UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
DEPARTAMENTO DE FÍSICA E MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM
FÍSICA APLICADA À MEDICINA E BIOLOGIA

CAMINHADAS DETERMINÍSTICAS EM MEIOS
DESORDENADOS:
PROBLEMA DA CAMINHADA DO TURISTA

Gilson Francisco de Lima

Tese apresentada à Faculdade de Filosofia, Ciências e Letras
de Ribeirão Preto da USP, como parte das exigências para
a obtenção do título de Doutor em Ciências: Área: Física
Aplicada à Medicina e Biologia.

RIBEIRÃO PRETO – SP

2002

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
DEPARTAMENTO DE FÍSICA E MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM
FÍSICA APLICADA À MEDICINA E BIOLOGIA

CAMINHADAS DETERMINÍSTICAS EM MEIOS
DESORDENADOS:
PROBLEMA DA CAMINHADA DO TURISTA

Gilson Francisco de Lima
Orientador: Alexandre Souto Martinez
Co-Orientador: Osame Kinouchi Filho

Tese apresentada à Faculdade de Filosofia, Ciências e Letras
de Ribeirão Preto da USP, como parte das exigências para
a obtenção do título de Doutor em Ciências: Área: Física
Aplicada à Medicina e Biologia.

RIBEIRÃO PRETO – SP

2002

FICHA CATALOGRÁFICA

Lima, Gilson Francisco

Caminhadas determinísticas em meios desordenados:
problema da caminhada do turista. Ribeirão Preto,
2002. 142 p.: il.; 30 cm

Tese (Doutorado), apresentada à Faculdade de
Filosofia, Ciências e Letras de Ribeirão Preto / USP.
Departamento de Física e Matemática.

Orientador: Martinez, Alexandre Souto

Co-orientador: Kinouchi Filho, Osame

1. Meios desordenados; 2. Sistemas complexos; 3. Ca-
minhadas determinísticas; 4. Caminhadas aleatórias.

Banca Examinadora:

Alexandre Souto Martinez

Antônio Carlos Roque da Silva Filho

Marco Aurélio Pires Idiart

Nestor Felipe Caticha Alfonso

José Roberto Drugowich de Felício

*Dedico este trabalho
aos meus familiares,
em especial à minha
esposa Irene e ao meu
filho Irison.*

Agradecimentos

Meus agradecimentos mais profundos são dirigidos ao meu co-orientador, Prof. Dr. Osame Kinouchi Filho que me ensinou os primeiros passos nesta área da Física, propôs os modelos da caminhada do turista determinística e estocástica (temperatura) e participou ativamente das discussões e da sistematização de todos os resultados numéricos apresentados nesta tese. Agradeço também ao meu orientador, Prof. Dr. Alexandre Souto Martinez por ter coordenado e orientado todos os trabalhos. A quem eu recorri nos momentos de maior dificuldade e sempre fui atendido com respeito e consideração. De certa forma, o resultado positivo do meu trabalho deveu-se a sua experiência e dedicação.

Desejo manifestar toda a minha gratidão aos meus colegas de trabalho na Escola Técnica Federal de Mato Grosso pelo apoio que recebi quando solicitei o meu afastamento para fazer pós-graduação. Agradeço também ao Sr. Gilmar Francisco de Lima, pelo auxílio financeiro. Isto permitiu que eu pudesse superar muitas dificuldades que surgiram ao longo de todo o curso de pós-graduação, em função da ausência de apoio das instituições de fomento científico ao meu trabalho.

Desejo manifestar a minha gratidão ao jovem Dr. Sebastian Risau-Gusman, por ter se implicado no assunto e ter indicado referências que nos ajudaram a compreender melhor os problemas estudados neste trabalho. Desejo manifestar a minha gratidão à secretária da pós-graduação em Física Aplica a Medicina e a Biologia, do Departamento de Física e Matemática, Gisele Maranghetti Lourenço, pela colaboração nos experimentos com dicionários. Estes experimentos foram fundamentais para a aplicação da dinâmica do turista em grafos de dicionários. Expresso a minha gratidão ao Prof. Dr. José Fernando Fontanari por seu posicionamento crítico em relação ao tema de minha tese e por suas sugestões

construtivas com relação ao uso dos modelos de mapeamentos aleatórios.

Agradeço ao Departamento de Física e Matemática da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, da Universidade de São Paulo, pela acolhida. Em especial ao Prof. Dr. Antônio Carlos Roque da Silva Filho pelo apoio moral incondicional que recebi na ocasião em que tive de mudar de orientador e pela orientação sobre o melhor caminho a ser seguido no decorrer de todo este trabalho.

Várias pessoas contribuíram de maneira construtiva, direta ou indiretamente, para a realização dos diferentes trabalhos que desenvolvi ao longo de todo o período. Por isso, desejo manifestar os meus sinceros agradecimentos aos seguintes profissionais:

- Prof. Dr. Adelaide de Almeida
- Prof. Dr. Carlos Alberto Pelá
- Carlos Alberto Brunello
- Prof. Dr. Evandro Eduardo Seron Ruiz
- Eldereis de Paula
- Élcio Aparecido Navas
- Prof. Dr. José Roberto Drugowich de Felício
- Prof. Dr. José Enrique Rodas Duran
- José Luis Aziani
- José Luiz Bruçó
- Prof. Dr. Nelson Augusto Alves
- Prof. Dr. Oswaldo Baffa Filho
- Prof. Dr. Robert Lee Zimmerman
- Prof. Dr. Ubiraci Pereira da Costa Neves

Resumo

O estudo de caminhadas aleatórias em meios desordenados é um assunto bastante explorado e pode modelar uma grande variedade de problemas, como por exemplo, problemas de transporte (difusão). O estudo de caminhadas determinísticas em meios desordenados é um assunto pouco explorado. Em uma paisagem composta de N sítios distribuídos aleatoriamente no espaço, um caminhante (“turista”) visita estes sítios seguindo a seguinte regra determinística: ir para o sítio vizinho mais próximo que não tenha sido visitado nos últimos τ passos. De cada sítio inicial, a trajetória obtida com esta dinâmica determinística apresenta inicialmente um tempo de transiente t , onde novos sítios são visitados, e no final um atrator de período p , onde os mesmos sítios são sempre revisitados. Apesar da simplicidade do modelo, a dinâmica é complexa e os resultados não são triviais. Para dimensionalidades $d \geq 2$, a distribuição de atratores de período p , obtida numericamente, pode ser descrita por uma lei de potência com um corte exponencial. Os modelos de ligações aleatórias simétricas (que representa o limite de alta dimensionalidade $d \gg 1$ do modelo proposto) e assimétricas indicam que o corte exponencial se torna menos importante à medida que N aumenta. O expoente da lei de potência independe da memória τ , sendo portanto uma distribuição robusta. A dinâmica do turista pode ser aplicada a problemas mais abstratos, onde apenas relações de ordem entre vizinhos são dados. O estudo (por amostragem) da estrutura de um dicionário de sinônimos é um exemplo que foi considerado. Mostrou-se que as palavras podem ser embebidas em um espaço Euclidiano de baixa dimensionalidade. Este resultado concorda com um recente estudo exaustivo realizado e questiona o modelo de análise semântica latente. Com a finalidade de entender a transição entre uma caminhada determinística e uma caminhada aleatória, generalizou-se o problema com memória nula designando uma distribuição de probabilidades para o turista visitar os diversos sítios. Esta distribuição é parametrizada por uma variável externa T (tem-

peratura) de modo que para $T = 0$ têm-se a caminhada do turista como caso limite e para $T \rightarrow \infty$ todos os sítios são visitados com igual probabilidade. Resultados analíticos ($d = 1$) e numéricos mostram a existência de uma região bem delimitada de transição entre os regimes não-ergódico (baixa temperatura) e ergódico (alta temperatura). Uma analogia é estabelecida com o modelo de vidros de Bouchaud. A eficiência da caminhada com relação aos novos sítios visitados, foi estudada e ela é máxima na borda da aleatoriedade, ou seja, ao redor da temperatura de transição.

Abstract

The study of random walks in disordered media is one well-developed subject and it can model a great variety of problems, for instance, problems of transport (diffusion). The study of deterministic walks in disordered media is a subject not too explored. In a landscape composed of N sites randomly distributed in of, a walker (“tourist”) visits these sites following the deterministic rule: going to the nearest site that has not been visited in the last τ steps. From each initial site, the trajectory, obtained with this deterministic dynamics, presents initially a time transient t , where new sites are visited, and, in the end, a p -period attractor, where the same sites are always revisited. In spite of the simplicity of the model, the dynamics is complex and the results are not trivial. For dimensionalities $d \geq 2$, the distribution of p -period obtained numerically can be described by a power law with an exponential cut. The models of symmetrical random connections (that represents the limit of high dimensionality $d \gg 1$ of the proposed model) and asymmetrical random connections indicate that the exponential cut turns out to be less important as N increases. The exponent law of the power law does not depend on the memory τ , being therefore a robust distribution. The tourist dynamics can be applied to more abstract problems, where just relationships of neighbor order are given. The study (by sampling) of the structure of a dictionary of synonyms has been considered. It has been shown that the words can be embedded in a Euclidean space of low dimensionality. This result agrees with a recent exhaustive study accomplished and it challenges the model of latent semantic analysis. With the purpose of understanding the transition between a deterministic and a random walk a generalization of the problem, with null memory has been performed by designating a distribution of probabilities for the tourist to visit the several sites. This distribution has the external variable T (temperature) as a parameter so that, when $T = 0$ it has the tourist walk as a limiting case and for $T \rightarrow \infty$ all of the sites are visited

with equal probability. Analytical numerical results ($d = 1$) show the existence of a well delimited transition between non-ergodic (low temperature) and ergodic (high temperature) regime. An analogy is established Bouchaud glass model. The walk efficiency, regarding the new visited sites to trajectory length, has been studied and it is maximum at the edge of stochasticity, in other words, around the temperature of transition.

Sumário

1	Introdução	1
1.1	Caminhadas determinísticas em redes regulares	3
1.1.1	Resultados obtidos com este modelo	6
1.2	Redes de variáveis lógicas	6
1.3	Mapeamento aleatório de Derrida-Flyvbjerg	10
1.3.1	Mapeamento aleatório de ligações simétricas	12
1.4	Problema da caminhada do turista	12
2	Alguns aspectos gerais	15
2.1	Descrição do modelo	15
2.2	Algoritmo da caminhada do turista	16
2.2.1	Aspectos geométricos	16
2.3	O grafo do turista	18
2.3.1	Distribuição de conectividade no grafo	20
2.3.2	Aspectos dinâmicos	21
2.3.3	Tratamento das trajetórias	24
2.3.4	Caracterização estatística das trajetórias	27
2.4	Resultados analíticos para $\tau = 0$	30
2.5	Conclusão	32
3	Sistemas unidimensionais	35
3.1	Estudo analítico	35
3.1.1	Número médio de sítios por atrator	35
3.2	Resultados	37
3.2.1	Efeito de tamanho finito para $\tau = 0$	38
3.2.2	Distribuição de tempos de transiente	39
3.2.3	Bacias de atração	42

3.2.4	Densidade de atratores	45
3.2.5	Distribuição de atratores	47
3.3	Conclusão	48
4	Sistemas bidimensionais	49
4.1	Características da paisagem	50
4.2	Efeito de tamanho finito	52
4.2.1	Estudo de tamanho finito para $\tau = 0$	53
4.3	Resultados para $\tau > 0$	53
4.3.1	Distribuição de tempos de transiente	54
4.3.2	Comportamento das bacias de atração	55
4.3.3	Bacias de atração de período p	59
4.3.4	Densidade de atratores de período p	60
4.3.5	Densidade total de atratores	61
4.3.6	Probabilidade de distribuição de atratores	62
4.3.7	Número médio de sítios por atrator	64
4.4	Conclusão	65
5	Efeito da dimensionalidade	67
5.1	Espaço de ligações aleatórias assimétricas	67
5.1.1	Tempos de transiente	68
5.1.2	Efeito de tamanho finito com $\tau = 0$	69
5.1.3	Distribuição de períodos de atratores	71
5.2	Espaço de ligações aleatórias simétricas	73
5.2.1	Tempos de transiente com um passo de memória	75
5.2.2	Distribuições de atratores	75
5.3	Interpolação entre os dois modelos	76
5.4	Efeito da dimensionalidade do espaço	77
5.4.1	Tempos de transiente	78
5.4.2	Comportamento das bacias de atração	78
5.4.3	Distribuição de atratores	79
5.4.4	Número médio de sítios por atrator	80
5.5	Conclusão	81

6	Aplicação a grafos de dicionários	85
6.1	Teste dos dicionários	87
6.1.1	Caminhadas com $\tau = 1$	89
6.2	Conclusão	91
7	Caminhadas tendenciosas	93
7.1	Modelo estocástico da dinâmica do turista	94
7.2	O sistema a altas temperaturas	96
7.3	O sistema a baixas temperaturas	96
7.3.1	Distribuição de distâncias entre vizinhos reflexivos	97
7.3.2	Tempos de residência nos casais	97
7.3.3	Tempo de residência	99
7.4	Experimentos numéricos	102
7.4.1	Sistemas unidimensionais	104
7.4.2	Sistemas bidimensionais	107
7.5	Conclusão	110
8	Conclusão	113

