dadas as freqüências Π_P para $P = 0, \dots, L$ na geração t , vamos calcular o valor esperado do número de moléculas do tipo P na próxima geração. Essa grandeza é dada por

$$\bar{n}_P(t+1) = \sum_{\mathbf{n}} \sum_{\mathbf{n}'} \sum_{\{n''_{PR}\}} \sum_R n''_{PR}(t) \mathcal{P}_m(\{n''_{PR}\} | \mathbf{n}') \mathcal{P}_s(\mathbf{n}' | \mathbf{n}) \mathcal{P}_\Pi(\mathbf{n}), \quad (3.3.17)$$

onde os somatórios levam em conta todas as possibilidades para os números de ocupação gerados pelas freqüências $\Pi_P(t)$ como também as subseqüentes transições devidas aos procedimentos de reprodução e mutação. Esse processo de média é, obviamente, análogo a média sobre diferentes *rodadas* das simulações ou sobre diferentes populações, exceto que aqui esta média é realizada a cada geração, enquanto que nas simulações do modelo de amostragens ela é efetuada somente depois que o estado estacionário foi atingido. Esse modo de tomar médias sobre todas as possíveis realizações do sistema estocástico a cada geração, naturalmente não leva em conta as flutuações de n_P em diferentes populações. De fato, estuda-se a evolução de uma *população média* que pode ser totalmente diferente de qualquer uma das populações específicas cujas evoluções são simuladas.

Lembrando que a frequência de moléculas do tipo P na geração $t + 1$ é simplesmente $\bar{\Pi}_P(t) = \bar{n}_P(t)/N$, poderíamos obter uma relação de recorrência *fechada* se substituíssemos Π_P pela sua média em $\mathcal{P}_{\Pi}(\mathbf{n})$. Essa substituição, que despreza as flutuações dessa frequência, caracteriza a aproximação de *campo médio* utilizada por diversos autores [21, 20].

Entretanto, essa aproximação é exata para o modelo utilizando o algoritmo evolucionário, uma vez que, conforme discutido na seção 5 do capítulo 1, a própria definição do modelo incorpora esse tipo de substituição e a conseqüente quebra de correlações.

Daí, a equação de recorrência fundamental, é a seguinte

$$\bar{\Pi}_P(t + 1) = \sum_{\mathbf{n}} \sum_R M_{PR} W_R(\mathbf{n}) \mathcal{P}_{\bar{\Pi}}(\mathbf{n}). \quad (3.3.18)$$

É fácil verificar a condição de normalização $\sum_P \bar{\Pi}_P(t + 1) = 1$, uma vez que $\sum_R W_R(\mathbf{n}) = 1$ e $\sum_P M_{PR} = 1 \forall R$, respectivamente.

É de fundamental importância que se enfatize o único tipo de aproximação utilizado na derivação de (3.3.18), a saber, a desconsideração das flutuações do número de ocupação n_P para diferentes populações a cada geração.

Conforme mencionado, isso nos leva a estudar a evolução de uma população média, que pode não ser representativa das populações simuladas. Para melhor apreciarmos as conseqüências deste modelo, consideramos o limite $N \rightarrow \infty$, na eq. (3.3.18) que então reduz-se a

$$\bar{\Pi}_P(t + 1) = \frac{\sum_{R=0}^L M_{PR} \bar{\Pi}_R(t) A_R}{\sum_{R=0}^L \bar{\Pi}_R(t) A_R}, \quad (3.3.19)$$

que é a solução exata (tem o mesmo estado estacionário) do modelo original de quase-espécies.

Portanto, a equação (3.3.18) sintetiza um processo evolucionário, obtendo-se no regime determinista ($N \rightarrow \infty$) a mesma fenomenologia observada na teoria de quase-espécie. Mais do que isso, podemos estudar os efeitos de uma população finita (onde as flutuações do processo reprodutivo são levadas em conta) e que no modelo original não é possível ser feito.

Cabe, ainda, ressaltar que a eq. (3.3.18) é facilmente iterada para vários tipos de relevos de replicação e pode nos fornecer a variação do *fitness* médio da população a cada geração [18].

Outro fenômeno interessante, característico de populações finitas, é o chamado *escape estocástico* [21], [18], onde a seqüência mestra e mesmo alguns mutantes devem ser perdidos devido à flutuações na população. Este fenômeno é importante particularmente quando estamos descrevendo a dinâmica na qual a população se aproxima do equilíbrio [18], onde a finitude da população é um ingrediente essencial.

Vale a pena fazer uma breve observação sobre suas propriedades. Para irmos mais adiante com nossa análise, devemos especificar a taxa de replicação A_P para cada indivíduo na população através de um relevo de replicação particular.

Assim, se especificarmos na equação (3.3.19) os dados relativos ao relevo de replicação de um pico, isto é, atribuímos $A_L = a$ e $A_P = 1$, para $P \neq L$. Isso nos leva a equação de recorrência para as frequências de seqüências do tipo P ,

$$\begin{aligned} \bar{\Pi}_P(t+1) &= M_{PL} [\bar{\Pi}_L(t)]^N \\ &+ \sum_{n_L=0}^{N-1} B_{n_L} \frac{\sum_{R=0}^{L-1} \bar{\Pi}_R(t) \left[M_{PR} + a \frac{r}{1-r} M_{PL} \right]}{1 + r(a-1)} \end{aligned} \quad (3.3.20)$$

para $P = 0, \dots, L$.

Aqui, introduzimos a notação

$$B_{n_L} = \binom{N-1}{n_L} [\bar{\Pi}_L(t)]^{n_L} [1 - \bar{\Pi}_L(t)]^{N-1-n_L}, \quad (3.3.21)$$

e $r = n_L/N$.

Então, dada a frequência molecular inicial média $\bar{\Pi}_P(t=0)$ para $P = 0, \dots, L$, a equação (3.3.20) é iterada até o regime estacionário ser alcançado.

A equação (3.3.20) no regime determinístico ($N \rightarrow \infty$) é a solução exata da equação cinética (3.2.12) para L finito. Nesse caso, a soma sobre n_L na equação (3.3.20) é dominada pelo inteiro mais próximo de $(N-1)\bar{\Pi}_L(t)$, de maneira que nesse regime $r \rightarrow \bar{\Pi}_L(t)$, e a equação de recorrência reduz-se a

$$\bar{\Pi}_P(t+1) = \frac{\sum_{R=0}^{L-1} M_{PR} \bar{\Pi}_R(t) + a M_{PL} \bar{\Pi}_L(t)}{1 + (a-1) \bar{\Pi}_L(t)}. \quad (3.3.22)$$

Esta equação apresenta o mesmo estado estacionário da equação (3.2.12) do modelo original de quase-espécies [18] quando nos restringimos ao relevo de replicação de um pico. Para ilustrar como o *modelo baseado amostragens* mantém as principais características do modelo de quase-espécies, iteramos a equação de recorrência determinística (3.3.22) até alcançar o estado estacionário e apresentamos os resultados na Figura 3.5 apresentada a seguir.

É importante lembrar, entretanto, que o modelo desenvolvido nesta seção é baseado em um sistema modelo (o método de transferência em série) diferente do adotado para derivar o modelo original de quase-espécies (o reator de fluxo). A diferença principal reside na forma de implementar a diluição: descontínua no experimento com transferências em série e contínua no reator de fluxo. Esta diferença é que faz com que o modelo descrito aqui seja discreto e estocástico como desejável para um modelo mais realístico, no sentido de que leva em conta os efeitos de população finita.

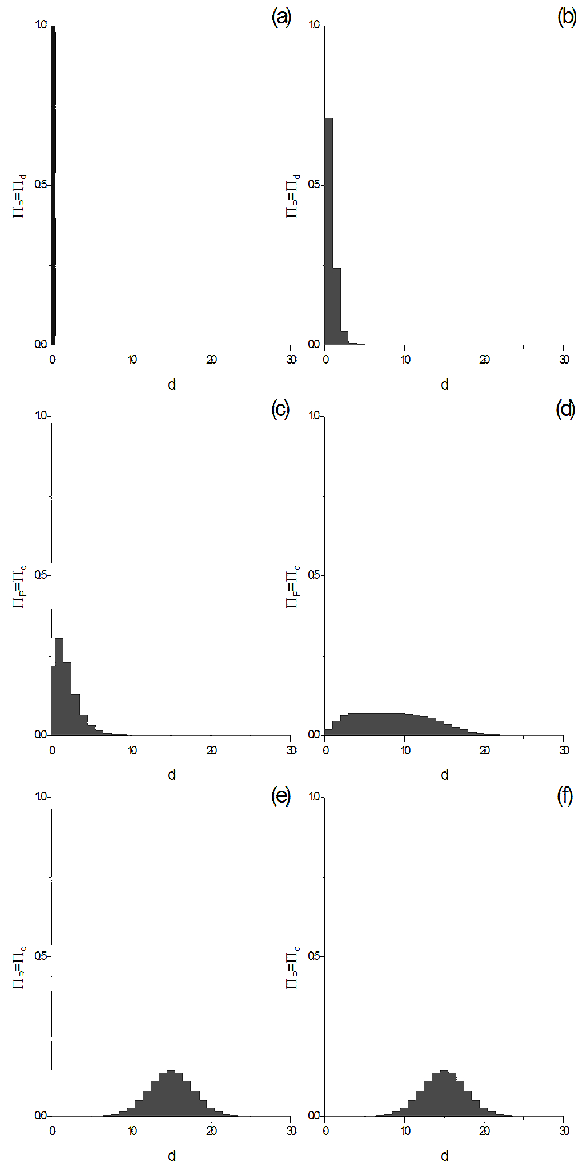


Figura 3.5: Distribuição das frequências de seqüências $\Pi_P = \Pi_d$ (P é o número de 1's e d é o número de zeros) em função da distância de Hamming de cada classe de seqüências à mestra no equilíbrio de seleção. (a) Destacamos a distribuição em que $\mu = 0$, para ser tomada como referência; (b) $\mu = 0.01$; (c) $\mu = 0.04$; (d) $\mu = 0.07$; (e) $\mu = 0.10$; e (f) $\mu = 0.13$. Aqui fizemos $a = 10$ e $L = 30$ e também tomamos as frequências iniciais como $\Pi_L(0) = 1$ e $\Pi_P(0) = 0$; $P = 0, \dots, L - 1$.

3.4 O paradigma da modelagem da AIDS

Talvez o exemplo mais importante de uma quase-espécie viral seja o HIV. Pacientes infectados com este vírus abrigam uma população de vírus extremamente diversa, com muitos mutantes diferentes. As mutações nesse caso são geradas na codificação da enzima transcriptase reversa do vírus, que é produzida com uma taxa de erro da ordem de 10^{-4} a 10^{-1} por base, ou seja, durante cada replicação de todo o genoma, de 1 a 10 erros acontecem, número este suficientemente grande para que o HIV opere muito perto do *limiar de erro de replicação*.

Assim, se acrescentarmos a informação de que o genoma do HIV contem 10000 nucleotídeos (cada um deles podendo assumir uma dentre 4 formas possíveis: A,U,G,C), é fácil deduzir que em cada ciclo de replicação, cada novo genoma é diferente de seu pai por mais de uma base. Estas observações levaram a formulação de um modelo [5, 29], que resume a importância potencial da diversidade dos antígenos como o principal fator da progressão da doença.