Lista de Tabelas

1.1	Regras de atualização para $k = 0$	7
1.2	Regras de atualização para $k = 1$	7
1.3	Regras de atualização para $k = 2$	7
1.4	Evolução do sistema com $N = 5$ e $k = 2$	9
2.1	Alguns valores de p_d e $D_0(2, d)$ em função de d	32
5.1	Expoente das distribuições de atratores com $\tau = 0$ e $\tau = 1$ para o modelo de ligações aleatórias assimétricas.	74
5.2	Quando fazemos a comparação dos valores usados para ajustar os resultados numéricos referentes aos modelos de ligações aleatórias simétricas (si) e assimétricas (assi), observamos que o comportamento das distribuições de transiente é robusto.	78
5.3	Comparação dos resultados numéricos das distribuições de atratores $P_\tau(p)$ referentes aos modelos de ligações simétricas (si) e assimétricas (assi). Para $\tau = 1$ observa-se que os expoentes são os mesmos para os dois modelos.	80
7.1	Probabilidade do viajante permanecer no par de sítios a e b , esse exemplo inicia-se em a	98

Lista de Figuras

2.1	Rede unidimensional: a) uma rede unidimensional com pontos distribuídos aleatoriamente (condições livres); b) o uso de condições periódicas simula um sistema infinito repetindo infinitamente uma mesma cela unitária; c) o efeito das condições periódicas de contorno em uma rede unidimensional é equivalente a unir os dois extremos da rede base a).	17
2.2	Exemplo de como são ordenados na matriz $V(i, k)$ os vizinhos mais próximos de um vértice i para $\tau = 4$. Aqui $D(i, j)$ representa a distância que separa o sítio i de seu k -ésimo vizinho, que possui índice j .	19
2.3	Distribuição de grau de entrada para um único grafo com $N = 3000$ vértices.	22
2.4	Desvio padrão do comportamento gaussiano apresentado na Figura 2.3. A linha cheia é dada pela Equação 2.4.	23
2.5	Exemplo de uma medida de tempo de transiente e de período de atrator. O período $p = 3$ do atrator e o transiente de tempo t são identificados no instante em que o valor absoluto da diferença sítio a sítio, de dois ciclos consecutivos da mesma trajetória é nula.	25

- 2.6 Exemplo de como determinar se os atratores encontrados são diferentes. Nesse caso a matriz de atratores tem os atratores de período $p = 2$ constituído dos sítios (6 100) e (3 10). Foi encontrado outro atrator de período $p = 2$ formado pelos sítios (100 5). Testa-se todas as combinações possíveis dos sítios pertencentes ao atrator encontrado com os sítios do primeiro atrator gravado na matriz de atratores. Depois dá-se um salto proporcional $2p.at(p)+1$ e repete-se os testes com o atrator seguinte. Neste exemplo, verifica-se que realmente o ciclo (100 5) é um ciclo novo 26
- 3.1 Exemplos dos diferentes tipos de atratores em $d = 1$: a) um bloco-base para $\tau = 1$; b) um bloco-base para $\tau = 2$; c) um ciclo permitido para $\tau = 1$ em que $n_I = 0$ e $p = 8$; d) um ciclo constituído de dois blocos-base e $n_I = 1$ sítio intermediário para $\tau = 1$ e $p = 10$; e) um exemplo de ciclo ímpar ($p = 11$) para $\tau = 2$ que é possível por causa de um laço interno (*loop*). 36
- 3.2 Número médio de sítios por atrator $\langle n(p) \rangle$: a) $\tau = 1$ (triângulo); b) $\tau = 6$ (retângulos). As linhas mostram a previsão teórica de cada τ 38
- 3.3 Distribuição de tempos de transiente medidos com $\tau = 0$ em $d = 1$, para $N = 100, 200, 400, 800, 1600, 3200$ e 6400 . O número de realizações foi escolhido de forma que variando N o número de trajetórias seja $N.N_R = 640000$ 39
- 3.4 O efeito de tamanho finito na densidade de atratores $D_0(2)$ em $d = 1$, com $\tau = 0$. A densidade de atratores converge com N^{-1} para o valor previsto teoricamente. 40
- 3.5 A distribuição dos tempos de transiente $P_\tau(t)$ para $d = 1$ decai exponencialmente. Da esquerda para a direita: $\tau = 0, 1, 2, 3, 4, 5$ e 6 , com $N = 10000$ e $N_R = 1000$ 41
- 3.6 Tempo de decaimento característico $\xi(\tau)$, dos tempos de transiente, para valores de memória $\tau = 0, 1, 2, 3, 4, 5$ e 6 42

3.7	Probabilidade de um viajante que parte de um sítio escolhido aleatoriamente pertencer a bacia de atração dos atratores de período p . Para $\tau = 1$ (triângulos), $\tau = 3$ (circunferência) e $\tau = 6$ (quadrados).	43
3.8	Ampliação da parte inicial da distribuição de probabilidade em escala mono-log. Para $\tau = 1$ (triângulos), $\tau = 3$ (circunferência) e $\tau = 6$ (quadrados).	44
3.9	Tamanho médio das bacias de atração para $\tau = 1, 3$ e 6 (de baixo para cima). Cada valor de memória tem um comportamento diferente.	44
3.10	Exemplos de densidade de atratores de período p em $d = 1$. Da esquerda para a direita $\tau = 1, 3$ e 6 .	45
3.11	A densidade total de atratores para $d = 1$ decai exponencialmente com a memória. Os resultados apresentados referem-se a $\tau = 1, 2, 3, 4, 5$ e 6 .	46
3.12	Produto da densidade total de atratores $D(\tau)$ pelos tempos de decaimento dos tempos transiente $\xi(\tau)$.	46
3.13	A probabilidade de distribuição de atratores de período p em $d = 1$ para $\tau = 1$ (triângulo), 3 (circunferência) e 6 (retângulo).	47
4.1	Exemplo de superposição de atratores para $d = 2$ e $\tau = 1$. Partindo de um sítio A , obtemos um atrator de período 4 ($ABCD$), mas partindo de C encontramos um atrator de período 3 (CBE).	50
4.2	Exemplos do aparecimento de diferentes atratores para $d = 2$ e $N = 400$, com $\tau = 1, 3, 5$ e 9 . A mesma paisagem foi usada para cada um destes gráficos, deste modo pode-se observar o crescimento de diferentes atratores com τ . Foram utilizadas condições livres de contorno. Os pontos que não pertencem a nenhum ciclo não foram mostrados.	51
4.3	Exemplos do aparecimento de diferentes atratores para $d = 2$ e $N = 400$, com $\tau = 2, 4, 6$ e 10 . A mesma paisagem da Figura 4.2 foi usada para cada um dos gráficos.	52

4.4	$D_0(2)$ em função de N^{-1} . Varia-se N e o número realizações N_R mantendo constante o número de caminhos ($N.N_R = 640000$). Os valores utilizados foram: $N = 100$ e $N_R = 6400$; $N = 200$ e $N_R = 3200$; $N = 400$ e $N_R = 1600$; $N = 800$ e $N_R = 800$; $N = 1600$ e $N_R = 400$; $N = 3200$ e $N_R = 200$; $N = 6400$ e $N_R = 100$	54
4.5	Distribuição de tempos de transiente $P_\tau(t)$ em $d = 2$ para $\tau = 0, 2, 3, 5, 7$ e 10 (da esquerda para a direita). Os tempos de transiente em $d = 2$ são menores que em $d = 1$	55
4.6	Tempo de decaimento característico $\xi(\tau)$ em função da memória τ . Diferentemente do caso unidimensional $\xi(\tau)$ tem um comportamento de lei de potência em função de τ	56
4.7	Probabilidade de um viajante pertencer a bacia de atração dos atratores de período p para $\tau = 1$ (triângulo), $\tau = 2$ (estrelas), $\tau = 4$ (circunferências), $\tau = 10$ (retângulos).	57
4.8	Comportamento inicial das distribuições de bacias de atração dos atratores de período p para $\tau = 1$ (triângulo), $\tau = 2$ (estrelas), $\tau = 4$ (circunferências), $\tau = 10$ (retângulos).	58
4.9	Comportamento dos períodos mínimos para $\tau \gg 1$. A linha sólida corresponde à Equação 4.6 e os quadrados correspondem aos valores numéricos.	58
4.10	Colapso dos dados da distribuição de probabilidade $P_A(p)$ usando a Equação 4.7, para $\tau = 1$ (triângulo), $\tau = 2$ (estrelas), $\tau = 4$ (circunferências), $\tau = 10$ (retângulos).	59
4.11	Tamanho médio das bacias de atração de período p para $\tau = 1$ (estrela), $\tau = 2$ (circunferências), $\tau = 4$ (retângulos), $\tau = 10$ (triângulo).	60
4.12	Exemplos de densidade de atratores $D_\tau(p)$: $\tau = 1$ (quadrado), $\tau = 4$ (estrelas) e $\tau = 10$ (circunferências).	60
4.13	Densidade total de atratores em função da memória τ . Existe um comportamento de lei de potência, que é função de τ , para $d = 2$. .	61

4.14	O produto da constante de decaimento pela densidade total de atratores de períodos p mostra que as contribuições de lei de potência se cancelam.	62
4.15	Probabilidade de distribuição de atratores de período p para $\tau = 1$ (triângulo), $\tau = 2$ (estrela), $\tau = 4$ (circunferências), $\tau = 10$ (retângulos).	63
4.16	A Equação 4.12 representa o colapso de todas as distribuições $P_\tau(p)$, como ilustrado em função de p/τ para $\tau = 1$ (triângulos), 2 (estrelas), 4 (circunferências) e 10 (retângulos).	64
4.17	Número médio de sítios por atrator: $\tau = 1$ (estrelas), $\tau = 2$ (triângulo), $\tau = 4$ (cruz), $\tau = 6$ (retângulos), $\tau = 10$ (circunferências).	65
5.1	Os tempos de transiente de uma caminhada determinística sobre uma configuração aleatória de sítios com conexões assimétricas escalam com $N^{1/2}$. Mostra-se $N = 50, 100, 200, 400, 800$ e 1600 para $\tau = 0$ (memória nula). A linha cheia é a curva analítica dada pela Equação 5.2 com $\xi_0 = 0, 94$; $A = 1, 2$ e $\gamma = 1, 6$	69
5.2	A densidade total de atratores $D(\tau) = \sum_{p=2}^{p=t_f} D_\tau(p)$ na caminhada do turista com $\tau = 0$, no modelo de ligações aleatórias assimétricas, tende para zero quando o tamanho do sistema aumenta. São mostrados $N = 50, 100, 200, 400, 800$ e 1600	70
5.3	Densidade total de período p de atratores em função do inverso do tamanho do sistema para $N = 50, 100, 200, 400, 800$ e 1600 . Os resultados são referentes $\tau = 0$, no modelo de ligações aleatórias assimétricas. De cima para baixo o período dos ciclos mostrados são: $p = 2, 3, 4, 6, 8$ e 16	71
5.4	No modelo de ligações aleatórias assimétricas com $\tau = 0$ o número de atratores (produto da densidade total de atratores pelo número de sítios), é proporcional a $\log(N)$. São mostrados $N = 50, 100, 200, 400, 800$ e 1600	71

- 5.5 Colapso das distribuições de atratores, referentes a caminhada do turista com $\tau = 0$, para $N = 100, 200, 400, 800$ e 1600 . A linha cheia é a curva analítica dada pela Equação 5.7 ajustada com $\tilde{p}_0 = 1,78$; $C = 0,35 \pm 0,05$; $\alpha = 1$ e $\beta = 2$ 72
- 5.6 Colapso das distribuições de atratores para $\tau = 1$ referentes a $N = 100, 200, 400, 800$ e 1600 no modelo de ligações assimétricas. A linha cheia é a curva analítica ajustada com $\tilde{p}_0 = 2,40$; $C = 0,35 \pm 0,05$; $\alpha = 1$ e $\beta = 2$ 73
- 5.7 Efeito de tamanho finito sobre a densidade de atratores, na caminhada do turista com $\tau = 0$, no grafo construído a partir do modelo de ligações aleatórias simétricas. São mostrados $N = 50, 100, 200, 400, 800$ e 1600 74
- 5.8 Na aproximação de ligações aleatórias simétricas os tempos de transiente também escalam com $N^{1/2}$ como foi descrito na Equação 5.3, com $\xi = 1,59$, $A = 0,70$ e $\gamma = 1,6$. Aqui estamos mostrando $N = 80, 160, 320, 400$ e 1600 com $\tau = 1$ 75
- 5.9 Efeito de tamanho finito na distribuição de atratores para $\tau = 1$ no modelo de ligações aleatórias simétricas. Da esquerda para direita tem-se: $N = 10, 20, 40, 320$ e 1600 , respectivamente. 76
- 5.10 Densidade total de atratores em função da probabilidade de assimetria das conexões do vértices do grafo, para $N = 1000$ e $N_R = 100$. São mostrados $\tau = 0$ e $\tau = 10$ 77
- 5.11 Distribuições de tempos de transiente para caminhadas do turista com $\tau = 1$. Da esquerda para direita tem-se $d = 1, 2, 3, 4, 8, 16, 32$ e 64 . Observa-se a convergência no limite de $d \rightarrow \infty$, para o modelo de ligações aleatórias simétricas (circunferências vazias). Com a finalidade de comparação, os quadrados vazios correspondem ao modelo de ligações aleatórias assimétricas (mapeamento aleatório assimétrico), que não é um bom limite para $d \rightarrow \infty$ 79

5.12	Convergência do expoente da distribuição de atratores $P_{\tau=1}(t)$ para o limite de $d \rightarrow \infty$. Da direita para esquerda tem-se $d = 1, 2, 4, 8, 16, 32$ e 64	80
5.13	Probabilidade do caminhante pertencer a bacia de atração dos atratores de período p para $\tau = 1$. Da esquerda para direita tem-se: $d = 1$ (triângulo), $d = 2, 4, 8, 16$ e 32 . Observa-se uma convergência lenta para o modelo de ligações aleatórias simétricas (circunferências vazias). Os quadrados vazios correspondem ao modelo de ligações aleatórias assimétricas (mapeamento aleatório de Derrida-Flyvbjerg). As curvas analíticas são dadas pela Equação 5.6.	81
5.14	Probabilidade $P_{\tau}(p)$ de distribuição de atratores de período p para $\tau = 1$. Da esquerda para direita tem-se: $d = 1, 2, 4, 8, 16, 32$ e 64 . Observa-se a existência de uma convergência lenta para o modelo de ligações aleatórias simétricas (as circunferências vazias). Os quadrados vazios correspondem ao modelo de ligações aleatórias assimétricas (mapeamento aleatório de Derrida-Flyvbjerg).	82
5.15	Expoente da distribuição de atratores $P_{\tau}(p)$ em função de $1/d$ para $\tau = 1$. Da direita para esquerda tem-se $d = 2, 4, 8, 16, 32$ e 64	83
5.16	Número médio de sítios por atrator para $\tau = 1$ e $d = 1, 2, 4, 8, 16$ e 32 . As maiores flutuações são para dimensionalidades intermediárias. Os retângulos correspondem ao modelo de ligações aleatórias assimétrica e a circunferência ao modelo de ligações simétricas.	83
6.1	Distribuição de tempos de transiente para o dicionário de inglês americano (MS98 USA) com $\tau = 0$ e $N = 1000$ palavras iniciais.	89

- 6.2 Comparação entre a fração de sítios pertencentes a ciclos de período p $P_A(p)$ formados com $\tau = 1$ em um grafo Euclidiano aleatório e a distribuição de atratores de um grafo formado por $N = 700$ palavras iniciais do dicionário de inglês da MS98 com $d = 2$ (círculos). O resultado para grafos com ligações aleatórias assimétricas também é mostrado (quadrado). 90
- 7.1 Sistema unidimensional com $N = 8$. Observe que D_0 é a distância entre o casal, D_1 a distância do primeiro vizinho do casal e D_2 a distância do segundo vizinho do casal. 100
- 7.2 Densidade de probabilidade de estabilidade para $d=1$, em função da distância dos primeiro e segundo vizinhos do casal. As linhas sólidas são os resultados analíticos das Equações 7.29 e 7.30 . . . 102
- 7.3 Tempo de residência médio para $\beta < \beta_c$. Dados obtidos com $N = 1000$, número de realizações $N_R = 200$ e $\beta_c = 2$. A linha sólida corresponde a $2/(2 - \beta)$ 103
- 7.4 Medidas da taxa de visitação $\langle R(\beta) \rangle$ e da probabilidade do caminhante ir para o sítio mais próximo $\langle P_{nn}(\beta) \rangle$. A variância $s_R^2(\beta)$ em $d = 1$ para $\langle R(\beta) \rangle$ em diferentes valores de temperatura $T = 1/\beta$. Da direita (estrelas) para a esquerda (circunferências) $N = 100, 200, 400$ e 800 a média foi obtida ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$), enviando um viajante por realização e usando número de passos $t_f = 4N^2$ 105
- 7.5 O deslocamento médio $\langle r(\beta) \rangle$ em $d = 1$ para diferentes valores de temperatura $T = 1/\beta$. São mostrados $N = 400$ (triângulo) e 800 (circunferência). As médias foram obtidas ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$), enviando um viajante por realização e usando número de passos $t_f = 4N^2$ 106

- 7.6 Medidas de eficiência $\langle \eta(\beta) \rangle$ em $d = 1$ para diferentes valores de temperatura β . Da direita (estrelas) para a esquerda (circunferências) $N = 100, 200, 400$, e 800 . Os resultados foram obtidos ao longo de $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$) e usando um número de passos $t_f = 4N^2$ 107
- 7.7 Medidas da taxa de visitaço $\langle R(\beta) \rangle$ e da probabilidade do caminhante ir para o sítio mais próximo $\langle P_{nn}(\beta) \rangle$. A variância $s_R^2(\beta)$ em $d = 2$ para $\langle R(\beta) \rangle$ em diferentes valores de temperatura $T = 1/\beta$. Da esquerda (triângulos) para a direita (circunferências) $N = 200, 400, 800$ e 1600 . As médias foram obtidas ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$) enviando um viajante por realização e usando número de passos $t_f = 4N^2$ 108
- 7.8 Medidas do deslocamento médio $\langle r(\beta) \rangle$ em $d = 2$ para diferentes valores de temperatura β . Da esquerda (estrelas) para a direita (circunferências) $N = 400, 800$ e 1600 . As médias foram obtidas ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$), enviando um viajante por realização com um tempo de vida $t_f = 4N^2$ 109
- 7.9 Medidas da eficiência $\langle \eta(\beta) \rangle$ em $d = 2$ para diferentes valores de temperatura β . Os valores de N são: 200 (triângulos), 400 (estrelas), 800 (circunferência) e 1600 (retângulos). As médias foram obtidas ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$) e usando um número de passos $t_f = 4N^2$ 110

Capítulo 1

Introdução

A essência de todos os valores humanos é a vida e o viver em toda a sua plenitude, em uma sociedade cujas regras de convívio social promovam o bem estar coletivo com respeito às liberdades individuais. Nesse contexto social a liberdade intelectual, ou o desenvolvimento de nossa capacidade de compreensão, é um estado de espírito. O desapego físico e de nossas próprias concepções de verdade nos permite observar e avaliar situações distintas com diferentes referenciais teóricos e experimentais. A partir da realidade observada, onde tudo que percebemos e sentimos é relativo aos nossos parâmetros de avaliação, podemos construir esquemas mentais, que no momento oportuno, nos permitirão operar sobre tais sistemas. Isto torna possível ao cérebro invocar atitudes com base na imagem de nossa concepção de mundo. Estas atitudes nos levam a interagir com diferentes realidades e permitem que nos adaptemos a diferentes situações. Cria-se condições para o nosso desenvolvimento cultural, onde materializamos a nossa própria fé.

Nesta perspectiva de realidade, as teorias nascem de experiências individuais colocadas dentro de uma lógica matemática. A abertura de novos horizontes impõe a necessidade de aperfeiçoamento e de ampliação destas estruturas, aplicando a metodologia científica em contextos individuais e coletivos. Do ponto de vista da dinâmica natural de todo evento físico ou intelectual, tudo na vida tem um custo, e embora nem sempre seja o capital, cada um tem a liberdade que pode pagar.

Em um mundo livre e democrático, onde as pessoas podem viver sem

medo ou coação, todo indivíduo tem a sua própria dinâmica e o seu determinismo é afetado pela sua história de vida. Podemos imaginar a vida como uma viagem por um mundo de propriedades abstratas, cujas possibilidades de ação, como no comportamento social mencionado acima, podem ser representadas por pontos. As ligações entre estas ações podem ser representadas por linhas. A representação de tal estrutura é chamada de rede ou de grafo. A dinâmica de cada indivíduo caminhando sobre tais grafos segue uma regra definida pelo seu determinismo individual e pode ser afetada por uma componente estocástica (individual ou coletiva).

Nós vivemos e navegamos em um mundo formado por redes: biológicas, sociais, lingüísticas, tecnológicas etc. O estudo e a caracterização de tais redes tem se intensificado recentemente devido ao surgimento de novas idéias e do desenvolvimento de instrumentos computacionais. As bases de dados têm aumentado muito rapidamente, em função do uso do computador para testar modelos e relacionar com os resultados experimentais (1).

Estas estruturas topológicas foram muito estudadas em duas situações em particular:

- redes ordenadas (como por exemplo uma rede cristalina);
- redes desordenadas (grafos aleatórios).

As redes ordenadas são caracterizadas pela simetria de translação enquanto que as redes desordenadas são assimétricas. Mais recentemente foi introduzido um novo tipo de redes que estão situadas entre as redes ordenadas (ordem local e global) e as redes aleatórias (desordem global), são as redes de mundo pequeno (1).

Nesta tese estuda-se uma dinâmica determinística de minimização local em um novo tipo de rede (ou grafo) aleatória, que seguindo a sugestão de Peter Stadler, denominamos de: *o problema da caminhada do turista*.

A seguir descrevemos alguns trabalhos que podem ser considerados precursores ao estudo que vamos apresentar.

1.1 Caminhadas determinísticas em redes regulares

No livro de Alice, provavelmente a passagem mais citada são as palavras da Rainha Vermelha para a menina: “Agora, como você vê, aqui você deve correr o máximo que conseguir para ficar no mesmo lugar. Se você quiser chegar em algum outro lugar, você deve correr pelo menos duas vezes mais rápido do que isso”.

Em um ecossistema, uma espécie que não muda fica vulnerável aos parasitas ou predadores. No caso de um predador, se ele não muda, as presas desenvolvem estratégias de defesa eficiente. Portanto, as espécies estão sempre evoluindo, seja geneticamente, seja em termos comportamentais. Considere o exemplo sugerido por Roger Lewin do livro *Complexidade* (2). Imagine uma mosca. Ela tem uma paisagem de aptidão. Agora imagine uma rã. Ela também tem uma paisagem de aptidão. Mas estas paisagens não são independentes. A rã lança a língua e captura a mosca. Imagine que a mosca desenvolva pés escorregadios, de modo que a língua da rã não possa prendê-la. A rã fica sem se alimentar e o seu pico na paisagem de aptidão diminui, ou seja, a rã se torna menos apta. Como a mosca é mais apta, o seu pico na paisagem de aptidão aumenta. Assim, paisagens acopladas mudam, uma interferindo na outra. O próximo passo é que a rã, por exemplo, desenvolve pêlos na língua e torna-se capaz de capturar a mosca novamente. As aptidões mudam e as paisagens mudam.

Estas mudanças podem ser forçadas por alterações no ambiente (mudanças exógenas) ou devido as interações entre as espécies (mudanças endógenas). Esta última pode ser exemplificada pelo efeito da Rainha Vermelha. Segundo Freund e Grassberger (3) mesmo que as mudanças endógenas sejam um efeito secundário, sua importância não deve ser desprezada.

O modelo proposto por Freund e Grassberger (3) demonstra explicitamente que a caminhada da Rainha Vermelha pode levar à evolução, em um sistema não-estacionário e não-aleatório.

A caminhada da Rainha Vermelha apresenta uma estrutura do espaço de fase muito diferente dos modelos de Lotka-Volterra (com mais do que 2 agentes), que é a maneira usual de formular as relações entre presa e predador, constituindo talvez em uma maneira alternativa para se abordar o problema.

O modelo é definido como segue. Considere um caminhante sobre uma rede quadrada de tamanho $L \times L$ com condições periódicas de contorno (3):

1. anexa-se a cada sítio i um número real, positivo, chamado de “profundidade” h_i ;
 - (a) no tempo $t = 0$, assume-se que a paisagem é uniforme, $h = 0$, e o caminhante (por exemplo, na origem) segue para um sítio e uma direção escolhida aleatoriamente;
 - (b) a profundidade h_i mede o quanto o caminhante danificou o sítio i no passado. Portanto, embora o caminhante permaneça sobre o sítio i , ele afunda como se estivesse em um atoleiro, e a profundidade do solo aumenta de uma quantia fixa por unidade de tempo. Assume-se que:

$$h_i(t + 1) = h_i(t) + 1. \quad (1.1)$$

2. os sítios vazios relaxam seguindo alguma lei do tipo:

$$h_j(t + 1) = f[h_j(t)]. \quad (1.2)$$

A função $f(h)$ estabelece que um sítio não visitado pelo caminhante não varia com relação a $h = 0$ e impede que o caminhante descanse tranquilo em uma mesma região, proibindo movimentos de 180° ;

3. o determinismo do modelo é assegurado por uma regra que prescreve que **em cada passo de tempo o caminhante sempre deve ir para o sítio de sua vizinhança que tenha a menor profundidade**;
 - (a) a Rainha Vermelha (o caminhante) quer evitar os danos no meio ambiente que ela própria causou no passado (por exemplo, a profundidade dos sítios simula o nível de presas);
 - (b) isto é feito localmente, i.e., não usando uma estratégia de longo alcance, e a decisão é baseada somente na profundidade do presente sítio e seus quatro primeiros vizinhos;
 - (c) se existir dois ou mais sítios com profundidade baixa, ela adota a regra determinística para definir qual sítio é o preferido. Se os sítios da vizinhança estiverem com mesma profundidade ela segue uma regra isotrópica (rotacionalmente simétrica), onde três situações distintas são definidas para a direção de preferência:
 - i. linear, esquerdo, direito (regra *LED*);
 - ii. esquerdo, linear, direito (regra *ELD*);
 - iii. direito, esquerdo, linear (regra *DEL*);
4. a dinâmica definida acima sugere os seguintes comportamentos:
 - (a) o caminhante não percorre a rede inteira, porque alguns sítios nunca são visitados;
 - (b) como a rede é finita, alguns sítios devem ser visitados muitas vezes;
 - (c) assumindo que cada um destes sítios tem um vizinho j que é visitado somente um número finito de vezes, depois de um tempo finito o caminhante passando pelo mesmo sítio sempre vai preferir seguir por uma direção diferente.