A idéia essencial é de que o vírus *evita* a pressão do sistema imunológico produzindo continuamente novos mutantes resistentes a esse ataque e resultando na acumulação de uma diversidade de antígenos durante o período assintomático. Particularmente, pode-se derivar um *limiar de diversidade de antígenos* [38] a partir das interações assimétricas entre a quase-espécie viral e o sistema imune, ou seja, as respostas do sistema imune específicas a uma linhagem, são direcionadas contra um antígeno específico do HIV; porém cada linhagem pode prejudicar todas as respostas imunológicas independentemente de sua especificidade.

Portanto, o aumento da diversidade do HIV, permite a população viral escapar ao controle imunológico. Nesse contexto então, a variabilidade genética da quase-espécie viral é responsável pelo fato de que o vírus estabelece uma infecção persistente sem ser dissolvida pela resposta imunológica e induz a AIDS depois de um longo tempo e com um tempo de incubação variável.

Para entender um pouco melhor essa questão vamos nos ater a infecção do HIV dentro de um indivíduo como um todo. Em geral, uma infecção pelo vírus HIV-1 segue um padrão com características quantitativas relativamente bem definidas. O *alvo* principal da infecção por HIV é uma classe de linfócitos, ou células brancas, chamadas células T4. Essas células secretam fatores de crescimento e diferenciação que são requeridos por outras populações celulares do sistema imune.

Assim, devido a sua importância na regulação desse sistema, a depleção de células T4 causada pelo ataque do vírus, tem um efeito deletério devastador no seu funcionamento como um todo, levando a imunodeficiência que caracteriza a AIDS: quando a contagem de células T4 é menor que 200 mm^{-3} em um paciente com HIV (a contagem normal é de aproximadamente 1000 mm^{-3}), então esse paciente é classificado como tendo AIDS.

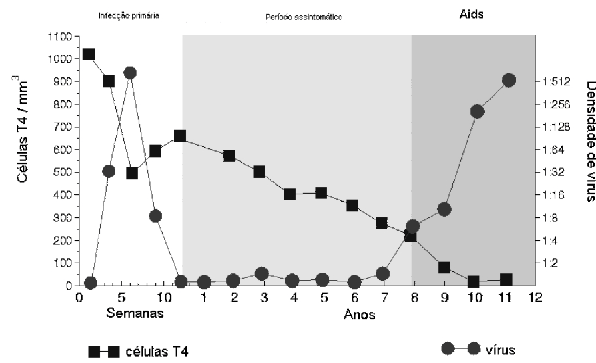


Figura 3.6: Padrão típico exibido por pacientes infectados por HIV. Densidade de vírus e da contagem de células T4 em função do tempo. São destacadas as três fases da dinâmica característica da doença

O tempo médio entre a infecção com o vírus e o início da AIDS é aproximadamente de 10 anos, sendo que, esse *time lag* esta relacionado à diminuição de células T4 circulando no sangue. Entretanto, não são conhecidos os processos que determinam a taxa de diminuição da contagem de células T. Essas células são normalmente recolocadas no corpo e a infecção pode afetar a fonte de novas células, ou o processo homeostático que controla o número delas no corpo. Embora o HIV possa matar células que ele infecta, somente uma fração pequena de células T4 (10^{-4} a 10^{-5}) são produtivamente infectadas em cada instante de tempo. Então, em adição a morte direta de células T, o HIV pode ter muitos outros efeitos indiretos. Para se ter uma noção do padrão dessa doença, o curso típico de uma infecção por HIV é mostrado na Figura 3.6.

Imediatamente depois da infecção, a quantidade de vírus detectada no sangue aumenta dramaticamente. Durante esse aumento da carga viral, aparecem alguns sintomas parecidos aos da gripe que desaparecem depois de algumas semanas, coincidindo com a diminuição da concentração do vírus a níveis muito baixos. Uma resposta imune então ocorre, podendo-se detectar anticorpos contra o vírus no sangue. Passado esse período de *infecção primária*, a concentração de carga viral se mantém relativamente constante e intrigantemente baixa por anos, durante os quais se observa uma lenta diminuição na contagem de células T4 no sangue enquanto o paciente não apresenta qualquer sintoma da doença. Com isso em mente, vamos focar nossa atenção na modelagem matemática de um aspecto particular, dentre os possíveis, que caracterizam a dinâmica de progressão da AIDS. Em especial, vamos discutir o que está acontecendo dentro de um indivíduo infectado durante esse *período assintomático*, analisando *criticamente* uma das possíveis explicações para a cinética de declínio das células T. É tentador pensar que, o fato de a doença demorar 10 anos em média para se desenvolver, implica que os componentes do processo devam também ser lentos. Isso é incorreto e na verdade as escalas de tempo que

acompanham a doença correspondem a processos biológicos importantes. Como já havíamos discutido antes, a idéia central do modelo baseado na diversidade viral é de que o vírus *evita* a pressão do sistema imunológico produzindo continuamente novos mutantes resistentes a esse ataque, resultando então em uma acumulação de uma diversidade de antígenos durante o período assintomático, a partir das interações assimétricas entre a *quase-espécie viral* e o sistema imunológico, ou seja, as respostas desse sistema, específicas a uma linhagem, são direcionadas contra um antígeno específico do HIV; porém cada linhagem pode prejudicar todas as respostas imunológicas independentemente de sua especificidade. Assim, as hipóteses básicas desse modelo de *evolução viral* durante a infecção podem ser resumidas como segue: (1) A resposta imune a uma infecção cria uma sub-população de células imunes específicas x_i , que direcionam o ataque imunológico a uma linhagem específica de vírus v_i . (2) A resposta imunológica é também caracterizada por uma resposta geral z , não específica, que atua contra todas as linhagens possíveis de vírus. (3) Cada tipo de vírus (mutante) pode matar todas as $y = z + \sum_i x_i$ células imunes: específicas e não específicas. Matematicamente, o modelo resultante dessas hipóteses pode ser expresso em forma de equações diferenciais não-lineares para as taxas de variação no tempo, da densidade de cada população envolvida no processo (onde $t = 0$ é o instante em que o hospedeiro adquire a infecção). Assim, para a população viral, temos:

$$\frac{dv_i}{dt} = f_i(v_i, y) - v_i(s_i z + p_i x_i), \quad (3.4.23)$$

onde $f_i(v_i, y)$ é a taxa de replicação de cada linhagem i do vírus definida como:

$$f_i(v_i, y) = (r'_i + r_i y) v_i. \quad (3.4.24)$$

Aqui, $r_i y$ representa a taxa *per capita* de novas linhagens que surgem da infecção das células T4, enquanto que r'_i modela uma reprodução eventual e baixa dos vírus em outras células que não as T4. O termo $v_i s_i z$ e o termo $v_i p_i x_i$ representam a morte do vírus do tipo i , com $i = 1, 2, \dots, n$, devido às células da resposta imune geral (z) e específica (x_i), respectivamente, com s_i e p_i , sendo as probabilidades de encontro entre o vírus com cada uma dessas células. Ainda podemos modelar a população de células do sistema imune através do seguinte conjunto de equações:

$$\frac{dy}{dt} = K - ay - uv y \quad (3.4.25)$$

representando o total de células T4, onde K é a taxa de reprodução dessas células no *timus*, a a taxa de morte *per capita* e o termo $uv y$ denota a taxa na qual essas células são mortas por qualquer um dos vírus da população total v . Da mesma maneira, podemos escrever para as células T4 específicas a uma linhagem:

$$\frac{dx_i}{dt} = kv_i y - uv x_i \quad (3.4.26)$$

e para a resposta imune geral:

$$\frac{dz}{dt} = k' v y - uv z. \quad (3.4.27)$$

Equivalentemente à equação (3.4.25), os termos kv_iy em (3.4.26) e $k'vy$ em (3.4.27) denotam novas células sendo ativadas e juntando-se as populações de linfócitos existentes da resposta específica e geral, assim como, os termos uvx_i e uvz representam a taxa de morte dessas células pelo vírus e para uma derivação detalhada do modelo ver [5, 29, 38]. O modelo inteiro é um processo estocástico no qual as equações que governam a dinâmica assumem a forma de equações diferenciais como explicadas acima. Apesar de Nowak e seus co-autores [5, 29], não especificarem o processo estocástico para a geração de uma nova linhagem de vírus, o número de tipos de vírus $n(t)$ é uma variável aleatória, que é implementada como tal em suas simulações. Assim, a taxa com que um novo tipo de vírus aparece na população, devido a mutação de um outro já existente, em um intervalo de tempo $[t, t + \Delta t]$ é dado por $\mu v(t)\Delta t$, onde, $v(t) = v_1 + v_2 + \dots + v_n$ é a população viral total e μ é a taxa de mutação do vírus. Colocado de outra maneira, *ativamos* a equação do vírus do tipo i toda vez que um gerador de números aleatórios produz um número p satisfazendo a condição

$$p \leq 1 - \exp[-\mu v(t_i)\Delta t]. \quad (3.4.28)$$

Por simplicidade, a população inicial de cada novo mutante é tomada como sendo a mesma quantidade de vírus v_0 da infecção inicial.

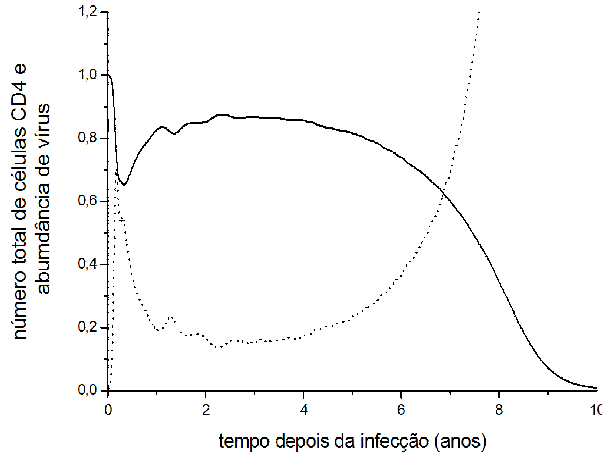


Figura 3.7: Simulação da variação temporal da quantidade de vírus e contagem de células T4. Parâmetros utilizados: $K = 100$, $k = k' = 0.01$, $u = 1$, $r'_i = 3r_i$, $s_i = 9.5r_i$, $p_i = 20r_i$ (todos com dimensão anos^{-1}). r_i foi tomado de uma distribuição exponencial, com média igual a 0.05. A taxa de mutação, quantidade inicial de vírus e número de vírus distintos no início da infecção são respectivamente: $\mu = 4$, $v(0) = 0.05$ e $n(0) = 1$.

A Figura 3.7, mostra uma simulação do modelo para a variação temporal da abundância do HIV e para a contagem de células T4 em um indivíduo. De uma maneira geral, pode-se notar que o modelo reproduz o padrão da doença discutido anteriormente: inicialmente observa-se um pico de viremia (característico da fase

de *infecção primária*) em um curto espaço de tempo, onde a resposta imune é capaz de reduzir drasticamente a quantidade de vírus. Contudo, à medida que o tempo passa, novos mutantes aparecem que não são reconhecidos pela população de células T4 pre-existentes, gerando mini-eclosões de viremia. Essas mini-eclosões, são outra vez suprimidas pela combinação de resposta imune específica e geral. À medida que o número de mutantes distintos vai crescendo no tempo, a população total de células T4 começa a declinar, até que depois de um período longo a população viral como um todo aumenta sem controle.

Com já comentado, pode-se derivar um *limiar de diversidade de antígenos*, ou seja, para entender os mecanismos não-lineares do modelo, podemos simplificá-lo assumindo que todos os tipos de vírus tem a mesma taxa de replicação r e que os parâmetros s e p são também constantes e independentes da linhagem do vírus. Com essas simplificações a equação (3.4.23) fica:

$$\frac{dv_i}{dt} = v_i(r - sz + px_i). \quad (3.4.29)$$

Com efeito, é fácil ver que, para um número fixo n de equações, o sistema imune pode controlar a população de vírus do tipo i , se $dv_i/dt < 0$, ou ainda, se $r - sz + px_i < 0$. Portanto, o sistema imune pode controlar toda a população de vírus v , se a desigualdade vale para todos os n valores de i . Assim, somando-se em i , ficamos com a restrição:

$$n < n_c = \frac{px}{r - sz}. \quad (3.4.30)$$

Nesse contexto, portanto, o aumento da diversidade do HIV acima de um limiar n_c permite à população viral escapar ao controle imunológico, ou seja, a variabilidade genética é responsável pelo fato de que o vírus estabelece uma infecção *persistente* sem ser dissolvida pela resposta imunológica, induzindo a AIDS depois de um longo tempo. Em outras palavras ainda, de acordo com esse modelo, a infecção por HIV é um processo evolucionário que determina a escala de tempo da infecção à doença [5, 29, 38]. Para o modelo mais geral o critério do limiar de diversidade pode ser generalizado facilmente como:

$$\frac{\sum_{i=1}^n (r_i - s_i z) u}{p_i k_i} > 1 \quad (3.4.31)$$

Aqui, é possível perceber que algumas variantes com uma taxa de reprodução alta, ou alta patogenicidade das células T4, podem influenciar desproporcionalmente na violação do limiar de diversidade. Apesar dessa classe de modelo estocástico reproduzir o padrão desejável da dinâmica da doença, ele pode ser criticado. Por uma questão de completeza, vamos nos ater as críticas específicas a esse modelo de evolução viral, que se referem à robustez de suas predições, explorando o comportamento dinâmico estocástico. Para isso, testamos o efeito de variar a taxa de mutação μ , mantendo todos os outros parâmetros constantes de tal forma que o limiar de diversidade de antígeno $n_c = 10$, também fique constante. Como esperado, as simulações mostram um forte viés da taxa de mutação sobre o período de

incubação do vírus. Assim, ao diminuirmos a taxa de mutação a partir de $\mu = 4$, observamos um aumento rápido desse tempo de incubação até o ponto em que, já para o valor de $\mu = 2$ observamos a supressão dos vírus na maioria das infecções. Essa intrigante característica indica que o número de mutantes que ocorrem na fase inicial da infecção é crucial para o processo como um todo, sendo que para poucos mutantes, o limiar de diversidade raramente será alcançado. É importante destacar que esse estranho padrão está intrinsecamente ligado a construção do modelo e à relevância do conceito de limiar de diversidade de antígenos. Uma idéia que seria interessante testar, é *recuperar* esse modelo de evolução viral de maneira a incluir alguns detalhes na dinâmica. O primeiro passo que pode ser dado é inserir no modelo uma descrição mais detalhada da estrutura da população viral como uma *quase-espécie* (como discutido na seção anterior). Apesar de populações de HIV-1 serem consideradas quase-espécies, para indicar a sua heterogeneidade genética, uma população desse vírus em um indivíduo infectado nunca alcança a distribuição de mutantes de equilíbrio por causa das mudanças rápidas que sofrem devido à pressão seletiva exercida pelo sistema imune. No modelo apresentado anteriormente, não existe nenhuma relação entre os diferentes mutantes do vírus, nenhuma distância genética. Assim, em um modelo subsequente poder-se-ia trabalhar inicialmente uma descrição *genotípica* da população viral de acordo com a seguinte alteração na equação (3.4.24),

$$f_i(v_i, y) = \left(\sum_{j=1}^n r'_j Q_{ij} + r_i y \right) v_i. \quad (3.4.32)$$

Aqui a matriz de mutação Q_{ij} define a probabilidade para obter o mutante i da replicação (errônea) da variante j e a distância genética entre mutantes diferentes. É importante destacar que esse tipo de detalhamento introduz complicações nas simulações que tem de ser adaptadas ao modelo em questão. De qualquer maneira, outras simulações demonstram ainda uma forte dependência do modelo às condições iniciais e juntamente com os resultados obtidos para a dependência com a taxa de mutação, devemos questionar a relevância do conceito de limiar de diversidade de antígenos, desde que ele caracteriza somente parcialmente a dinâmica do modelo.

3.5 A batalha entre o sistema imune e os patógenos

Vamos aproveitar a discussão iniciada na seção anterior para estudar um conceito bastante interessante. O conceito de *memória imunológica* é importante em qualquer consideração da interação entre parasita e hospedeiro. Se tomarmos como exemplo uma infecção viral, a imunidade humana pode ser duradoura contra um novo ataque do mesmo patógeno caso a resposta imune seja suficientemente forte (como no caso de doenças infantis), ou, devido a estratégias evolutivas do patógeno, um novo processo infeccioso se iniciará, independente da resposta à infecção anterior (como no caso da gripe e da cólera). Se essa imunidade é mantida através de constante exposição à infecção, ou através de clones de linfócitos duradouros, ou

ainda, devido à persistência do patógeno em baixos níveis de abundância dentro do hospedeiro (baixo nível de anticorpos do hospedeiro) permanece obscuro até o presente. Por exemplo, no caso da malária; existe um mecanismo de resposta mais complexo que implica numa imunidade lentamente adquirida através da apresentação constante do antígeno, o que indica uma aparente desregulação no desenvolvimento da memória sobre este antígeno [2]. A hipótese mais comum para explicar a memória imunológica é baseada na presença das células de memória com vida longa no organismo, todavia, alguns trabalhos questionam essa hipótese, sugerindo que a memória imunológica pode ser um estado dinâmico de uma rede regulatória dentro do sistema imune, ou da persistência do agente infeccioso dentro do hospedeiro [2]. A idéia que trabalhamos aqui é o conceito de memória imunológica fazendo-se uma comparação entre essas duas hipóteses, em dois níveis de modelagem.

Em um primeiro nível de aproximação tenta-se modelar como o organismo constrói uma memória imunológica de maneira a prevenir a re-invasão do hospedeiro pelo mesmo agente patogênico. Assim, para descrever a interação entre o sistema imune e o patógeno, consideramos um sistema de antígeno dirigido. A densidade do vírus e da célula efetora (célula que exerce funções imunológicas) é definida por $V(t)$ e $E(t)$, respectivamente, em um instante t . Portanto, podemos assumir de forma simplificada que as células efetoras são recrutadas a uma taxa constante Λ (mimetizando a produção de linfócitos na medula espinhal) e morrem a uma taxa μ , ou seja, $1/\mu$ é a expectativa de vida das células efetoras. Para a população viral podemos ter uma taxa de nascimento e morte, b e d , respectivamente; e definimos a taxa de crescimento natural do vírus como sendo $r = b - d$. Matematicamente, as mudanças nessas quantidades com relação ao tempo, são representadas pelas seguintes equações diferenciais:

$$\frac{dE}{dt} = \Lambda - \mu E + \varepsilon V E, \quad (3.5.33)$$

$$\frac{dV}{dt} = rV - \sigma V E. \quad (3.5.34)$$

A equação 3.5.33 representa a variação temporal de células efetoras. Nela, o termo $\varepsilon V E$ é a taxa de proliferação líquida das células efetoras devido ao contato com os vírus, onde ε é o coeficiente de proliferação dessas células. Em contrapartida, a equação 3.5.34 representa a variação temporal da quantidade de vírus. Como pode ser observado, modelamos que as células efetoras matam os vírus existentes nesse compartimento a uma taxa $\sigma V E$, onde σ é o grau de contato entre essas células e o vírus e modela a eficácia com que elas matam os vírus. Assim, dadas as quantidades iniciais de células efetoras $E(0)$ e de vírus $V(0)$, podemos gerar os valores de $E(t)$ e $V(t)$, para qualquer valor de t . Analiticamente, podemos estudar os pontos de equilíbrio do sistema, primeiramente, estabelecendo que $V(0) = V_0$ e em seguida qual critério fará com a população de vírus irá crescer acima de V_0 , ou seja, estabelecemos que:

$$\frac{dV}{dt} > 0 \Rightarrow r > \frac{\sigma \Lambda}{\mu} \quad (3.5.35)$$

Portanto, se essa inequação é satisfeita, o sistema irá alcançar o equilíbrio com o

vírus persistindo, onde:

$$E = \frac{r}{\sigma} \text{ e } V = \frac{\mu\varepsilon}{r} \left(r - \frac{\Lambda\sigma}{r} \right) \quad (3.5.36)$$

definem os pontos de equilíbrio do sistema de equações diferenciais estabelecidas para o modelo, com uma grande propensão do sistema exibir um comportamento oscilatório até o equilíbrio. De uma maneira geral, da interação do sistema imune com o antígeno replicante, podemos extrair duas conclusões (levando em consideração que esse modelo é uma primeira aproximação). A primeira é que fornecida a expectativa de vida das células efetoras, a memória imunológica criada para sua proliferação é suficiente para suprir a alta quantidade de vírus para os níveis iniciais da epidemia dentro do hospedeiro. Em outras palavras, o período de latência da infecção pode ser visto como o estágio (mais recente) de crescimento da população viral, enquanto o período de infecção pode ser visto como um período que coincide com as fases de alta quantidade viral na primeira viremia do hospedeiro. Em seguida, quando ocorre a supressão da população de vírus para baixos níveis, a memória imunológica é criada pela abundância de células efetoras com vida longa (ou relativamente longa). O segundo ponto relevante a ser observado, envolve a previsão de que células efetoras, em alta quantidade, são incapazes de eliminar completamente o vírus do hospedeiro. Assim, o conceito de memória imunológica nesse modelo é (pelo menos em parte), consequência da persistência viral.

Uma vez que a resposta imune é ativada pela infecção viral, os componentes dessa resposta, tais como as células B e T, são capazes de intensa replicação. As experiências que envolvem transferência de linfócitos que são estimulados em hospedeiros irradiadores demonstram que, se dado uma oportunidade, um clone de células B continuará a se expandir indefinidamente. É evidente assim que alguma forma de mecanismo de feedback deve agir para regular o sistema imune. No modelo discutido anteriormente, o controle desse feedback estava relacionado a uma densidade de antígenos. Uma hipótese alternativa, proposta por Jerne [2], é que a regulação da resposta imune é conseguida mesmo na ausência de antígenos através da interação entre as células e alguns tipos de anticorpos dentro do hospedeiro. Em sua forma mais simples, tal regulação considera uma rede simétrica em que células do tipo 1 agem como antígeno a células do tipo 2, e vice-versa. Assim, 1 monta uma resposta imune a 2 e suprime sua abundância e 2 monta uma resposta imune a 1, que limita concomitantemente a abundância das células do tipo 1. Uma analogia conceitual útil deve considerar que 1 e 2 operam em uma interação do tipo predador-presa, em dois sentidos: em que 1 é predador da presa 2, e vice-versa.