5. A ordem de profundidade dos vizinhos de i é exatamente a ordem de visitas no tempo. Como consequência a ordem em que os vizinhos serão visitados outra vez quando a Rainha Vermelha residir no sítio i novamente é a mesma, a não ser que um vizinho tenha sido visitado em um tempo inferior por um caminhante que não passou através do sítio i .

1.1.1 Resultados obtidos com este modelo

1. Para redes pequenas as caminhadas são periódicas;
2. os resultados dependem muito dos detalhes da implementação, i.e., da preferência entre os sítios que têm uma profundidade pequena;
3. a periodicidade tem uma dependência muito forte dos valores de L . Em geral foi observado um comportamento diferente para valores diferentes de L ;
4. foi feita uma estimativa dos tempos de transiente em que ambas as distribuições de profundidade coincidem. As caminhadas foram realizadas sobre redes pequenas, para as três regras consideradas, partindo das mesmas condições iniciais;
5. a dinâmica independe dos detalhes da função $f(h)$.

1.2 Redes de variáveis lógicas

Considere n variáveis binárias σ_i . Estas variáveis podem ser lógicas $(0, 1)$ ou podem ser vistas como spins de Ising $(-1, 1)$. Um estado ou configuração é definido pelo vetor:

$$\vec{C}(t) = \{\sigma_1(t), \sigma_2(t), \dots, \sigma_n(t)\} , \quad (1.3)$$

de modo que existem $M = 2^n$ estados possíveis para cada instante t . Se a evolução temporal do sistema for tal que o valor de cada variável σ_i , no instante $t + 1$, dependa de k variáveis σ_{i_k} no instante precedente, têm-se uma atualização em paralelo:

$$\vec{C}(t+1) = f \left[\vec{C}(t) \right] , \quad (1.4)$$

que corresponde à:

$$\sigma_i(t+1) = f_i[\{\sigma_{i_1}(t), \sigma_{i_2}(t), \dots, \sigma_{i_k}(t)\}] . \quad (1.5)$$

A função f_i resulta nos valores 0 ou 1. Para k spins existem 2^{2^k} regras possíveis conforme os exemplos apresentados nas Tabelas 1.1, 1.2 e 1.3:

- para $k = 0$,

$f^{(1)}$	$f^{(2)}$
0	1

Tabela 1.1: Regras de atualização para $k = 0$

- para $k = 1$

$\{\sigma_{i_1}\}$	$f^{(1)}$	$f^{(2)}$	$f^{(3)}$	$f^{(4)}$
$\{0\}$	0	1	0	1
$\{1\}$	0	0	1	1

Tabela 1.2: Regras de atualização para $k = 1$.

- para $k = 2$

$\{\sigma_{i_1}, \sigma_{i_2}\}$	$f^{(1)}$	$f^{(2)}$	$f^{(3)}$	$f^{(4)}$	$f^{(5)}$	$f^{(6)}$	$f^{(7)}$	$f^{(8)}$	$f^{(9)}$	$f^{(10)}$	$f^{(11)}$	$f^{(12)}$	$f^{(13)}$	$f^{(14)}$	$f^{(15)}$	$f^{(16)}$
$\{0, 0\}$	0	1	0	0	0	0	0	1	0	1	1	1	1	1	0	1
$\{0, 1\}$	0	0	1	0	0	0	1	0	1	0	1	1	1	0	1	1
$\{1, 0\}$	0	0	0	1	0	1	0	0	1	1	0	1	0	1	1	1
$\{1, 1\}$	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1

Tabela 1.3: Regras de atualização para $k = 2$.

Em geral este problema é tratado utilizando condições periódicas de contorno, de tal modo que: $\sigma_{n+1} = \sigma_1$, $\sigma_{n+2} = \sigma_2, \dots$ e o sistema é inicializado com uma configuração aleatória de spins. As variáveis congeladas, que não variam com o tempo, são as regras $f(\{\sigma_{i_1}, \sigma_{i_2}, \dots, \sigma_{i_k}\})$ e as variáveis dinâmicas são os spins σ_i . No início, em $t = 0$, determina-se o valor de k . Se o valor de k referir-se aos primeiros vizinhos, têm-se um autômato celular. Por exemplo, o autômato celular com $k = 3$ é bastante conhecido e consiste em fazer cada spin $\sigma_i(t + 1)$ interagir com $\sigma_{i-1}(t)$, $\sigma_i(t)$ e $\sigma_{i+1}(t)$ com suas 256 regras (classificação de Wolfram). Na versão de rede regular, os spins são conectados com seus primeiros vizinhos. A estocasticidade pode ser introduzida nos autômatos por exemplo nas regras (4) que apresentam diversas aplicações biológicas (5) e medicinais (6).

Uma versão aleatória pode ser imaginada (redes de Kauffman), para cada spin i , selecionando-se de modo aleatório e uniforme os k spins $\sigma_{i_1}, \sigma_{i_2}, \dots, \sigma_{i_k}$ com os quais o spin σ_i interage. Para n spins, existem $\binom{n}{k}$ possibilidades distintas e 2^{2^k} possíveis regras, resultando em $\binom{n}{k} 2^{2^k}$ possíveis realizações.

Observe que se $k = 0$, o sistema sempre converge após um passo de tempo para o estado estacionário definido pelas regras de atualização $\{f_1, f_2, \dots, f_n\}$. Para $k = 1$ o problema foi tratado na referência (7).

Considere por exemplo, um sistema em uma rede aleatória, com $n = 5$ e $k = 2$ com a seguinte estrutura:

$$\begin{aligned}\sigma_1(t + 1) &= f_1^{(8)}[\{\sigma_1(t), \sigma_3(t)\}] \\ \sigma_2(t + 1) &= f_2^{(3)}[\{\sigma_1(t), \sigma_3(t)\}] \\ \sigma_3(t + 1) &= f_3^{(9)}[\{\sigma_2(t), \sigma_5(t)\}] \\ \sigma_4(t + 1) &= f_4^{(15)}[\{\sigma_1(t), \sigma_4(t)\}] \\ \sigma_5(t + 1) &= f_5^{(7)}[\{\sigma_5(t), \sigma_4(t)\}]\end{aligned}\tag{1.6}$$

onde os valores de $f^{(j)}$ são dados pela Tabela 1.3 e a evolução do sistema é mostrada na Tabela 1.4.

Vê-se aqui que $t = 0$, $t = 1$ e $t = 2$ é um tempo de transiente e a partir de $t = 3$ o sistema entra em um ciclo (atrator) de período $p = 4$, caracterizado

$\vec{C}(0)$	$\{0, 1, 1, 0, 1\}$
$\vec{C}(1)$	$\{0, 1, 0, 0, 0\}$
$\vec{C}(2)$	$\{1, 0, 1, 0, 0\}$
$\vec{C}(3)$	$\{1, 0, 0, 1, 1\}$
$\vec{C}(4)$	$\{0, 0, 1, 1, 1\}$
$\vec{C}(5)$	$\{0, 1, 1, 1, 1\}$
$\vec{C}(6)$	$\{0, 1, 0, 1, 1\}$
$\vec{C}(7)$	$\{1, 0, 0, 1, 1\}$
$\vec{C}(8)$	$\{0, 0, 1, 1, 1\}$
$\vec{C}(9)$	$\{0, 1, 1, 1, 1\}$
$\vec{C}(10)$	$\{0, 1, 0, 1, 1\}$
$\vec{C}(11)$	$\{1, 0, 0, 1, 1\}$

Tabela 1.4: Evolução do sistema com $N = 5$ e $k = 2$.

pelas configurações: $\vec{C}(3)$, $\vec{C}(4)$, $\vec{C}(5)$ e $\vec{C}(6)$. Este é um exemplo de rede de Kauffman.

Nas redes de Kauffman todas as configurações iniciais levam a uma trajetória composta de um tempo de transiente e um atrator no final. Sistemas onde o tempo de transiente é nulo, são chamados de sistemas reversíveis (8; 9). As redes lógicas são ditas dissipativas pois vários estados iniciais diferentes podem levar a um mesmo atrator com transiente não nulo (a informação sobre a condição inicial é perdida).

Bhattacharjya e Liang (10) estudaram a distribuição de probabilidade dos períodos dos atratores e dos comprimentos de transiente destas redes críticas ($k = 2$) e encontraram que estas distribuições de probabilidades são uma lei de potência com um expoente que depende fortemente do tamanho da rede.

Um sistema dinâmico é caracterizado por um período característico que cresce em função do número de spins n . Uma vez que a distribuição de períodos de ciclos segue uma lei de potência, o período característico não é dado pelo valor médio, mas sim pelo valor mediano. Assim para $k > 2$, o valor mediano de p

cresce exponencialmente com n , e para $k < 2$, p mediano cresce polinomialmente com n . Se $p_{med} \propto e^n$ diz-se que o sistema é caótico, enquanto que $p_{med} \propto n^\alpha$ diz-se que o sistema é ordenado, com $k = 2$ sendo o valor da conectividade crítica.

Kauffman argumenta que além da seleção natural de Darwin e da mutação aleatória, a auto-organização em sistemas complexos (11) é também responsável pela complexidade observada em sistemas biológicos. A questão crucial então é: quais são as características de sistemas dinâmicos complexos que melhor se adaptam e são mais convenientes para a evolução? Tem sido especulado que os sistemas ótimos estão posicionados na borda da transição entre o regime ordenado e caótico, como para $k = 2$ na rede de Kauffman (10).

Assumindo que as redes de Kauffman com $k = 2$ (na borda do caos) descrevem adequadamente as propriedades estatísticas de uma rede regulatória genética, este resultado implica que o tempo de replicação das células de um organismo com o mesmo tamanho de gene obedece uma lei de potência.

Resultados analíticos para a rede de variáveis aleatórias podem ser encontrado na referência (12).

1.3 Mapeamento aleatório de Derrida-Flyvbjerg

Um modelo interessante de mapeamento aleatório com dinâmica determinística foi estudado por Derrida e Flyvbjerg (13). Este mapeamento aleatório corresponde ao limite em que n e $k \gg 1$ nas redes lógicas aleatórias descritas anteriormente.

O modelo de mapeamento aleatório (13) foi definido considerando um sistema cujo o espaço de fase consiste de M pontos. A dinâmica do sistema é governada por um mapeamento assim definido:

1. para cada ponto $1 \leq i \leq M$ do espaço de fase, escolhe-se um ponto $T(i)$ entre os M pontos existentes de maneira aleatória e uniforme. Se o sistema está na configuração (ou ponto) i no tempo t , estará sobre o ponto $T(i)$ no

tempo $t + 1$. Observe que $T(i) = i$ é permitido;

2. se após t iterações o sistema voltar à configuração i , ele visitará outra vez a configuração $T(i)$ no tempo $t + 1$, ou seja, $T(i)$ é independente do tempo (as variáveis são congeladas).

Este modelo é aleatório uma vez que o mapeamento T é aleatório. Por outro lado, a dinâmica é determinística uma vez que T não se modifica (aleatoriamente) em função do tempo. A evolução temporal de qualquer estado inicial termina em ciclos, uma vez que o espaço de fase é finito. O conjunto dos M pontos é a união das bacias de atração dos vários atratores periódicos (13).

É interessante observar que as propriedades estatísticas da configuração aleatória, de um conjunto de M pontos, foram estudadas nos anos 50 do século XX e os resultados estão compilados na publicação de Harris (14). Estes pesquisadores mostraram que a probabilidade $P_A(p)$ de uma configuração inicial terminar em um atrator de período p é:

$$P_A(p) = \sum_{i \geq p} \frac{\Gamma(M)}{\Gamma(M - i + 1)} \frac{1}{M^i} . \quad (1.7)$$

No limite de $M \rightarrow \infty$, $P_A(p)$ pode ser escrita em termos da função erro complementar

$$P_A(p) = \frac{1}{\sqrt{M}} \int_{p/\sqrt{M}}^{\infty} dy e^{-y^2/2} = \sqrt{\frac{\pi}{2M}} \operatorname{erfc}\left(\frac{p}{\sqrt{2M}}\right) , \quad (1.8)$$

que resulta em um tamanho médio de atrator $\langle p \rangle = \sqrt{\pi M/8}$ e variância $\langle p^2 \rangle - \langle p \rangle^2 = M(2/3 - \pi/8)$.