Portanto, podemos pensar que uma população E_1 reage com o antígeno (de maneira similar ao modelo introduzido anteriormente), enquanto outra população E_2 reage com a presença de E_1 . Assim, para descrever a interação entre essas células efetoras na presença do antígeno numa concentração inicial $V(0)$, escrevemos as

seguintes equações diferenciais:

$$\frac{dE_1}{dt} = \Lambda_1 - \mu_1 E_1 + \varepsilon V E_1 + \frac{a_1 E_1 E_2}{1 + b_1 E_1 E_2}, \quad (3.5.37)$$

$$\frac{dE_2}{dt} = \Lambda_2 - \mu_2 E_2 + \frac{a_2 E_1 E_2}{1 + b_2 E_1 E_2}, \quad (3.5.38)$$

onde as células efetoras são recrutadas em uma taxa constante Λ_i da medula e morrem em uma taxa per capita μ_i . Os termos a_i e b_i são constantes arbitrárias que determinam o grau de proliferação devido a interação entre as células e os níveis no qual a taxa líquida de proliferação satura. Ainda, o antígeno é eliminado por uma reação com os anticorpos da seguinte forma:

$$\frac{dV}{dt} = rV - \sigma V (E_1 + E_2). \quad (3.5.39)$$

Apesar de deixarmos este estudo propositalmente incompleto, podemos verificar algumas propriedades interessantes desse modelo simples representado por essas equações. Nesse cenário o hospedeiro é infectado pelo patógeno viral que possui um antígeno que é idêntico ao expressado pela célula tipo 1. Uma vez que o vírus tenha se submetido à fase de replicação, o resultado do sistema será substancialmente mudado para a densidade da célula tipo 1 e, portanto, o antígeno tipo 1, se move do estado inicial para o estado imune, uma vez que o hospedeiro é ativamente imune. Este estado é estável e, portanto, a eliminação rápida do vírus deixará o sistema com elevados níveis de células efetoras. Em outras palavras, a memória imunológica é a essência do estado dinâmico criado pela rede regulatória via interação das células. As lições que devemos aprender com este exemplo são duas. Primeiro, a de gerar uma hipótese alternativa para explicar a existência da memória imunológica. Segundo, argumentar que a interpretação observada da resposta imunológica para agentes infecciosos pode ser mais facilmente investigada pela dinâmica de populações de célula-célula e célula-antígeno que interagem dentro do hospedeiro individualmente.

De qualquer maneira é bom frisar que uma teoria baseada num modelo matemático simples não resolve a controvérsia sobre a memória imunológica, mas ajuda a esclarecer as formas de comportamento dinâmico que podem levar a um sistema de rede regulatória.

3.6 O conceito de memória imunológica revisitado

Nesta seção vamos retomar o modelo baseado em autômato celular probabilístico desenvolvido na seção 2 do capítulo 2 tentando explorar até que ponto podemos simular com esse aparato metodológico a competição entre vírus e células do sistema imune, particularmente tentando explorar o conceito de memória imunológica dentro desse tipo de modelagem. Assim, primeiramente temos de adaptar o autômato original ao problema que estamos lidando. A nossa rede agora mimetiza um mundo microscópico um tecido humano virtual, onde ocorrerá a interação entre os

patógenos e células da resposta imune. Os estados que os sítios dessa rede vão poder assumir, são, um estado V ocupado com um patógeno, um estado E_1 ocupado com uma célula do sistema imune do tipo 1 e para podermos entrar na controvérsia gerada pelas duas formulações discutidas na seção anterior vamos considerar um estado E_2 , representando um sítio dessa rede sendo ocupado por um célula efetora diferente. Aqui temos de considerar também um estado 0, ou vazio que pode ser gerado devido a interação entre os elementos dessa competição.

O próximo passo é definir as regras que mimetizam as interações dos elementos presentes nessa rede assumindo que os fenômenos que ocorrem nesse tecido virtual são: mobilidade dos agentes, competição entre as espécies e replicação de cada um desses elementos. Um fato curioso e honestamente ainda não totalmente entendido é que essa regras têm de ser mantidas nessa ordem durante a simulação, de maneira, por exemplo, que uma célula viral só se reproduz se não for eliminada pelas células efetoras, o que nesse caso evita a superpopulação de patógenos precocemente.

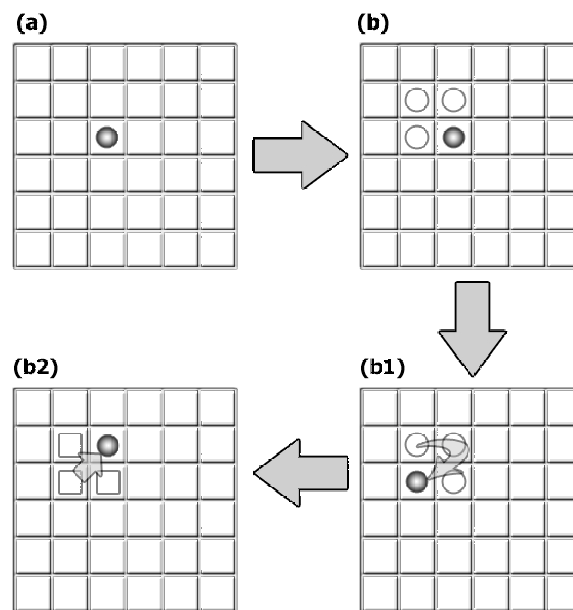


Figura 3.8: Algoritmo conhecido como vizinhança de Margolus. Passos consecutivos de tempo alternam a vizinhança de um determinado sítio.

Para simular a mobilidade das espécies interagentes, utilizamos um algoritmo conhecido como vizinhança de Margolus [39] e que tipicamente é utilizado para gerar um processo de difusão na rede. É importante observar que ao optarmos por esse procedimento abandonamos a interpretação de mobilidade atribuída a p_G no autômato construído no capítulo 2. Isso é interessante já que devemos assumir que os indivíduos tratados para este problema em particular (tipicamente células) têm um movimento na rede sem nenhuma correlação (ou intenção) enquanto que os

indivíduos modelados na versão macroscópica (pessoas) tem uma mobilidade mais estruturada (temos de preservar algum grau de intencionalidade no modelo).

A Figura 3.8, apresentada na página anterior, mostra esquematicamente o funcionamento desse algoritmo e seus passos podem ser resumidos da seguinte maneira: (1) dividimos a rede em blocos de 2×2 , de maneira que não ocorra sobreposição entre eles e fazendo com que nenhuma informação seja trocada entre blocos vizinhos, (2) esses blocos podem ser rotacionados de 90° a mesma probabilidade p_R (este mecanismo pode ser entendido indo da parte (b) para (b1) da figura, onde o ponto escuro foi para a posição da esquerda em relação onde ele estava inicialmente), (3) após a rotação, cada elemento dos blocos podem ser deslocados (transladados) na diagonal com probabilidade p_T , de maneira que agora possamos trocar informação entre os blocos vizinhos na figura (observamos isso indo de (b1) para (b2), onde podemos reparar que a grade já não esta mais no mesmo lugar que estava inicialmente, pois a disposição da rede mudou, ou seja, houve troca de informação entre blocos vizinhos).

Cada elemento tem sua própria probabilidade de rotacionar e transladar na diagonal, independente uma da outra, ou seja, se uma célula rotacionar, ela pode não necessariamente transladar. Para o sistema que estamos estudando ao aplicarmos este algoritmo, observamos que as células do sistema imune e a do patógeno vão se difundindo pela rede até estarem completamente misturadas, mimetizando o tipo de mobilidade esperada para esses elementos *in vivo*.

O mecanismo de competição entre os elementos da rede pode ser introduzido de maneira bem simples, todavia é importante observar que em primeira instância, as células do sistema imune irão disputar entre si e com o patógeno (que por sua vez também irão competir entre si) por um lugar na rede. Além dessa competição (por espaço), que aparece devido aos processos de difusão e reprodução como veremos, introduzimos um mecanismo adicional que nos diz como uma célula efetora elimina cada vírus, ao entrar em contato com ele (uma competição devido à interação entre eles, do tipo predador-presa). Assim, modelamos esse mecanismo através de uma probabilidade que mostra o grau de contato entre as células dos dois tipos de sistema imune e o vírus e a eficácia com que cada célula efetora (do tipo 1 ou 2) elimina o vírus. A equação a seguir

$$p_C = 1 - (1 - \lambda_1)^{n_1} (1 - \lambda_2)^{n_2},$$

portanto, define a probabilidade de que um vírus ocupando um sítio particular seja eliminado (o sítio, então passará para o estado vazio, nesse caso) devido à presença de n_1 células efetoras do tipo 1 e n_2 células efetoras do tipo 2 em sua vizinhança, onde λ_i é a probabilidade de uma célula do tipo i do sistema imune matar o vírus.

Particularmente, a replicação é dividida em duas fases: replicação do vírus e do sistema imune, cada um se replicando de maneira própria. A replicação em si consiste basicamente em fazer com que um dos elementos ocupando um sítio arbitrário (aleatoriamente) se multiplique, ou seja, gere filhos. A cada passo de tempo, então, esse elemento gera em média um número fixo de filhos distribuídos aleatoriamente pela rede nos sítios vazios existentes.

Em todos os casos que descrevemos a seguir, mantemos esse número fixo (escolhidos de uma distribuição uniforme). Todavia é possível alterar essa hipótese

de maneira a considerar cenários mais complexos de simulação onde, por exemplo, a reprodução das células do sistema imune dependa de uma probabilidade delas reconhecerem o vírus.

Assim, no caso da replicação do vírus, consideramos que cada um tem um tempo médio de vida t_v e que a cada passo de tempo ele gere n_v filhos. No caso do sistema imune, temos dois tipos de células para se reproduzir e entrar no sistema. Particularmente, as células do tipo i começam a entrar no sistema a partir de um instante particular t_i em que o vírus já mantinha uma população apreciável no sistema.

Respeitando esses tempos, permitimos que as células efetoras ocupem um espaço vazio com probabilidade p_i de acordo com a disponibilidade de espaços vazios em nosso tecido virtual. Ou seja, primeiramente percorremos toda a rede e ocupamos os sítios vazios com as células efetoras do tipo 1. Posteriormente, procuramos dentre os sítios vazios remanescentes e ocupamos com células efetoras do tipo 2. Observe-se que esse modo de ocupar a rede, introduz uma competição adicional por espaço (a cada passo de tempo).

Na Figura 3.9, mostramos um exemplo da simulação da evolução temporal da competição entre o vírus e dois tipos de células efetoras, a segunda ativada pela presença da primeira, efetuadas ao variar os parâmetros p_1 e p_2 , que são as probabilidades de ocupar os sítios vazios da rede com as células efetoras do tipo 1 e 2, respectivamente, e como foi discutido anteriormente, cria uma competição adicional entre elas e o patógeno por espaço, limitando o crescimento deste.