Também é possível obter a probabilidade de que uma configuração aleatória de sítios tenha exatamente A atratores, de modo que o número médio de atratores cresce logaritmicamente com o tamanho do sistema,

$$\langle A \rangle = \frac{1}{2} \log(M) . \quad (1.9)$$

É interessante observar como o modelo de mapeamento aleatório é o caso limite das redes de Kauffman. O ponto importante é que na rede de Kauffman os n spins estão conectados com k outros spins, e que estes k spins são escolhidos de modo aleatório, independente, e não são necessariamente diferentes. Isto implica que no limite de $k \rightarrow \infty$, as imagens das dinâmicas de duas configurações no espaço de fase não estão correlacionadas, mesmo que estas duas configurações estejam muito próximas.

1.3.1 Mapeamento aleatório de ligações simétricas

No modelo de mapeamento aleatório de Derrida e Flyvbjerg (13), as ligações entre duas configurações i e j não são necessariamente iguais. Um modelo de ligações aleatórias simétricas pode ser construído, mas neste caso a ligação da configuração i com j e da configuração j com i são iguais. Este modelo de ligações aleatórias simétricas foi explorado por Percus e Martin (15) e pode representar distâncias entre dois sítios em espaços com alta dimensionalidade. Ele foi utilizado no estudo do problema do caixeiro viajante, fornecendo alguns resultados analíticos para este sistema (16).

1.4 Problema da caminhada do turista

Nesta tese apresentamos um modelo original de caminhadas determinísticas em paisagens aleatórias, que denominamos de problema da caminhada do turista. Este modelo, apesar de ser simplesmente enunciado, apresenta uma dinâmica muito rica, comum aos sistemas complexos.

Este modelo difere da caminhada da Rainha Vermelha por ser essencialmente um modelo construído no espaço contínuo e não discretizado. Apesar do modelo não modificar a paisagem, este efeito é simulado através da memória que impede o caminhante de visitar lugares recentemente visitados.

O modelo de caminhada do turista também difere das redes lógicas (de

Kauffman) por haver uma regra de proximidade para as variáveis dinâmicas e a memória. No entanto, os resultados da caminhada do turista convergem para os resultados do modelo de ligações aleatórias simétricas no limite de $d \rightarrow \infty$.

No Capítulo 2 apresentamos o modelo da caminhada do turista e enfatizamos as grandezas a serem estudadas numericamente. Em seguida descrevemos resumidamente os aspectos mais relevantes do algoritmo do programa com o qual realizamos os experimentos numéricos, as medidas de tempos de transiente, período de atratores, grafo do turista, etc.

No Capítulo 3 apresentamos alguns resultados analíticos para o cálculo do número médio de sítios por atrator, que posteriormente foram validados pelos resultados numéricos. Mostramos que em $d = 1$ as distribuições estatísticas consideradas tem uma forte dependência com a memória do sistema.

No Capítulo 4 apresentamos os resultados numéricos da caminhada do turista em duas dimensões. Mostramos que a densidade de probabilidade da distribuição de atratores tem um comportamento robusto, em função da memória, como sendo uma lei de potência com um corte exponencial.

No Capítulo 5 consideramos o efeito da dimensionalidade do espaço na caminhada do turista. Neste estudo observamos que a lei que rege os tempos de transiente também são robustas em função da dimensionalidade e que a medida que a dimensionalidade do sistema cresce, os nossos resultados se aproximam do resultado do modelo de ligações aleatórias simétricas. Neste último modelo, resultados analíticos indicam que o corte exponencial na densidade de probabilidade da distribuição de períodos dos atratores depende do tamanho do sistema.

No Capítulo 6 enfatizamos, através de uma aplicação, que as distâncias introduzidas no modelo da caminhada do turista não são variáveis fundamentais, mas somente um mecanismo para produzir o ordenamento entre os sítios. A caminhada do turista é aplicada na caracterização de dicionários de sinônimos. Através deste estudo mostramos que a estrutura dos grafos de dicionários pode ser embebida em espaços Euclidianos de baixa dimensionalidade, ao contrário do

preconizado pelo modelo de análise de semântica latente (17; 18).

No Capítulo 7 consideramos a caminhada do turista sem memória, mas com um parâmetro adicional, o equivalente a temperatura (ruído) de sistemas dinâmicos. Apresentamos então, resultados da caminhada do turista em função deste novo parâmetro. Para altas temperaturas, o sistema se torna ergódico, o viajante pode percorrer todo o espaço. Mas, para baixas temperaturas o viajante fica aprisionado em determinadas regiões, regime não-ergódico. A transição entre estes dois regimes, é uma transição que lembra uma transição vítrea. Uma analogia com o modelo de vidros do Bouchaud (19; 20) é feita.

Finalmente no Capítulo 8 apresentamos as principais conclusões do trabalho.

Capítulo 2

Alguns aspectos gerais

Estamos propondo um novo modelo que pode ser expresso de modo simples, ainda que os resultados da dinâmica inerente sejam bastante intrigantes. Trata-se do estudo das consequências de uma dinâmica de otimização local em uma estrutura aleatória.

2.1 Descrição do modelo

Nosso objetivo aqui é propor e estudar um problema simples e ainda intrigante: uma caminhada determinística sobre um grafo aleatório orientado de N vértices. Estamos interessados nas propriedades estatísticas de uma dinâmica onde a movimentação no grafo segue sempre a mesma regra determinística.

O modelo é definido como segue: N pontos são distribuídos de modo aleatório e uniforme em cada uma das coordenadas de R^d , onde d é a dimensão do espaço Euclidiano. Estes pontos podem ser imaginados como “sítios”, ou como os vértices de um grafo aleatório. Em cada passo de tempo um caminhante, situado em um dado sítio, segue a seguinte regra determinística: ***Vá para o sítio vizinho mais próximo que não tenha sido visitado nos últimos τ passos de tempo.***

Utilizando esta regra o caminhante quer minimizar somente a distância que o separa do sítio mais próximo. Este procedimento produz uma otimização local, não uma otimização da soma de todas as distâncias na trajetória como no problema do caixeiro viajante, ou a otimização de alguma outra função custo

global.

Partindo de um sítio qualquer o viajante faz uma caminhada parcialmente auto-repulsiva. A janela de memória τ regula o grau de repulsão na caminhada do turista e só permite as trajetórias se interceptarem fora desse intervalo. A memória τ pode ser imaginada como um tempo refratário, necessário para que o sítio se torne atrativo novamente ao visitante.

A dinâmica será definida usando-se passos discretos de tempo t , não como uma medida física do tempo gasto quando o viajante anda entre os sítios, mas de quantos sítios ele visitou.

2.2 Algoritmo da caminhada do turista

A seguir vamos descrever o algoritmo utilizado para simular numericamente o sistema proposto e obter as variáveis de entrada para o cálculo das distribuições estatísticas de interesse. Considerando uma janela de memória τ , a dinâmica é bem definida se conhecermos quais são os $\tau + 1$ sítios mais próximos do sítio onde se encontra o caminhante. Os movimentos fora dessa vizinhança nunca são permitidos pela regra determinística.

Este algoritmo é dividido em três partes. Na primeira considera-se todos os aspectos geométricos do problema, ou seja, a construção da paisagem aleatória. Na segunda parte implementa-se a movimentação de um (ou vários viajantes) gerando a trajetória (as trajetórias) de acordo com a regra de minimização local das distâncias. Finalmente, na terceira parte faz-se uma análise de dados para retirar informações de interesse das trajetórias obtidas.

2.2.1 Aspectos geométricos

Embora em simulações numéricas trabalhe-se com sistemas finitos, fenômenos físicos inerentes aos mesmos aparecem somente quando o número de sítios que o compõe for muito grande comparado com a unidade.

Condições periódicas de contorno

O uso de condições periódicas de contorno nos permite minimizar os efeitos de tamanho finito que aparecem devido ao número de sítios considerados. Um sistema relativamente pequeno, como a cela unitária em uma rede cristalina, pode ser usado para simular um sistema infinito repetindo infinitamente a sua estrutura (Figura 2.1b).

Figura 2.1: Rede unidimensional: a) uma rede unidimensional com pontos distribuídos aleatoriamente (condições livres); b) o uso de condições periódicas simula um sistema infinito repetindo infinitamente uma mesma cela unitária; c) o efeito das condições periódicas de contorno em uma rede unidimensional é equivalente a unir os dois extremos da rede base a).

Consideremos um sistema unidimensional onde N pontos são distribuídos aleatoriamente de modo uniforme (Figura 2.1a). Uma maneira de minimizar o efeito de tamanho finito é replicar esse sistema conectando uma extremidade a outra (Figura 2.1b). Este procedimento é equivalente a considerar condições

periódicas de contorno (Figura 2.1c).

Na Figura 2.1a, imagine um caminhante que parte do sítio A para o sítio B sobre uma rede unidimensional seguindo a direção AB (condições de contorno livres),

$$AB = OB - OA. \quad (2.1)$$

A Figura 2.1c mostra que o efeito de unir os dois extremos da rede unidimensional é formar uma paisagem continua e permitir que se possa ir do sítio A para o B pelo caminho alternativo AB' . Portanto, um caminhante sobre a circunferência formada por uma rede unidimensional com condições periódicas pode ir de um sítio à outro por dois caminhos distintos. Usando condições de contorno livres o caminho natural a ser percorrido é AB . Como os sítios estão distribuídos no intervalo de $[0, 1]$ se ele seguisse pelo sentido contrário a distância a ser percorrida seria $AB' = 1 - AB$.

Para aplicar condições periódicas de contorno na dinâmica do turista, toma-se o mínimo entre AB e AB' . Essa idéia pode ser generalizada para dimensões $d > 1$.

2.3 O grafo do turista

A construção da paisagem aleatória começa pela distribuição aleatória das coordenadas dos sítios $x(j, m)$ $j = 1, 2, \dots, N$ com uma densidade uniforme para cada uma das $m = 1, \dots, d$ coordenadas, no intervalo entre $[0, 1]$. Calcula-se as distâncias Euclidianas de cada um dos sítios i com relação a todos os outros sítios da paisagem,

$$D(i, j) = \sqrt{\sum_{m=1}^d [x(i, m) - x(j, m)]^2} \quad (2.2)$$

onde $j(\neq i) = 1, \dots, N$ e d é a dimensão final do espaço a ser considerado. Caso haja condições periódicas de contorno, deve-se levar isso em conta na determinação de $D(i, j)$. Estas distâncias são utilizadas para determinar os $\tau + 1$ vizinhos do sítio i , em ordem crescente, e construir uma tabela de vizinhança $V(i, k) = j(k)$, onde $j(k)$ é o k -ésimo vizinho mais próximo do sítio i .

Enfatiza-se que as distâncias Euclidianas foram introduzida em nosso modelo somente como uma maneira de identificar os sítios vizinhos. Toda a dinâmica da caminhada é desenvolvida sobre a tabela de vizinhos e não sobre a matriz de distâncias. A caminhada do viajante sempre é feita nesse grafo (tabela de vizinhos) que apresenta somente um conjunto ordenado de vizinhos $V(i, k)$ para cada sítio (Figura 2.2).

Figura 2.2: Exemplo de como são ordenados na matriz $V(i, k)$ os vizinhos mais próximos de um vértice i para $\tau = 4$. Aqui $D(i, j)$ representa a distância que separa o sítio i de seu k -ésimo vizinho, que possui índice j .

Cabe aqui um comentário sobre a construção da tabela de vizinhos, ou grafo do turista $V(i, k)$. Inocentemente pode-se construir uma matriz de distâncias de dimensão $N \times N$ e em seguida construir uma tabela de vizinhos de tamanho $N \times (N - 1)$. No entanto este procedimento exige uma grande quantidade de memória, de modo que os sistemas considerados são necessariamente pequenos ($N \sim 10^3$). Um procedimento mais eficaz no uso da memória é gerar um vetor de distâncias $D(j)$ de tamanho N para um dado sítio. Através da procura das menores distâncias constrói-se em seguida a tabela de vizinhos do sítio considerado. Em seguida, calcula-se $D(j)$ para o próximo sítio (apagando-se $D(j)$ anterior). Assim evita-se o armazenamento da matriz $D(i, j)$ total. Note que somente os $\tau + 1$ primeiros vizinhos são necessários. A medida que se considera todos os sítios gerados, tem-se a tabela de vizinhos $V(i, k)$, que tem dimensão de $N \times (\tau + 1)$. Este espaço de memória economizado permite que possam ser simuladas caminhadas em sistemas maiores ($N \sim 10^4$ sítios). A diminuição do número de operações, necessárias para determinar o vizinho mais próximo do caminhante em cada passo de tempo, resulta em economia de tempo durante o processamento dos dados. No limite de $\tau/N \ll 1$, através deste procedimento os

espaços de memória alocados são da ordem de N e não de N^2 .

A tabela $V(i, k)$ define um grafo orientado cujas características estudaremos a seguir. Lembramos que as caminhadas do turista serão trajetórias feitas em um grafo.

2.3.1 Distribuição de conectividade no grafo

Um grafo é dito orientado quando a conexão entre cada par de vértice não é simétrica. Em uma família, por exemplo, se o vértice A representa o pai de B , o vértice B não pode representar o pai de A . Existe portanto uma orientação com um correspondente efeito na representação gráfica do grafo.

Uma aresta é uma ligação não-orientada entre dois vértices e um grafo contendo somente arestas é denominado de não-orientado. Os grafos não-orientados são usados para representar relações simétricas, como por exemplo, uma relação entre irmãos (v é irmão de w , então w é irmão de v). Um arco é uma ligação direcionada entre dois vértices e um grafo conectado somente por arcos é denominado de grafo orientado. Todo grafo orientado corresponde naturalmente a um grafo não-orientado sobre os mesmos vértices, obtido pela substituição de cada arco por uma aresta entre os mesmos pares de vértices.

O grau k de um vértice é dado pelo número de arestas que lhe são incidentes. Quando a rede é orientada, o grau de entrada k_i^{in} e o grau de saída k_i^{out} de um vértice referem-se ao número de arcos que chegam e que saem do referido vértice, respectivamente.

O grafo do turista é definido pela matriz de vizinhos $V(i, k)$, onde:

- i é o conjunto de vértices ou nodos do grafo;
- os pares ordenados de vizinhos $V(i, k)$ são os arcos;
- cada vértice i possui $k = 1, \dots, \tau + 1$ conexões. O grau de saída $k^{out} = \tau + 1$ mede o número de arcos que saem de cada vértice e independe do sítio i ;
- o grau de entrada k_i^{in} é variável como veremos a seguir.

Como um exemplo, considere apenas o grafo contendo o vizinho mais próximo de cada sítio, ou seja, $\tau = 0$ e $k^{out} = 1$. Neste caso, uma conexão é:

- simétrica se os sítios A e B são vizinhos mutuamente mais próximos;
- assimétrica se o sítio A é o vizinho mais próximo do B , mas o vizinho mais próximo de B é o C , por exemplo.

Verifica-se que para valores grandes de τ ($\tau \gg 1$), a distribuição de arcos em cada grafo se aproxima de uma distribuição normal com média $\tau + 1$ (ver a Equação 2.3):

$$N(k^{in}) = \frac{N}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{k^{in} - (\tau + 1)}{\sigma}\right)^2\right]. \quad (2.3)$$

Observe que mesmo para valores menores de τ , o ponto de máximo da distribuição é $k^{out} = \tau + 1$. O desvio padrão pode ser descrito por uma expressão analítica do tipo:

$$\sigma(k^{in}) = C\sqrt{\tau + 1}, \quad (2.4)$$

com $C \approx 1,5$. Este comportamento é mostrado na Figura 2.4.

É interessante observar que este comportamento é parecido com o comportamento observado em uma distribuição binomial, onde o valor médio do número de sucesso é mp (m é o número de passos independentes de Bernoulli e p é a probabilidade de sucesso) e a variância é $mp(1-p)$. Relacionando $mp = \tau + 1$, vemos que o desvio padrão é proporcional a $\sqrt{\tau + 1}$.

2.3.2 Aspectos dinâmicos

A saída da primeira parte do algoritmo é a tabela de vizinhos $V(i, k)$, com $i = 1, 2, \dots, N$ e $k = 1, 2, \dots, \tau + 1$. Esta tabela é a paisagem, ou grafo do turista, sobre o qual um ou vários viajantes vão caminhar. Observe que esta tabela não é mais modificada na execução do programa de simulação numérica. As outras variáveis de entrada na segunda parte da rotina são:

- a memória τ ;

Figura 2.3: Distribuição de grau de entrada para um único grafo com $N = 3000$ vértices.

- o número de paisagens;
- o tamanho das caminhadas t_f (tempo de vida de cada viajante) e;
- o número de trajetórias por paisagem¹.

Todos os sítios por onde o viajante passa, seguindo a regra do turista, são gravados em um vetor que armazena a trajetória $traj(-\tau : t_f)$ com $t = -\tau, -\tau + 1, \dots, 0, 1, 2, \dots, t_f$. Observe que os $\tau + 1$ primeiros valores são inicializados com sítios escolhidos ao acaso. Ao final de um tempo de caminhada t_f , a análise de cada trajetória individual permite determinar quanto tempo t o caminhante

¹Pode-se pensar que um viajante faz n_v trajetórias, saindo de condições iniciais diferentes ou que n_v viajantes faça somente uma trajetória cada um.

Figura 2.4: Desvio padrão do comportamento gaussiano apresentado na Figura 2.3. A linha cheia é dada pela Equação 2.4.

andou até entrar em um ciclo (atrator), o período p do atrator e todas as outras grandezas de interesse (ver a Seção 2.3.4).

A caminhada pode ser desenvolvida sem que se estabeleça um número máximo de passos de tempo t_f para o caminhante. Nesse caso em cada passo de tempo é preciso fazer os testes para saber se o viajante já entrou em um atrator. Para uma janela de memória pequena ($\tau \sim 1$), onde os tempos de transiente e os períodos dos atratores são relativamente pequenos, realmente existe uma economia de tempo computacional. Porém, a medida em que τ aumenta, o tempo computacional gasto nesses testes tornam a realização das simulações numéricas inviáveis. Por esta razão, a melhor estratégia é estimar o tempo de caminhada t_f necessário para o viajante com memória τ ser aprisionado em um atrator. Nos códigos numéricos utilizamos $t_f \sim 110\tau + 100$ para $d \leq 2$ e $N = 10000$. O estudo com $d > 2$ foi realizado com $N = 1000$ e $t_f \sim N/2$. Embora no modelo de alta dimensionalidade os tempos de transiente sejam maiores, foi constatado que tal tempo de caminhada é grande o suficiente para que se detecte um atrator em cada

trajetória. Na análise da trajetória, se um atrator não for encontrado, guarda-se esta informação em uma variável destinada a este fim e reinicializa-se a simulação com valores maiores de t_f .

A saída desta parte da simulação é a trajetória de um viajante.

Observando que parte um viajante de cada sítio, ao final de cada realização terão sido produzidas N trajetórias.

2.3.3 Tratamento das trajetórias

Nesta parte do programa as variáveis de entrada são: t_f , e a trajetória. Observe que se todas as trajetórias fossem guardadas em uma matriz, nesta etapa o número de viajante também seria uma variável de entrada. Para determinar o período p do atrator e o tempo de transiente t da trajetória monta-se uma rotina que estuda a caminhada do viajante tomado como variável de entrada a sua trajetória $traj(-\tau : t_f)$.

Lembrando que:

1. o número de passos dados antes do turista ser aprisionado em um atrator define um tempo de transiente t ;
2. o número de passos necessários para o turista completar uma volta no ciclo define o período p do atrator.

São realizadas as seguintes operações no estudo das trajetórias:

1. varia-se o período p de 2 até $t_f - 1$ para testar todos os possíveis valores de p ;
2. varia-se it de zero até $t_f - 2p$ para testar todos os possíveis tempos de transientes. O valor $-2p$ é subtraído de t_f porque cada atrator é gravado duas vezes. Caso contrário termina-se a execução do programa numérico com um código de erro:

- (a) para cada valor de p e it calcula-se a soma do valor absoluto da diferença entre dois sítios deslocados de p (ver a Figura 2.5):

$$S = \sum_{j=0}^{p-1} | traj(it + j) - traj(it + j + p) | ; \quad (2.5)$$

Figura 2.5: Exemplo de uma medida de tempo de transiente e de período de atrator. O período $p = 3$ do atrator e o transiente de tempo t são identificados no instante em que o valor absoluto da diferença sítio a sítio, de dois ciclos consecutivos da mesma trajetória é nula.

- (b) se $S = 0$, conclui-se que um atrator de período p foi encontrado e que o tempo de transiente foi de $t = it - 1$ ²;
3. verifica-se se a trajetória considerada é a primeira de um dado mapeamento;
 4. em caso afirmativo, armazena-se os sítios do atrator encontrado duas vezes na variável $A(p, k)$ (matriz de atratores), com $k = 1, 2, \dots, 2 * p$. Observe que não é possível determinar a priori o número de atratores que serão encontrados na paisagem. Verificamos que ao dimensionar a matriz de atratores é preciso fazer $k \leq 3N$;
 5. o número de vezes que os atratores de período p foram encontrados é incrementado $at(p) = at(p) + 1$;
 6. se não for o primeiro viajante, verifica-se se o atrator é diferente dos demais já gravados em $A(p, k)$.

²se $t = 0$ a trajetória se inicia no atrator.

- (a) a grandeza que mede a diferença de todas as combinações possíveis dos sítios do atrator considerado, e o atrator encontrado é definida como:

$$T(\tilde{p}, na) = \sum_{i=0}^{p-1} | A(p, 2p \cdot na + \tilde{p} + i) - traj(it + p + i) |, \quad (2.6)$$

onde $na = 0, 1, \dots, at(p)$, $i = 1, \dots, p - 1$ e $\tilde{p} = 1, 2, \dots, p$. Como os sítios são representados por números $s = 1, \dots, N$, este procedimento permite testar todas as combinações possíveis de atrator (veja a Figura 2.6);

Figura 2.6: Exemplo de como determinar se os atratores encontrados são diferentes. Nesse caso a matriz de atratores tem os atratores de período $p = 2$ constituído dos sítios (6 100) e (3 10). Foi encontrado outro atrator de período $p = 2$ formado pelos sítios (100 5). Testa-se todas as combinações possíveis dos sítios pertencentes ao atrator encontrado com os sítios do primeiro atrator gravado na matriz de atratores. Depois dá-se um salto proporcional $2p \cdot at(p) + 1$ e repete-se os testes com o atrator seguinte. Neste exemplo, verifica-se que realmente o ciclo (100 5) é um ciclo novo

- (b) se $T(\tilde{p}, na) \neq 0$ o atrator é diferente. Depois de um salto proporcional a $2at(p) + 1$ armazena-se duas vezes a sequência de sítios pertencentes ao atrator, na matriz $A(p, k)$;
- (c) incrementa-se a variável dos atratores diferentes $at(p) = at(p) + 1$;
- (d) para cada sítio pertencente ao atrator coloca-se o valor 1 na matriz de sítios $cities(p, k)$, para $k = 1, 2, \dots, N$;
- (e) se $T(\tilde{p}, na) = 0$ significa que o atrator já foi contado. Então não se incrementa as variáveis $at(p)$ e nem se grava o atrator em $A(p, k)$;
- (f) incrementa-se a variável que guarda informação sobre as bacias de atração $XP(p) = XP(p) + 1$;

- (g) incrementa-se a variável $XT(t)$ que guarda informação sobre os tempos de transientes $XT(t) = XT(t) + 1$;
7. após cada trajetória conta-se o número de cidades que pertencem ao atrator de período p e acumula-se em um vetor $S0(p)$;
8. depois da análise da trajetória deste viajante deixa-se os resultados guardados nas variáveis mencionadas anteriormente e envia-se um novo viajante $j = 1, \dots, N$;

2.3.4 Caracterização estatística das trajetórias

As grandezas medidas na análise estatística da trajetória de todos os viajantes que partiram na realização anterior são as variáveis de entrada para o cálculo dos primeiros e segundos momentos com relação aos mapeamentos.

Cálculo em tempo real do primeiro e segundo momento de uma distribuição de probabilidade

Uma grandeza a ser observada para estimar a confiabilidade de um resultado numérico é a variância. Mas para economizar tempo computacional é importante que os cálculos do primeiro e do segundo momento sejam efetuados em tempo real de simulação. Com este tipo de tratamento estatístico consegue-se trabalhar com sistemas da ordem de $N \sim 10^4$. Isto permite que os cálculos do desvio padrão (ou da variância) sejam obtidos usando vetores.

A média de cada uma das grandezas descritas anteriormente, que vamos denominar genericamente de $X(p)$, na realização i é obtida da média na realização $i - 1$, ponderada pelas $i - 1$ realizações precedentes. Depois da análise estatística da trajetória de todos os N viajantes enviados em cada mapeamento, segue que:

$$MX(p) = \frac{(i-1)MX(p) + X(p)}{i}. \quad (2.7)$$

Este mesmo procedimento é utilizado para o cálculo dos quadrados de $X(p)$

$$VMX(p) = \frac{(i-1)VMX(p) + X(p)^2}{i}. \quad (2.8)$$

Calcula-se a variância como:

$$S_R^2 = \left(\frac{N_R}{N_R - 1} \right)^2 [VMX(p) - MX(p)^2] , \quad (2.9)$$

e o desvio padrão é a raiz quadrada positiva da variância. Para grandes valores de N_R ($N_R \gg 1$) pode-se desprezar o préfator $N_R/(N_R - 1)$. Esses cálculos foram usados como parâmetro para avaliar os resultados numéricos. Mas, não aparecem nos resultados numéricos referentes ao modelo determinístico porque a barra de erro é menor que o símbolo, ou, em certos casos, para não carregar demais a figura.

Observe que as variáveis usadas para acumular a média referente aos diferentes mapeamentos só devem ser reinicializadas após as N_R realizações, enquanto que aquelas que acumulam a média da análise das trajetórias dos N viajantes enviados em cada mapeamento devem ser reinicializadas (com zero) em cada nova realização. Já o vetor que guarda a informação da trajetória dos viajantes deve ser reinicializado (zerado) antes de cada vez que um viajante parte.

A dinâmica iterativa da caminhada do turista sempre leva à trajetórias que são compostas inicialmente de um tempo de transiente e um atrator no final (ciclo de período p), onde o viajante é aprisionado. Se $\tau = 0$ (sem memória) o viajante vai simplesmente para o sítio vizinho mais próximo até encontrar dois sítios que sejam reciprocamente os mais próximos e entrar em um atrator de período 2.