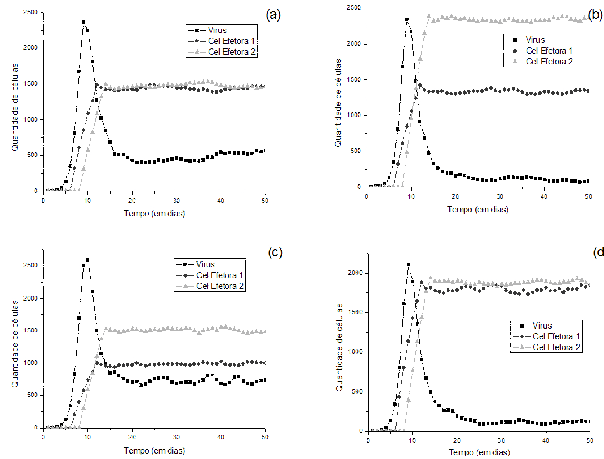


Figura 3.9: Quatro cenários de simulação até atingir o equilíbrio entre as espécies competindo em uma rede 100×100 . Aqui os valores de p_1 e p_2 são tomados como sendo, respectivamente: (a) 0,03 e 0,03, (b) 0,03 e 0,05, (c) 0,02 e 0,03, (d) 0,04 e 0,04.

De uma maneira geral, à medida que aumentamos p_1 o valor máximo da quantidade de células efetoras aumenta e o valor máximo da quantidade de patógenos

diminui. Após atingir esses máximos observamos um outro comportamento interessante ao variarmos esse parâmetro. Com efeito, nas partes (a), (b) e (c) da figura podemos observar um equilíbrio similar ao discutido no modelo baseado em equações diferenciais que foi interpretado como a memória imunológica da presença do antígeno. Esse equilíbrio é caracterizado pelas oscilações das quantidades de células efetoras e de patógenos sendo que o tamanho dessas oscilações podem ser modificados ao variar os tempos de vida relativos desses elementos interagentes e a probabilidade de uma célula efetora matar um vírus particular. Particularmente, a parte (d) da figura mostra que podemos manter a quantidade de células efetoras alta e a quantidade de vírus baixa, eventualmente eliminando-o. Esse fato é importante já que no modelo baseado em equações diferenciais as oscilações amortecem até os pontos de equilíbrio e um segundo pico de viremia (observado nesses modelos) é um tanto artificial já que não dependia de uma re-infecção adquirida do meio externo, mas de uma diminuição das células efetoras. Outro fato importante de destacar desses resultados é que com esse modelo alternativo é possível garantir níveis de células efetoras mesmo depois do desaparecimento do vírus, já que a densidade de cada uma das células depende da outra, mimetizando a hipótese de rede regulada discutida nos modelos baseados em equações diferenciais.

Capítulo 4

Velhos Problemas e Novas Fronteiras

Apesar dos benefícios dos antibióticos no tratamento de uma vasta variedade de patógenos, não há dúvida que seu uso indiscriminado tem levado à emergência de novas linhagens resistentes desses patógenos. Hospitais e outras comunidades locais são especialmente férteis no espalhamento desses novos tipos. Em particular segundo a Organização Mundial de Saúde [43], 5% da população mundial com tuberculose, tem o patógeno resistente a qualquer antibiótico. Desse modo, fica claro que considerar uma estratégia compreensível para o uso de antibióticos é essencial. Tais estratégias por sua vez, devem ser fundamentadas com base teórica quantitativa, derivada de um firme entendimento de como o uso de antibióticos por uma comunidade, traduz-se na emergência de linhagens resistentes. Com uma tal teoria disponível, pode-se administrar uma estratégia que possa balancear, custos e benefícios para a saúde, de modo a permitir a disponibilidade de antibióticos àqueles que mais necessitam, sem cortar o efeito da droga por um longo período. Infelizmente, uma tal teoria ainda não existe, mas alguns modelos epidemiológicos tem sido propostos ultimamente que fundem elementos de evolução das bactérias e dinâmica de transmissão para examinar a emergência de resistência a antibióticos. Neste capítulo vamos tentar apurar a idéia original que vínhamos discutindo durante toda esta monografia: a de tentar buscar um modelo epidemiológico que leve em consideração fatores mais complexos como a evolução do patógeno invasor. Isso foi feito de maneira declaradamente desconexa até aqui já que na maioria das vezes estávamos preocupados em estabelecer o método e entender o fenômeno. Com esse problema sobre a resistência a antibióticos, vamos tentar levantar possibilidades neste sentido de mostrar algum tipo de modelagem e dar sugestões sobre os próximos passos a serem dados.

4.1 Modelos de controle farmacológico

Particularmente, um dos primeiros trabalhos nessa direção foi desenvolvido por Massad e colaboradores [42], que propuseram um modelo epidemiológico para hospitais que captura, de maneira bastante completa, as condições de coexistência de bactérias sensíveis e resistentes a droga, apesar de eles não explicitarem a ligação entre resistência e o consumo de drogas. Mais recentemente, outros modelos simples utilizando genética de populações de resistência a antibióticos tem sido elaborados [44, 45], considerando a pressão seletiva imposta pelo tratamento (prescrito em uma taxa *per capita* definida) sobre uma bactéria comensal. Suas análises mostram que existe uma relação entre consumo de antibióticos e a frequência de resistência, mesmo quando a resistência tem uma desvantagem seletiva para níveis específicos de tratamento. O modelo que vamos discutir na seqüência reúne esses ingredientes e é uma variante do modelo proposto por Austin e colaboradores [45], onde distinguimos cada hospedeiro de acordo com seu *status* de doença e se eles estão sujeitos ao tratamento com antibióticos. Assim, descrevemos a dinâmica de transmissão de bactérias entre hospedeiros, considerando que dentro de cada hospedeiro essas bactérias, resistentes (r) e sensíveis (s) a droga, crescem competindo entre si segundo um modelo clássico de genética de populações. Dessa maneira, mostramos como a introdução de antibióticos em uma população altera a dinâmica de transmissão, ao reduzir o *fitness* das bactérias sensíveis frente às resistentes e, por consequência, produzindo novos hospedeiros para as bactérias resistentes.

O modelo é desenhado para deprever a dinâmica de transmissão de qualquer das muitas espécies de bactérias que comumente residem de maneira comensal no corpo humano (*staphylococcus*, *enterococcus*, *escherichia coli*, etc.). Estas bactérias podem ser transmitidas entre pacientes em hospitais, por contato direto, através da contaminação de todo o ambiente, ou por vetores humanos como os enfermeiros que cuidam dos pacientes. As bactérias são continuamente confrontadas com antibióticos usados para a profilaxia ou tratamento de alguma doença. Além de acarretarem um risco de infecções oportunistas, essas espécies de bactérias servem como reservatórios de plasmídios resistentes a antibióticos, que são horizontalmente transmitidos entre linhagens e espécies de bactérias e são os verdadeiros responsáveis por resistência múltipla à drogas [46]. Tanto em hospitais como em outras comunidades, o padrão de emergência da resistência tende a ser similar; tipicamente, um período longo de baixos níveis de resistência precede uma fase de aumento rápido na frequência, seguido de uma lenta aproximação a frequência de equilíbrio ($< 100\%$). Esse padrão sigmoidal, das mudanças observadas na frequência da resistência freqüentemente tem sido atribuído ao aumento do volume de consumo de drogas no período de vida de um indivíduo. De fato, podemos obter esse padrão epidemiológico utilizando a teoria clássica de genética de populações sob uma pressão seletiva constante (= consumo de drogas), considerando que dentro de um hospedeiro particular podem existir bactérias com dois genótipos de interesse: s (sensível) e r (resistente). Na ausência de uso de antibióticos o genótipo da bactéria

sensível tipicamente tem uma vantagem seletiva,

$$w_s = 1 > w_r. \quad (4.1.1)$$

Assim, as freqüências das bactérias resistentes (p_t) e sensível ($q_t = 1 - p_t$) depois de t gerações sucessivas é determinado por

$$p_{t+1} = \frac{w_r p_t}{w_r p_t + w_s q_t}. \quad (4.1.2)$$

A pressão seletiva imposta pelo uso de antibiótico é medida em termos de *doses diárias definidas (DDDs)/1000* indivíduos da comunidade em estudo, ou equivalentemente, a proporção da comunidade recebendo tratamento a ($a \simeq$ taxa de prescrição \times duração do tratamento). Assim, o uso de antibiótico reduz o fitness da bactéria sensível de 1 para $1 - a$, deixando o fitness das resistentes constante (assumimos resistência completa). Portanto, a denota a probabilidade de que a seleção esteja presente.

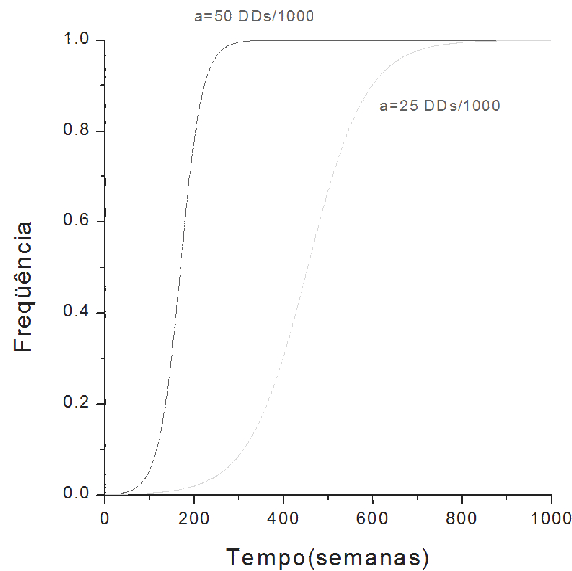


Figura 4.1: Padrão sigmoidal da variação da freqüência de bactérias resistentes p_t sob pressão seletiva imposta pelo consumo de drogas mantida constante. A freqüência inicial de bactérias resistentes foi tomada como $p_0 = 10^{-3}$ e o fitness relativo das bactérias resistentes permanece constante $w_r = 0.99$.

A Figura 4.1 mostra as simulações para a evolução temporal da freqüência da resistência para dois regimes de pressão seletiva devido ao consumo de antibióticos, medida em *doses diárias definidas (DDDs)/1000* indivíduos. Ali, consideramos que o uso de antibióticos reduz a vantagem seletiva da bactéria sensível ($w_s = 1 - a$)

independente da frequência das bactérias. Dessa maneira, se p_0 e q_0 forem as respectivas frequências iniciais, é fácil mostrar que

$$\frac{p_t}{q_t} = \frac{p_0}{q_0} \cdot \left(\frac{w_r}{1-a} \right)^t, \quad (4.1.3)$$

ou seja, nesse regime de consumo constante de antibiótico, uma vez que a ultrapassa o limiar $a > 1 - w_r$, a emergência de resistência tem o padrão sigmoidal observado, tendendo a fixação quando $t \rightarrow \infty$. O mesmo tipo de padrão é obtido quando a prescrição de antibiótico é dependente da frequência das bactérias, ou seja, se modelarmos que o consumo é reduzido a medida que a frequência de bactérias resistentes aumenta, uma frequência de equilíbrio intermediária ($< 100\%$) é obtida. Assim, se a diminuição de consumo de antibiótico cai de uma forma não-linear como $w_s = 1 - aq_t^n$ (onde n é uma constante), um novo ponto fixo é alcançado quando, $w_s(\hat{q}) = w_r$ ou, equivalentemente

$$\hat{p} = 1 - \left(\frac{(1 - w_r)}{a} \right)^{\frac{1}{n}}. \quad (4.1.4)$$

Para descrever a dinâmica de transmissão de um patógeno dentro da comunidade de hospedeiros, nós classificamos os indivíduos na comunidade (ou pacientes em um hospital) de maneira similar ao que foi concebido por Austin e colaboradores [45]. Daí, dada uma população com N indivíduos, em qualquer instante de tempo ela estará dividida nas seguintes classes: susceptíveis a colonização ou não colonizados ($S(t)$), infectado pela bactéria sensível ao antibiótico ($I(t)$) e colonizado pela bactéria resistente ($R(t)$). Consideramos também que uma fração da população a esta sendo tratada com antibiótico independente da classe a que pertence, ou seja, independente da colonização. O tratamento em geral, acarreta a eliminação das bactérias sensíveis na maioria dos hospedeiros, que passam para a classe de indivíduos tratados e não colonizados ($T(t)$). Todavia, esse mesmo tratamento pode selecionar uma fração pequena dos hospedeiros para serem resistentes, com uma probabilidade σ , ou por mutação ou pela expansão de cepas resistentes pré-existentes no hospedeiro (resistência adquirida). Como os organismo resistentes não são afetados pelo tratamento (resistência completa) os hospedeiros tratados nesse caso permanecem susceptíveis de colonização por bactérias resistentes ($TR(t)$). Uma simplificação adotada aqui é a de considerar que a população total de hospedeiros N é mantida constante. Esta é uma hipótese realística quando nós queremos modelar os efeitos da colonização dentro de hospitais onde o número de vagas, camas e número de pessoas disponíveis para o atendimento, são fatores limitantes. Essa formulação *compartimental* incorpora a hipótese clássica de uma *mistura homogênea* de hospedeiros e os pacientes no hospital tornam-se colonizados devido a um modo global de transmissão: por via aérea, por exemplo, ou por causa da mobilidade de enfermeiras. Contudo, como colocado por Massad e colaboradores [42], pacientes em um hospital estão confinados a suas camas e não podem ser considerados como entidades livres para movimentar-se, consideração essa inerente à *lei de ação das massas* por detrás da hipótese de mistura homogênea. Para fazer um modelo

mais detalhado de transmissão, distribuímos os N indivíduos da população em uma rede bidimensional quadrada de comprimento linear L ($N = L \times L$), de maneira que, cada indivíduo da população ocupa um sítio dessa rede (consideramos que toda rede é ocupada). Assim, um indivíduo na posição (i, j) , ($i, j = 1, \dots, L$) pode pertencer a uma das classes designadas acima de acordo com seu *status* de colonização ou de tratamento. Além disso, consideramos que cada indivíduo em um dado instante de tempo pode estar sendo colonizado por $m_s(i, j)$ bactérias susceptíveis e $m_r(i, j)$ bactérias resistentes. Dessa maneira, se $m_s(i, j) + m_r(i, j) = 0$, o indivíduo pertence a classe S ou T , se $m_s(i, j) = 0$ e $m_r(i, j) \neq 0$, o indivíduo pertence a classe R ou TR e se $m_s(i, j) \neq 0$ e $m_r(i, j) = 0$, o indivíduo necessariamente pertence a classe I . Outra diferença em relação aos modelos existentes é introduzida ao considerarmos o caso de coexistência, em que um hospedeiro particular pode ser colonizado por bactérias susceptíveis e resistentes ($m_s(i, j) \neq 0$ e $m_r(i, j) \neq 0$). Obviamente, já que queremos manter a divisão da população em classes de hospedeiros como feito em [45], adotamos o critério de que um paciente particular pertence a classe R , se $m_r(i, j) > K$ e pertencente a classe I , caso contrário. A evolução temporal da população em cada passo de tempo é feita através da atualização da rede sincronicamente segundo as seguintes regras: (1) A colonização por bactérias espalha-se aos primeiros vizinhos susceptíveis de um indivíduo infectado com uma probabilidade p_L dependendo do número de vizinhos infectados que cada susceptível tem e do número de bactérias colonizando cada um desses vizinhos. Portanto, estamos considerando aqui somente transmissão horizontal e nós iremos nos referir a esse tipo de transmissão como *transmissão direta*.

(2) No caso de um indivíduo já estar colonizado, ele continuará nesse estado sem alterar a quantidade de bactérias que o colonizam. (3) Os sítios com indivíduos colonizados podem tornar-se susceptíveis devido a uma taxa natural de eliminação de bactérias f , correspondendo a um período de colonização de $1/f$. Além disso, qualquer hospedeiro permanece na comunidade por um período de $1/\mu$, onde μ é a taxa de mortalidade ou de remoção do hospedeiro. (4) Estudamos também a possibilidade de um indivíduo susceptível torna-se infectado com uma probabilidade global p_G . Nós nos referimos a esse fenômeno como *transmissão indireta*. As regras do modelo podem, portanto, ser explicadas com segue. Com relação a regra 1, um indivíduo susceptível pode ser colonizado devido a seus n_I vizinhos infectados (com $0 \leq n_I \leq 8$) e uma probabilidade p_L para a transmissão da doença (estimada como uma função do número de vizinhos infectados). Dessa maneira, se o indivíduo tornar-se infectado ele ficará colonizado, por migração dos *reservatórios* em seus vizinhos, com uma quantidade \overline{m}_s de bactérias sensíveis e \overline{m}_r de bactérias resistentes, que são as médias sobre todos os vizinhos colonizados, do número de bactérias de cada tipo. A segunda regra pode ser vista como uma restrição da primeira no caso em que o indivíduo já está colonizado e estabelece que uma vez que isso aconteça, o indivíduo não pode receber bactérias de seu vizinho (por migração). Essa restrição pode ser relaxada criando-se uma hierarquia de transmissão entre hospedeiros, dependendo da quantidade de bactérias que eles contém e dessa maneira a segunda regra seria um caso particular da primeira. A interpretação da terceira regra esta ligada aos resultados da doença. Qualquer um dos indivíduos da

população tem uma probabilidade de morte natural μ (fixada constante), ou ainda, podemos pensá-la como uma taxa média de entrada e saída em um hospital. Somado a isso, a colonização pode ser letal ou é eliminada em um indivíduo infectado com uma probabilidade f o que significa que novos indivíduos podem entrar na população para ocupar seu lugar. Se esse não é o caso, a terceira regra estabelece que o sítio contém um indivíduo que, por causa de morte natural ou porque eliminou a bactéria, não participa no espalhamento da doença. Finalmente, a quarta regra foi introduzida, para considerar transmissões por via aérea, comida, contato com os enfermeiros de uma maneira geral, ou algum outro vetor secundário. A transmissão indireta também é adotada no modelo para incorporar o modo global de transmissão que depende da presença de um ou mais pacientes colonizados na população aumentando a probabilidade p_G de qualquer um dos pacientes ser colonizado.

Na Figura 4.2 (a) apresentamos a evolução temporal da população durante dez anos, anotando a cada semana a proporção de indivíduos de acordo com o seu *status* de colonização: os indivíduos susceptíveis ($S(t) + T(t)$), os indivíduos que só contém bactérias sensíveis ao tratamento ($I(t)$) e a proporção de indivíduos com a bactéria resistente ($R(t) + TR(t)$). Durante esse período, 10% da população permanece em um regime de consumo de antibiótico constante. Quando a resistência ao antibiótico é introduzida em uma comunidade (ou desenvolve resistência através de mutação ou transferência de plasmídeo) ela irá se estabelecer ou diminuir dependendo do fitness relativo da bactéria resistente (w_r) e da taxa de prescrição de antibióticos (γ). Essa taxa de prescrição é estimada a partir de proporção de pacientes recebendo tratamento a e o tempo médio $1/g$ que o tratamento é prescrito, de maneira que $\gamma \simeq ag$. Assim, a parte (b) da figura mostra a proporção endêmica de hospedeiros colonizados com a bactéria resistente, definida após a população alcançar o equilíbrio como:

$$p_R = \frac{\text{Número de indivíduos tratados e resistentes}}{\text{Número total de indivíduos tratados}} = \frac{TR(\infty)}{T(\infty) + TR(\infty)}, \quad (4.1.5)$$

em função da taxa de prescrição de antibióticos. Obviamente, os gráficos da figura foram ajustados de maneira a não mostrar as flutuações inerentes da natureza estocástica das simulações. Assim, os gráficos mostram apenas os valores médios das grandezas medidas, para que possamos comparar esses resultados com a formulação determinística [45]. Por simplicidade ignoramos também mutações e transferência de plasmídeos, e observamos o limiar de taxa de prescrição acima do qual a resistência se estabelece. Aumentando-se o custo da transmissão deslocamos a curva para taxas de prescrição mais altas, como também a transição observada fica mais rápida. Além disso, usando a proporção endêmica de indivíduos carregando a bactéria resistente como uma medida, nós podemos observar o efeito de usar diferentes tratamentos. Por exemplo, comparando o efeito sobre a resistência de um tratamento curto e intensivo administrado na maioria das pessoas infectadas, comparado com um tratamento longo prescrito para poucas pessoas.

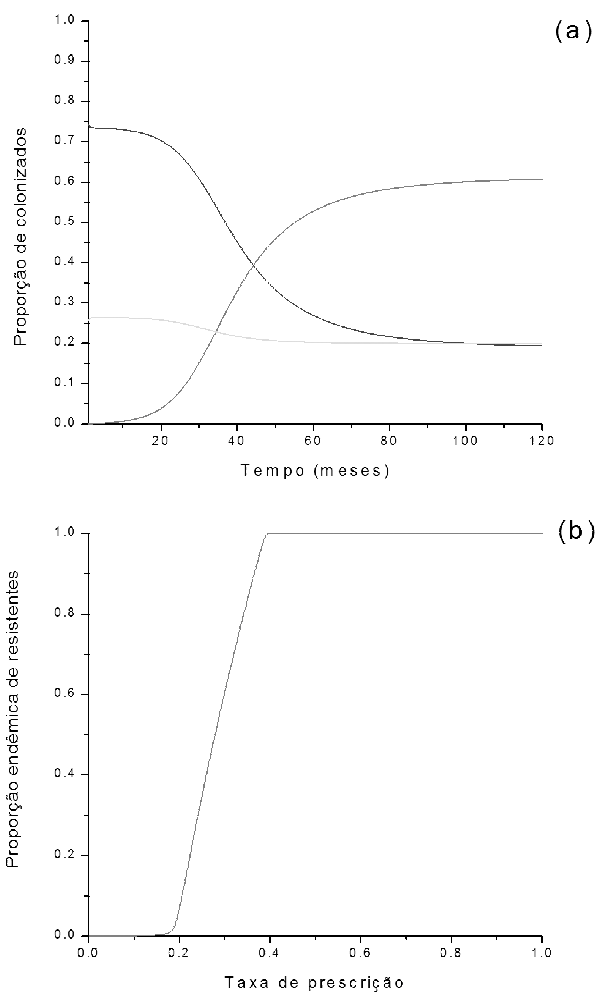


Figura 4.2: (a) Emergência de resistência a antibiótico em uma comunidade em dez anos de consumo de antibiótico. A prevalência global permanece aproximadamente constante enquanto que resistência aumenta até aproximadamente 72%. Antibióticos são consumidos em uma taxa de uma vez a cada três meses. (b) Proporção endêmica colonizada com bactérias resistentes a antibiótico em função da taxa de prescrição com um custo de transmissão de 50%. Aumentando-se a taxa de prescrição acima do limiar 0.2 por mês a resistência alcança a fixação.