Usando-se a média de todas as realizações, para uma dada memória, o modelo pode ser caracterizado pelas seguintes distribuições estatísticas:

1. a probabilidade $P_\tau(t)$ do caminhante dar t passos de tempo antes de entrar no atrator é estimada usando-se a razão entre o número total de tempos de transiente com t passos de tempo $TXT(t)$ e o número total de trajetórias nas N_R realizações:

$$P_\tau(t) = \frac{TXT(t)}{N_R n_v}, \quad (2.10)$$

onde $TXT(t)/N_R$ é a média dos tempos de transiente e n_v é o número de viajantes (ou trajetórias) por realização;

2. a densidade de atratores $D_\tau(p)$ é estimada usando-se a razão entre o número total de atratores diferentes de período p e o número total de sítios considerados nas N_R realizações:

$$D_\tau(p) = \frac{TAT(p)}{N_R N}, \quad (2.11)$$

onde $TAT(p)/N_R$ é definida como a média do número total de atratores diferentes de período p , encontrados na paisagem, em N_R realizações;

3. a probabilidade de distribuição de atratores de período p na paisagem, é a razão entre a densidade de atratores de período p e a densidade total $(\sum_{p=0}^{p=t_f} D_\tau(p))$ de atratores:

$$P_\tau(p) = \frac{D_\tau(p)}{\sum_{p=2}^{p=t_f} D_\tau(p)}, \quad (2.12)$$

onde t_f é o período máximo dos atratores que o viajante pode encontrar em sua trajetória, supondo que o tempo de caminhada seja suficiente para ele entrar em um atrator de período p ;

4. a probabilidade $P_A(p)$, de um caminhante pertencer a bacia de atração dos atratores de período p , é estimada usando-se a razão entre o número total de caminhantes $TXP(p)$ que são atraídos para o atrator de período p e o número total de caminhantes nas N_R realizações:

$$P_A(p) = \frac{TXP(p)}{N_R n_v}, \quad (2.13)$$

onde $TXP(p)/N_R$ é a média do número total de caminhantes atraídos para os atratores de período p ;

5. o tamanho médio das bacias de atração $S(p)$ de período p é definido como sendo a razão entre $P_A(p)$ e $D_\tau(p)$,

$$S(p) = \frac{P_A(p)}{D_\tau(p)}; \quad (2.14)$$

6. a densidade média de sítios situados em atratores de período p , é definida como a razão entre o número total de sítios $TS0(p)$ que pertencem a atratores de período p , em N_R realizações, e o número total de sítios na paisagem:

$$\langle d_S(p) \rangle = \frac{TS0(p)}{N_R N}, \quad (2.15)$$

onde $TS0(p)/N_R$ é a média do número total de sítios pertencentes aos atratores de período p , encontrado nas N_R realizações; A densidade média de sítios em atratores de período p , dividida pela densidade de atratores de período p , nos permite calcular o número médio de sítios por atrator de período p :

$$n(p) = \frac{\langle d_S(p) \rangle}{D_\tau(p)}. \quad (2.16)$$

Observe que a variável p nestas definições refere-se a um conjunto de atratores diferentes, com mesmo período.

2.4 Resultados analíticos para $\tau = 0$

Os resultados obtidos por Clark e Evans (21) e (22), cujo trabalho posteriormente foi corrigido por (23) e generalizado por Cox (24) podem ser utilizados na determinação da densidade de atratores. A ligação entre estes dois problemas foi colocada em evidência por Sebastian Risau-Gusman (25). Clark e Evans definiram como vizinhos reflexivos de ordem 1 os sítios que são mutuamente os mais próximos e determinaram analiticamente a probabilidade p_{rn} de se ter vizinhos reflexivos em processos de Poisson. A fração de vizinhos reflexivos mais próximos de uma população de N pontos, em todas as dimensões, corresponde a densidade de atrator de período 2 no problema do turista. Aumentando a dimensionalidade, a proporção de vizinhos reflexivos (casais) diminui e P_{rn} converge para $1/2$ quando $d \rightarrow \infty$. Portanto, a dinâmica do modelo de caminhada do turista com $\tau = 0$ leva à atratores idênticos a estes casais, de tal modo que aplicando a fórmula de Cox (24), a densidade de ciclos p , $D_\tau(p; d)$ é dada por:

$$D_0(2; d) = \frac{P_{rn}(d)}{2} = \frac{1}{2(1 + p_d)}, \quad (2.17)$$

onde

$$p_d = \frac{2\Gamma(d/2 + 1)}{\sqrt{\pi}\Gamma[(d+1)/2]} \int_0^{1/2} (1-x^2)^{(d-1)/2} dx, \quad (2.18)$$

Sendo que $P_{rn}(d)$ é a proporção de vizinhos reflexivos e $\Gamma(z)$ a função gama:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt. \quad (2.19)$$

A função gama $\Gamma(z)$ é usada para expressar fatoriais de números não inteiros com $Re(z) > 0$. A equação 2.18 pode ser trabalhada para obtermos uma função padrão observando que:

$$\frac{\Gamma(d/2 + 1)}{\sqrt{\pi}\Gamma[(d+1)/2]} = \frac{1}{B[1/2, (d+1)/2]}, \quad (2.20)$$

onde a função beta é definida como:

$$B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (2.21)$$

para $Re(a) > 0$ e $Re(b) > 0$.

Mudando as variáveis na Equação 2.18 ($t = x^2$ e $dt = 2xdx$) encontramos que:

$$\int_0^{1/2} (1-x^2)^{(d-1)/2} dx = \frac{1}{2} \int_0^{1/2} t^{-1/2} (1-t)^{(d-1)/2} dt, \quad (2.22)$$

e comparando a Equação 2.22 com a 2.21 concluímos que:

$$a - 1 = -1/2 \rightarrow a = 1/2,$$

além disso,

$$(d-1)/2 = b-1 \rightarrow b = (d+1)/2.$$

Substituindo a e b na Equação 2.18, obtemos:

$$p_d = I_{1/4}\left(\frac{1}{2}, \frac{d+1}{2}\right). \quad (2.23)$$

onde $I_{1/4}$ é a função beta incompleta que é definida como segue

$$I_x = \frac{1}{B(a, b)} \int_0^x t^{a-1} (1-t)^{b-1} dt. \quad (2.24)$$

Desta maneira conseguimos expressar a probabilidade p_d , e consequentemente de ciclos 2, em termos de uma função conhecida, a função beta incompleta.

d	p_d	$D_0(2, d)$
0	$1/3$	$3/8$
1	$1/2$	$1/3$
2	$(2\pi + 3^{3/2})/(6\pi)$	$3\pi/(2\pi + 3^{3/2})$
3	$11/16$	$8/27$
$\vdots$	$\vdots$	$\vdots$
∞	1	$1/4$

Tabela 2.1: Alguns valores de p_d e $D_0(2, d)$ em função de d .

Alguns valores de p_d e as respectivas densidade de vizinhos reflexivos $D_0(2, d)$ são mostrados a seguir em função da dimensionalidade do espaço na Tabela 2.1. Observa-se a convergência para $D_0(2, \infty) = 1/4$.

Nesse caso simples, os atratores podem ser identificados com propriedades geométricas (vizinhos reflexivos). Para $\tau = N - 2$, as trajetórias são totalmente repulsivas e temos o “algoritmo de Greedy” do caixeiro viajante (TSP) (16). Os casos de interesse são as situações de memórias intermediárias. Por exemplo, se $\tau = 1$, o último sítio visitado na trajetória não pode ser visitado novamente e somente atratores de período p com $p \geq 3$ podem existir. A relação entre os possíveis períodos p para um valor genérico de memória τ é:

$$p \geq \tau + 2, \quad (2.25)$$

pois o ciclo mínimo é composto pelo sítio atual, pelos τ últimos sítios (proibidos) e pelo próximo sítio a ser visitado.

2.5 Conclusão

Estamos propondo um modelo simples de caminhadas determinísticas em meios aleatórios. A regra determinística que orienta a trajetória do viajante é similar a outros problemas de otimização (algoritmo de greedy) e apesar de sua simplicidade, associada ao efeito da memória, produz um comportamento muito rico que está longe de ser trivial. O problema do turista introduz uma nova classe de problemas de otimização local onde o caminhante vai de sítio em sítio,

respeitando uma janela de exclusão (memória), ou seja, um “limited self-avoiding walk”.

Nesse Capítulo foram introduzidas as grandezas que estudamos para a caracterização de uma caminhada determinística. Estas grandezas foram descritas conceitualmente e na apresentação do algoritmo do programa foi enfatizado os aspectos gerais da rotina de análise de dados, de modo que o estudo pode ser realizado em uma dimensão d arbitrária.

Também descrevemos neste capítulo como as propriedades geométricas de conjuntos de pontos (“vizinhos reflexivos”) se relacionam com o caso simples sem memória ($\tau = 0$).

Capítulo 3

Sistemas unidimensionais

Neste capítulo vamos apresentar um estudo analítico para o número médio de sítios por atrator em $d = 1$. A comparação dos resultados desta análise teórica com os resultados numéricos nos fornece uma boa indicação da confiabilidade do programa. Mostraremos também que em $d = 1$ todas as distribuições de probabilidades discutidas no Capítulo 2 têm uma forte dependência com a memória τ considerada. Nos capítulos 4 e 5 mostraremos que este não é o caso quando $d > 1$.

3.1 Estudo analítico

O resultado analítico apresentado no Capítulo 2 mostra que para $\tau = 0$, a densidade de atratores em uma paisagem unidimensional vale $D_0(2, 1) = 1/3$. Este valor é importante na validação dos códigos numéricos desenvolvido.

A seguir, um cálculo analítico é apresentado para quando $\tau > 0$. Este resultado ajuda a explicar alguns comportamentos das distribuições de probabilidades obtidas numericamente.

3.1.1 Número médio de sítios por atrator

O período p do atrator depende da memória τ e do número de pontos intermediários n_I entre os dois extremos. A memória determina o período mínimo dos atratores, mas não o máximo. Os atratores aparecem devido a dinâmica introduzida e restrições geométricas podem determinar a existência ou não de alguns períodos de atratores.

Os atratores mais simples para um dado valor de τ possuem períodos $p_{min} = \tau + 2$ e o número mínimo de sítios $n_{min} = \tau + 2$, como ilustrado nas Figuras 3.1a e 3.1b. Chamamos tais estruturas de blocos base.

Figura 3.1: Exemplos dos diferentes tipos de atratores em $d = 1$: a) um bloco-base para $\tau = 1$; b) um bloco-base para $\tau = 2$; c) um ciclo permitido para $\tau = 1$ em que $n_I = 0$ e $p = 8$; d) um ciclo constituído de dois blocos-base e $n_I = 1$ sítio intermediário para $\tau = 1$ e $p = 10$; e) um exemplo de ciclo ímpar ($p = 11$) para $\tau = 2$ que é possível por causa de um laço interno (*loop*).

Quando o número de sítios que compõem o atrator n é muito grande, restrições geométricas impõem que os ciclos mais comuns sejam constituídos de dois blocos-base (um em cada extremidade do atrator) conectados por n_I sítios intermediários como mostra as Figuras 3.1c e 3.1d. Em um atrator com n sítios existem $n_I = n - 2n_{min}$ pontos intermediários. Estes sítios intermediários contribuem para o período total com $p_I = 2n_I + 2$ passos (se $n_I = 0$, a junção dos dois blocos-base contribui com dois passos). Portanto, o período total é $p = 2p_{min} + p_I = 2(\tau + 2) + 2[n - 2(\tau + 2)] = 2(n - \tau - 1)$, que leva a seguinte equação:

$$n(p) = \frac{p}{2} + \tau + 1, \quad (3.1)$$

onde $n(p)$ é o número médio de sítios por ciclo. Esta relação vale para ciclos com $n \geq 2n_{min} = 2(\tau + 2)$, porque somente estes ciclos podem incorporar dois

blocos-base independentes.

Se $\tau = 1$, esta é a única maneira possível de construir ciclos em $d = 1$. Isso implica na proibição por restrições geométricas de ciclos ímpares maiores que 3. Verifica-se também que o ciclo-6 é proibido pelo fato que não é possível construir um atrator com dois blocos base que possuam um ponto em comum.

Para $\tau > 1$, é possível construir ciclos ímpares pelo uso de *laços* internos (*loops*) e em condições iniciais muito específicas (um exemplo com $\tau = 2$ é dado na Figura 3.1e).

Do mesmo modo que a distribuição de probabilidade $P_A(p)$, a probabilidade de distribuição de atratores na paisagem $P_\tau(p)$ está fortemente correlacionada com a memória do caminhante e não pode ser descrita por uma expressão analítica simples. Dos resultados numéricos podemos deduzir que o comportamento da parte intermediária e posterior ($p \gg 1$) também pode ser descrito pela equação 3.7.

O resultado numérico do número médio de sítios que pertencem a atratores de período p pode ser comparado com o resultado analítico obtido na seção 3.1. Esta comparação é apresentada na Figura 3.2, onde vemos uma concordância relativamente boa, que melhora à medida que aumentamos τ e conseqüentemente p .