Apesar dos resultados apresentados nesta seção, obviamente este tópico ainda tem de ser estudado com mais detalhes e realmente este é um problema em aberto. Particularmente, a política de prescrição considerada no modelo é a de manter uma proporção constante de pacientes sob tratamento com antibiótico. Essa hipótese serviu para calcular um nível de tratamento limiar, mas certamente não é uma

descrição realística do tipo de política que pode ser adotada por um médico durante o *tempo de vida* de um antibiótico. Mesmo assim, um raciocínio simples aqui ainda pode nos levar a conclusões sobre outros tipos de terapias. Por exemplo, com uma vantagem seletiva para resistência de 50% por geração, uma frequência inicial de bactérias resistentes de 10^{-2} ou 10^{-8} , não são números muito diferentes. No primeiro caso, levaria aproximadamente 9 gerações para que a frequência da resistência alcançasse 0.5, enquanto que no segundo caso levaria aproximadamente 37 gerações para alcançar o mesmo nível. Esta é no mínimo parte da razão para *falha* dos programas terapêuticos que alternam antibióticos ciclicamente em hospitais, ou seja, a não ser que as bactérias sejam totalmente eliminadas, quando o mesmo antibiótico é usado outra vez, a resistência irá aparecer rapidamente.

4.2 Como fazer uma modelagem afinal?

Epidemiologia estuda a dinâmica populacional de uma doença, contudo a infecção *per se* não é equivalente à doença. A infecção é a presença de um agente infeccioso no hospedeiro a doença é um desvio do *estado* saúde. Então a doença é um possível mas não necessário resultado da infecção. O estado de imunidade de um indivíduo também é um fator importante na progressão de uma doença. Com a infecção da malária, por exemplo, os indivíduos com imunidade clínica podem ter a infecção da malária, mas não ter nenhuma doença ou patologia, em contraste, aquele sem essa imunidade irá ter a doença e a patologia para níveis similares de infecção. Portanto, um modelo epidemiológico não deve seguir somente o progresso de uma infecção, mas também seguir as possíveis consequências da infecção chamadas doença. Até certo ponto o impacto da infecção é considerado nos modelos epidemiológicos que monitoram o estado imunológico de um hospedeiro ou aqueles que consideram diferentes tipos de cepas de um patógeno. Na prática, a doença associada com a infecção de um indivíduo será um resultado complexo de vários fatores incluindo a evolução do hospedeiro e seus parasitas e a história da exposição do hospedeiro a infecção. Não existe nenhuma relação simples ou linear entre esses fatores e a doença: doença é, portanto, melhor vista como uma propriedade emergente da infecção e pertence à mesma classe de problema. Como vimos, aproximações computacionais e matemáticas; imaginativas e eficientes são essenciais para lidar com essas complexas relações entre fatores que contribuem para uma possível doença associada à infecção. Essas aproximações, obviamente devem manter um compromisso com os dados reais, no sentido de que para que esses modelos, e métodos de simulação possam ser utilizados de forma preditiva para auxiliar e avaliar uma tomada de decisão quando da disseminação de uma epidemia, eles devem ser cuidadosamente parametrizados e validados com dados epidemiológicos coletados. Entretanto, em todo este curso nos ocupamos em estabelecer varias maneiras de modelar o mesmo problema no sentido de paulatinamente entender como podemos introduzir complexidades nesse modelo. Cada passo de modelagem discutida aqui então tinha o intuito de mostrar, até um certo ponto, como que cada hipótese adicional diferenciadas entre si pelo modo de modelar (ou mais propriamente, pela técnica disponível

ao pesquisador no momento da interação com o problema real) tem um efeito crucial sobre as conclusões. Com efeito, todo o curso pode ser visto como uma justificativa para o *pluralismo* no modo pelo qual modelamos no sentido de que o fenômeno em estudo pode ser entendido de diferentes modos e uma comparação desses *modos* de modelar pode revelar mais sobre o fenômeno do que um dos métodos em particular. Ainda no âmbito geral de modelagem, podemos entrever que qualquer intenção de formalizar um problema biológico como os de transmissão de uma infecção ou mesmo evolucionário, matematicamente, obriga-nos a fazer hipóteses, implícita ou explicitamente, sobre quais são as *unidades que devemos modelar*.

Daí a resposta à pergunta desta seção: não existe um modo de modelar um fenômeno biológico, já que o que queremos é entendê-lo. Para uma compreensão honesta desses fenômenos devemos deixar de lado a ambição especialista e humildemente delatar nossa ignorância ao investigar o pluralismo dos modos de modelar.

Finalmente, parece que tudo o que nos restou é admitir o erro. Sim, porque modelar é fazer uma caricatura do processo real em estudo, que no caso do processo epidêmico e evolucionário optamos por criar uma realidade virtual em que são simulados os elementos que se supõem implicados na gênese e difusão desses processos. Como consolo nos resta o diálogo entre a Rainha de copas e Alice, criado por Lewis Carrol, em *Através do Espelho*, onde podemos finalmente definir em que consiste a metodologia discutida neste curso:

“Eu não posso acreditar nisso!”, disse Alice.

“Não pode?”, disse a rainha com pena. “Tente de novo: respire fundo e feche os seus olhos”.

Alice riu. “Não tem qualquer sentido tentar”, ela disse: “não se pode acreditar em coisas impossíveis”.

“Eu ousou dizer que você não tem muita prática”, disse a rainha.

“Quando eu era da sua idade, sempre praticava durante meia hora por dia. Algumas vezes, cheguei a acreditar em seis coisas impossíveis antes do café da manhã.”

Bibliografia

- [1] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and its Applications* (Griffin, London, 1975).
- [2] R. M. Anderson e R. M. May, *Infectious Diseases of Humans* (Oxford University Press, Oxford, 1991).
- [3] N. C. Grassly et al, Population dynamics of HIV-1 inferred from gene sequences, *preprint* (1998).
- [4] J. M. Coffin, *Current Topics in Microbiology*, **176**, 143 (1992).
- [5] M. A. Novak, R. M. Anderson and R. M. May, The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease, *AIDS*, **4**, 1095-1103 (1990).
- [6] D. Mollison, *Epidemic models: their structure and relation to data*, Cambridge University Press (1995).
- [7] W. O. Kermack and A. G. MacKendrick *A contribution to the mathematical theory of epidemics*, *Proc. R. Soc. Lond. A*, **115** (1927) 700.
- [8] C. T. Codeço, *Modelagem Computacional de Processos Infecciosos*, Escola de Verão em Métodos Computacionais em Biologia, Petrópolis: LNCC, 2000.
- [9] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Chem. Phys.* **81**, 2340–2361 (1977).
- [10] D. T. Gillespie, A general method for numerically simulating the stochastic evolution of coupled, *J. Comp. Phys.* **22**, 403–434 (1976).
- [11] B. T. Grenfell and A. Dobson, *Ecology of infectious diseases in natural populations*, Cambridge University Press (1995).
- [12] J. E. Keymer, P. A. Marquet, J. X. Velasco-Hernández and S. A. Levin, Extinction thresholds and metapopulation persistence in dynamics landscapes, *Am. Natur.* **156**, 478–494 (2000).

- [13] N. Boccara, K. Cheong, Automata Network SIR Models for Spread of Infectious Diseases in Populations of Moving Individuals, *J. Phys. A: Math and Gen.*, v.25, p.2447-2461, 2000.
- [14] C. J. Rhodes, R. M. Anderson, Power laws governing epidemics in isolated populations, *Nature*, **381**, 600–602 (1996).
- [15] D. J. Watts, S. H. Strogatz, Collective dynamics of small-world networks, *Nature*, **393**, 440-442, 1998.
- [16] D. Alves, Análise estatística da teoria de quasi-espécies de evolução molecular *Tese de Doutorado*, São Carlos, IFSC - USP (1999).
- [17] D. R. Mills, F. R. Kramer, S. Spiegelman, Complete nucleotide sequence of a replicating RNA molecule, *Science*, **180**, 916 (1973).
- [18] D. Alves, J. F. Fontanari, Error threshold in finite populations, *Phys. Rev. E*, **57**, 7008–7013, june, (1998).
- [19] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
- [20] D. Bonnaz, A. J. Koch. *Stochastic model of evolving populations*, *J. Phys. A: Math. Gen.*, **31**, 417 (1998).
- [21] G. Woodcock, P. G. Higgs. *Population evolution on a mutiplicative single-peak fitness landscape*, *J. Theor. Biol.*, **179**, 61 (1996).
- [22] D. Alves, V. J. Haas, A. Caliri, The Predictive Power of R_0 in an Epidemic Probabilistic System, *J. Biol. Phys.* **29**, 63–75 (2003).
- [23] H. M. Yang, Epidemiologia da transmissão da dengue, em *TEMA - Seleta do XXV CNMAC*, (A. Sri Ranga, ed.), 2003.
- [24] C. P. Ferreira, H. M. Yang, Estudo da transmissão da dengrue entre indivíduos em interação com a população de mosquitos *Aedes aegypti*, em *TEMA - Seleta do XXV CNMAC*, (A. Sri Ranga, ed.), 2003.
- [25] H. F. Gagliardi, F. A. B. Silva, D. Alves. Automata network simulator applied to the epidemiology of urban Dengue fever. Lectures Notes in Computational Science. V.N. Alexandrov et al. (Eds.): ICCS 2006 Part III, LNCS 3993, pp. 297–304. Springer-Verlag Berlin Heidelberg 2006.
- [26] M. Eigen. *Selforganization of matter and the evolution of biological macromolecules*, *Naturwissenschaften* **58**, 465 (1971).
- [27] M. Eigen, J. McCaskill, P. Schuster. *The molecular quasi-species*, *Adv. Chem. Phys.*, **75**, 149 (1989).
- [28] M. Eigen, *Viral quasispecies*, *Sci. Am.*, 269, 42–49 (1993).

- [29] M. A. Novak, R. M. Anderson, Mclean, T. F. W. Wolfs, J. Goudsmit and R. M. May, Antigenic diversity threshold and the development of AIDS, *Science*, **254**, 963-969.
- [30] C. J. Thompson, J. L. McBride. *On the Eigen's theory of the selforganization of molecules and the evolution of biological macromolecules*, *Math. Biosci.*, **21**, 127 (1974).
- [31] B. L. Jones, R. H. Enns, S. S. Rangnekar. *On the theory of selection of coupled macromolecular systems*, *Bull. Math. Biol.*, **38**, 12 (1976).
- [32] S. Wright. *Evolution and the genetics of populations* Vol. I (The University of Chicago Press, Chicago, 1968).
- [33] S. Galluccio. *Exact solution of the quasispecies model in a sharply peaked fitness landscape*, *Phys. Rev. E*, **56**, 4526 (1997).
- [34] D. Alves, J. F. Fontanari, Populational genetics approach to the quasispecies model. *Phys. Rev. E*, **54**, number 4, 4048-4053 (1996).
- [35] J. Swentina, P. Schuster. *Self-replication with errors - a model for polynucleotide replication*, *Biophysical Chemistry*, **16**, 329 (1982).
- [36] J. M. Smith. *Models of Evolution*, *Proc. R. Soc. Lond. B*, **219**, 315, (1983).
- [37] M. Eigen, P. Schuster. *The Hypercycle - A Principle of Natural Self-Organization*, Springer-Verlag, Berlin, (1979).
- [38] M. A. Nowak, R. M. May, Mathematical biology of HIV infections: Antigenic variation and diversity threshold, *Math. Biosci.* **106**, 1-21 (1991).
- [39] C. P. Ferreira, Dinâmica de replicação na rede: aplicações em modelos de evolução pré-biótica e de formação úlceras, 2001, 118 p., Tese de Doutorado em Física Básica, IFSC – Universidade de São Paulo.
- [40] G. Abramson, M. Kuperman, Small world effect in an epidemiological model, *Preprint arXiv.nlin.AO/0010012* (2000).
- [41] M. Nowak, P. Schuster. *Error thresholds of replication in finite populations mutation frequencies and the onset of Muller's ratchet*, *J. Theor. Biol.* **137**, 375 (1989).
- [42] E. Massad, S. Lundberg and H. M. Yang, The evolution of resistance to antibiotics, *Int. J. Biomed. Comp.*, **33**, 65–81 (1993).
- [43] Reports, *World Health Organization*, <http://www.who.int/>, (1998).
- [44] B. R. Levin, et al, The population genetics of antibiotic resistance, *Clin. Infect. Dis.*, **24**, S9 (1997).

- [45] D. J. Austin, K. G. Kristinsson and R. M. Anderson, The transmission dynamics of antibiotic-resistant bacteria: the relationship between resistance in commensal organisms and antibiotic consumption, *Proc. R. Soc. Lond. B*, **264**, 1629 (1997).
- [46] J. E. MacGowan, Antimicrobial resistance in hospital organisms and its relation to antibiotic use, *Rev. Infect. Dis.*, **5**, 1033–1048 (1983).

Índice

- AIDS, 12, 57, 60
- antibiótico, 73, 76, 79
 - resistência ao, 78
- antibióticos, 74, 75
 - prescrição de, 78
 - resistência a, 73, 74
- aproximação, 11
- arbovírus, 37
- autômato, 34, 35, 38, 39, 41, 42
- autômatos, 20–22
- autômatos celulares, 20

- back mutations, 51
- bactéria, 27
 - resistente, 76
 - sensível, 76
- bactérias, 73–75, 77
 - resistentes, 79
- bacteriófago, 27

- célula efatora, 70
- células efetoras, 66
- células T, 60
- catástrofe de erro, 49
- cinética química, 45
- colonização, 76, 77
- colonizado, 76
- colonizados, 76, 77
- compartimental, 13, 16, 17, 20–23, 76
- compartimentalização, 37
- complexos, 11
- controle farmacológico, 73

- Dengue, 12, 16, 37, 39–41
- desvantagem seletiva, 74
- dinâmica de populações, 12

- dinâmica de transmissão, 73, 76
- distribuição
 - de mutantes, 49
 - largura, 49
- diversidade
 - limiar de, 63
- diversidade de antígenos, 57, 63
- diversidade viral, 61
- doença, 12, 13, 15, 16, 19, 21, 24, 26, 27, 53, 62, 63, 74, 77, 80
- droga, 74
- drogas, 74

- enterococcus, 74
- enzima, 52, 59
- epidemia, 12, 34, 37, 39–41, 43, 66, 80
- equilíbrio, 49, 66, 74
- escherichia coli, 74
- espaço de genomas, 49, 54
- espacial, 34
- estocástico
 - escape, 56
 - processo, 62
- estocático, 34
 - modelo, 63
- evolução, 45, 47, 49, 55, 56, 63, 64, 70, 75, 78
 - das bactérias, 73
 - viral, 53, 63
- evolucionários, 13, 27, 30, 31
- expostos, 16, 37, 38

- fitness, 28, 49, 74, 75, 78

- genética de populações, 74
- genótipo, 74

- genoma, 29, 46, 48, 49, 54, 59
- HIV, 12, 57, 60, 62
- hospedeiro, 12, 27, 32, 33, 38, 66, 74, 76, 77
- hospedeiros, 76
- humanos, 37–41
- imunes, 61
- in vivo, 29
- infecção, 13, 18, 22, 23, 26, 27, 29, 80
- infecções, 63
- infectados, 13–19, 21, 22, 24, 34, 35, 37–39, 41, 42, 59, 77
- irrecuperados, 37, 41
- limiar de diversidade, 63
- limiar de erro, 50–52
- modelagem, 11–13, 16–18, 23–25, 27, 33, 34, 39, 41, 43, 57, 73, 80
- Modelagem de Sistemas Biológicos
via da simulação, 11
via experimental, 11
- Modelo
SIR, 13, 18
- modelo, 12–16, 18, 20–27, 29, 31–37, 39–43, 45, 57, 61, 62, 73, 74, 76, 77, 79, 80
estocástico, 63
SEIR, 16, 33
SIR, 13, 16, 18, 20, 21, 26, 33, 35
SIRS, 33
- mosquitos, 37–41
- mundo pequeno, 25–27, 35, 41–43
- mutação, 31, 45, 46, 49, 50, 54, 55, 62–64, 78
matriz, 64
taxa, 49
- número de reprodução básica, 15
- nucleotídeo, 57
- otimização evolucionária, 49
- patógeno, 45, 73, 76
- patógenos, 12, 27, 29, 45
- plasmídeo, 78
- política de prescrição, 79
- pre-bioótica, 45
- predador-presa, 69
- preditivo
modelos, 11
- pressão seletiva, 74, 75
- probabilidade, 17–19, 21–23, 25, 26, 29, 30, 34, 35, 37, 39–41, 49, 51, 64, 69–71, 75–77
- protocolo de re-conexão, 25
- quase-espécie, 48–54
modelo de, 57
teoria de, 56
viral, 57
- quase-espécies, 45, 49
teoria, 45
virais, 53
- reator de fluxo, 57
- rede, 76, 77
- redes, 38, 42
- relevo de replicação, 50
- replicação, 57, 64, 69
erro de, 49
taxa, 50
- replicase, 28, 31
- reprodução, 15, 16, 20, 23, 28, 30, 31, 55
taxa, 63
- resistência, 13, 73–75, 78, 79
- resistente, 74
- resistentes, 76
- resistência
completa, 75
- RNA, 28, 29
- SEIR, 16, 33, 38
- SEIRS, 41
- seleção, 45, 47–51, 75
- sensível, 74
- seqüência mestra, 49
- simulação, 11, 62
- SIR, 13, 16, 18, 20, 21, 26, 27, 33, 35, 41, 42

SIRS, 33, 41
sistema imune, 12, 57, 66
staphylococcus, 74
susceptíveis, 13, 14, 16–19, 22, 26, 27,
34, 35, 37, 38, 41, 43, 76
susceptível, 77

taxa de prescrição, 78
transcriptase reversa, 57
transmissão, 35, 37–41, 76–78, 80
 de bactérias, 74
 dinâmica de, 73, 74, 76
 direta, 77
 horizontal, 77
 indireta, 77, 78
tratamento, 53, 73–76, 78, 79
tratamentos, 13
tuberculose, 73

vírus, 27, 29, 37, 38, 45, 54, 57, 60–63,
69, 70
vantagem seletiva, 49, 74, 75, 79
viremia, 38
vizinhança, 20–22, 24, 26, 38, 39

NOTAS EM MATEMÁTICA APLICADA

1. Restauração de Imagens com Aplicações em Biologia e Engenharia
Geraldo Cidade, Antônio Silva Neto e Nilson Costa Roberty
2. Fundamentos, Potencialidades e Aplicações de Algoritmos Evolutivos
Leandro dos Santos Coelho
3. Modelos Matemáticos e Métodos Numéricos em Águas Subterrâneas
Edson Wendlander
4. Métodos Numéricos para Equações Diferenciais Parciais
Maria Cristina de Castro Cunha e Maria Amélia Novais Schleicher
5. Modelagem em Biomatemática
Joyce da Silva Bevilacqua, Marat Rafikov e Cláudia de Lello Courtoque Guedes
6. Métodos de Otimização Randômica: algoritmos genéticos e “simulated annealing”
Sezimária F. Pereira Saramago
7. “Matemática Aplicada à Fisiologia e Epidemiologia”
H.M. Yang, R. Sampaio e A. Sri Ranga
8. Uma Introdução à Computação Quântica
Renato Portugal, Carlile Campos Lavor, Luiz Mariano Carvalho e Nelson Maculan
9. Aplicações de Análise Fatorial de Correspondências para Análise de Dados
Dr. Homero Chaib Filho, Embrapa
10. Modelos Matemáticos baseados em autômatos celulares para Geoprocessamento
Marilton Sanchotene de Aguiar, Fábria Amorim da Costa, Graçaliz Pereira Dimuro e Antônio Carlos da Rocha Costa

11. Computabilidade: os limites da Computação
Regivan H. N. Santiago e Benjamín R. C. Bedregal
12. Modelagem Multiescala em Materiais e Estruturas
Fernando Rochinha e Alexandre Madureira
13. Modelagem em Biomatemática (Coraci Malta ed.)
 - 1 - “Modelagem matemática do comportamento elétrico de neurônios e algumas aplicações”
Reynaldo D. Pinto
 - 2 - “Redes complexas e aplicações nas Ciências”
José Carlos M. Mombach
 - 3 - “Possíveis níveis de complexidade na modelagem de sistemas biológicos”
Henrique L. Lenzi, Waldemiro de Souza Romanha e Marcelo Pelajo-Machado
14. A lógica na construção dos argumentos
Angela Cruz e José Eduardo de Almeida Moura
15. Modelagem Matemática e Simulação Numérica em Dinâmica dos Fluidos
Valdemir G. Ferreira, Hélio A. Navarro, Magda K. Kaibara
16. Introdução ao Tratamento da Informação nos Ensinos Fundamental e Médio
Marcilia Andrade Campos, Paulo Figueiredo Lima
17. Teoria dos Conjuntos Fuzzy com Aplicações
Rosana Sueli da Motta Jafelice, Laércio Carvalho de Barros, Rodney Carlos Bassanezi
18. Introdução à Construção de Modelos de Otimização Linear e Inteira
Socorro Rangel
19. Observar e Pensar, antes de Modelar
Flavio Shigeo Yamamoto, Sérgio Alves, Edson P. Marques Filho, Amauri P. de Oliveira
20. Frações Contínuas: Propriedades e Aplicações
Eliana Xavier Linhares de Andrade, Cleonice Fátima Bracciali