3.2 Resultados

Consideremos os resultados numéricos de caminhadas determinísticas sobre um conjunto de $N = 10000$ pontos ao longo de um número de realizações $N_R = 1000$ em uma paisagem unidimensional. O estudo foi realizado somente no limite de $\tau/N \ll 1$. Os sítios (“pontos”) foram espalhados aleatoriamente em R^1 seguindo uma distribuição uniforme com condições periódicas de contorno. Um viajante (turista), em cada passo de tempo, vai para o sítio vizinho mais próximo que não tenha sido visitado nos últimos τ passos. Considerando todas as condições iniciais possíveis, de cada sítio da rede parte um viajante que descreve uma trajetória compostas inicialmente de um tempo de transiente t e de um atrator de período p no final.

Figura 3.2: Número médio de sítios por atrator $\langle n(p) \rangle$: a) $\tau = 1$ (triângulo); b) $\tau = 6$ (retângulos). As linhas mostram a previsão teórica de cada τ .

3.2.1 Efeito de tamanho finito para $\tau = 0$

Um estudo do efeito de tamanho finito foi realizado, variando o número de sítios N e o número de realizações diferentes N_R de modo que o número de trajetórias $N.N_R$ permanecesse constante: $N = 100$ e $N_R = 6400$; $N = 200$ e $N_R = 3200$; $N = 400$ e $N_R = 1600$; $N = 800$ e $N_R = 800$; $N = 1600$ e $N_R = 400$; $N = 3200$ e $N_R = 200$; $N = 6400$ e $N_R = 100$. As medidas foram realizadas usando condições periódicas de contorno.

As distribuições de tempos de transiente, para diferentes valores de N foram superpostas, como ilustra a Figura 3.3. Não observamos diferenças significativas ($> 10^{-3}$) que possam ser atribuídas a efeitos de tamanho finito.

Além disso, pode-se comparar a densidade de ciclos de período 2 para $\tau = 0$ com o resultado analítico apresentados na seção 2.4 [$D_0(2) = 1/3$]. Esta comparação é ilustrada na Figura 3.4. Percebe-se que o valor numérico de $D_0(2)$ se aproxima de $1/3$ com N^{-1} . Neste caso, $N = 10000$ fornece um resultado com precisão de três casas decimais.

Embora os efeitos de tamanho finito tenham sido minimizados com o uso

Figura 3.3: Distribuição de tempos de transiente medidos com $\tau = 0$ em $d = 1$, para $N = 100, 200, 400, 800, 1600, 3200$ e 6400 . O número de realizações foi escolhido de forma que variando N o número de trajetórias seja $N.N_R = 640000$.

de condições periódicas de contorno, observamos que permanece um resíduo da ordem de N^{-1} .

3.2.2 Distribuição de tempos de transiente

Uma questão natural é se no limite de pouca memória ($\tau/N \ll 1$) existe um valor característico que produza uma transição de fase, por exemplo a emergência de um estado de transiente com um comportamento do tipo lei de potência. A distribuição de probabilidade dos tempos de transiente não sugere esta possibilidade porque é do tipo exponencial (Figura 3.5),

$$P_\tau(t) \propto \exp\left(\frac{-t}{\xi(\tau)}\right). \quad (3.2)$$

Figura 3.4: O efeito de tamanho finito na densidade de atratores $D_0(2)$ em $d = 1$, com $\tau = 0$. A densidade de atratores converge com N^{-1} para o valor previsto teoricamente.

O decaimento exponencial pode ser entendido como segue. A distribuição acumulada dos tempos de transiente define a probabilidade de um viajante não ser aprisionado em um atrator antes de t passos de tempo (independente do período p do atrator). Seja q a probabilidade do viajante não ser aprisionado em algum atrator e $\tilde{p} = 1 - q$ a probabilidade do viajante entrar no ciclo. Considerando o regime de transientes grandes em que $\tilde{p}$ é pequeno e tem o mesmo valor ao longo de uma trajetória, a probabilidade do caminhante dar t passos e só então entrar em algum atrator é:

$$P(t) = \tilde{p}q^t = \tilde{p}(1 - \tilde{p})^t, \quad (3.3)$$

como $\tilde{p}$ é pequeno, temos então:

$$P(t) \approx \tilde{p}e^{-\tilde{p}t}, \quad (3.4)$$

ou seja,

$$\tilde{p} = 1/\xi(\tau), \quad (3.5)$$

Figura 3.5: A distribuição dos tempos de transiente $P_\tau(t)$ para $d = 1$ decai exponencialmente. Da esquerda para a direita: $\tau = 0, 1, 2, 3, 4, 5$ e 6 , com $N = 10000$ e $N_R = 1000$.

Desta maneira vemos que a probabilidade de um viajante ficar preso em um atrator é inversamente proporcional ao “tempo” de decaimento $\xi(\tau)$ (número de passos).

Encontramos que $\xi(\tau)$ depende fortemente de τ . Para os resultados numéricos, verificamos que a melhor expressão para o tempo de decaimento é:

$$\xi(\tau) = c\tau \exp(\gamma\tau) . \quad (3.6)$$

Este comportamento observado para $d = 1$ é ilustrado na Figura 3.6. O fato de $\xi(\tau)$ depender exponencialmente de τ limitou o nosso estudo para valores de $\tau \leq 6$.

Figura 3.6: Tempo de decaimento característico $\xi(\tau)$, dos tempos de transiente, para valores de memória $\tau = 0, 1, 2, 3, 4, 5$ e 6 .

3.2.3 Bacias de atração

A probabilidade $P_A(p)$ de um viajante que parte de um sítio escolhido aleatoriamente, pertencer a bacia de atração dos atratores de período p apresenta dois tipos de comportamentos distintos que são exemplificados na Figura 3.7. Estas distribuições estão indicando que a probabilidade do caminhante pertencer a bacia de atração dos atratores de período p cai muito rapidamente para os ciclos no intervalo $\tau + 1 < p \leq 2(\tau + 1)$ e são seguidos de três ciclos proibidos. Observe também que em $\tau = 1$ os ciclos ímpares com $p \geq 2(\tau + 1)$ são proibidos. Isto sugere o aparecimento de um comportamento diferente, com o aumento da memória.

Existe uma evidência de que o comportamento inicial de $P_A(p)$ cai exponencialmente conforme mostra a Figura 3.8. Mas no limite de memória con-

Figura 3.7: Probabilidade de um viajante que parte de um sítio escolhido aleatoriamente pertencer a bacia de atração dos atratores de período p . Para $\tau = 1$ (triângulos), $\tau = 3$ (circunferência) e $\tau = 6$ (quadrados).

siderado não é possível determinar uma expressão analítica para a parte inicial.

A probabilidade $P_A(p)$ depende fortemente da memória do caminhante e acreditamos que não seja possível obter analiticamente uma expressão para descrever o comportamento observado numericamente. Dos resultados numéricos podemos observar que o comportamento da parte intermediária e posterior ($p \gg 1$) da distribuição de probabilidade é do tipo:

$$P_A(p) = C'(\tau) p^{\alpha'(\tau)} \exp\left(-\frac{p}{p_0(\tau)}\right), \quad (3.7)$$

onde o expoente α' apresentam uma dependência muito forte da memória τ .

Figura 3.8: Ampliação da parte inicial da distribuição de probabilidade em escala mono-log. Para $\tau = 1$ (triângulos), $\tau = 3$ (circunferência) e $\tau = 6$ (quadrados).

O tamanho médio das bacias de atração em $d = 1$ (ver seção 2.2), depende da memória conforme ilustrado na Figura 3.9.

Figura 3.9: Tamanho médio das bacias de atração para $\tau = 1, 3$ e 6 (de baixo para cima). Cada valor de memória tem um comportamento diferente.

3.2.4 Densidade de atratores

A densidade de atratores de período p , $D_\tau(p)$, possui dois comportamentos distintos que dependem da memória τ e, como ilustra a Figura 3.10, são de difícil caracterização.

Figura 3.10: Exemplos de densidade de atratores de período p em $d = 1$. Da esquerda para a direita $\tau = 1, 3$ e 6 .

Este resultado fornece uma boa intuição de como a memória τ afeta a densidade total de atratores (ver seção 2.2). A Figura 3.11 mostra que a densidade total de atratores $D(\tau)$ decai exponencialmente com a memória.

Este decaimento exponencial é uma característica de $d = 1$ e nós con-

Figura 3.11: A densidade total de atratores para $d = 1$ decai exponencialmente com a memória. Os resultados apresentados referem-se a $\tau = 1, 2, 3, 4, 5$ e 6 .

jeturamos que é devido a probabilidade de ocorrência de barreiras $P(d_{i+1} > d_i + d_{i-1} + \dots + d_{i-\tau})$, onde d_i é a distância viajada no i -ésimo passo que aprisiona o caminhante em ciclos.

Em sistemas unidimensionais observamos que o produto da grandeza $D(\tau)$ pelos tempos de decaimento dos tempos de transiente $\xi(\tau)$ cresce com uma lei de potência em função de τ (veja a Figura 3.12).

Figura 3.12: Produto da densidade total de atratores $D(\tau)$ pelos tempos de decaimento dos tempos transiente $\xi(\tau)$.

Os termos exponenciais se cancelam. Isto indica que a probabilidade de ocorrência de barreira $\tilde{p} = 1/\xi$ em primeira aproximação é proporcional a $D(p)$.

Portanto, $\gamma = \gamma'$, e permanece apenas um resíduo de lei de potência,

$$\begin{aligned}\xi(\tau)D(\tau) &= c\tau \exp(\gamma\tau) \exp(-\gamma'\tau) \\ &= c\tau .\end{aligned}\tag{3.8}$$

3.2.5 Distribuição de atratores

Do mesmo modo que em $P_A(p)$, também não é possível caracterizar a parte inicial da distribuição de probabilidade $P_\tau(p)$ (ver a Figura 3.13). Uma característica interessante da distribuição de atratores unidimensionais é que além da memória proibir os períodos com $p < \tau + 2$, os ciclos permitidos com $(\tau + 1) < p \leq 2(\tau + 1)$ apresentam um decaimento muito rápido e são seguidos de três ciclos proibidos.

Figura 3.13: A probabilidade de distribuição de atratores de período p em $d = 1$ para $\tau = 1$ (triângulo), 3 (circunferência) e 6 (retângulo).

A probabilidade de distribuição de atratores na paisagem $P_\tau(p)$ está fortemente correlacionada com a memória do caminhante e não pode ser descrita por uma expressão analítica simples. Dos resultados numéricos podemos deduzir que o comportamento da parte intermediária e posterior ($p \gg 1$) também pode ser descrito pela Equação 3.7.

3.3 Conclusão

Os resultados numéricos apresentados para a distribuição de atratores encontrados pelo turista com $\tau = 0$ foram comparados com os resultados analíticos para validar o programa e mostraram excelente concordância. Além disso, este estudo do caso unidimensional nos permitiu identificarmos outras características surpreendentes e específicas a estas distribuições de probabilidade. Estes resultados diferem dos de dimensionalidade mais alta porque a distribuição de atratores apresenta dois comportamentos distintos e dependem da memória do viajante, sendo um decaimento muito rápido para atratores compostos de um bloco-base e outro muito mais lento para os que têm dois blocos-bases. Isto deve-se a restrições geométricas que, como foi mostrado na seção 3.1 em alguns casos, a existência de blocos-base proíbe alguns atratores, como por exemplo a inibição dos ciclos ímpares para $\tau = 1$.

Capítulo 4

Sistemas bidimensionais

Considera-se as mesmas grandezas introduzidas no Capítulo 2 e estuda-se caminhadas determinísticas em meios aleatórios bidimensionais. Mostraremos que existem diferenças qualitativas entre o caso $d = 1$ e caso $d = 2$. Já para dimensões maiores, o comportamento é similar ao caso $d = 2$ (2).

Como motivação específica para estudarmos detalhadamente o problema de caminhadas determinísticas em duas dimensões, lembramos que os roedores apresentam dois tipos de comportamentos: um exploratório (curiosidade) e outro defensivo. Nos referimos a aversão de roedores a locais abertos e preferência por lugares onde seus bigodes podem tocar superfícies verticais ou objetos, que aparentemente lhes proporcionam uma sensação de proteção (26). Os sítios (cidades) do modelo poderiam representar esses lugares protegidos.

É provável que, para agentes biológicos (e robôs inspirados biologicamente) muitas vezes seja mais importante minimizar a distância viajada entre dois locais seguros do que otimizar alguma função custo global.

Um modelo determinístico também poderia descrever, por exemplo, comportamento migratório ou nômade de seres humanos, elefantes e outros animais com memória espacial bem desenvolvida. Ciclos podem ser relacionados às rotas migratórias estáveis em ambientes com recursos localizados, como por exemplo, ilhas oceânicas, oásis e nascentes de água (27). A necessidade de otimização local emerge devido ao pequeno desenvolvimento das capacidades sensoriais e cognitivas dos animais em geral. Longas distâncias podem também implicar em haver um custo adicional não proporcional à distância percorrida: viajantes (e animais) necessitam de local seguro para permanecerem durante o período de descanso, o

que determina a distância máxima que pode ser viajada em cada intervalo tempo.

4.1 Características da paisagem

Embora as propriedades geométricas possam afetar o período dos atratores, eles somente são produzidos devido à regra dinâmica introduzida no modelo. Considere como exemplo um caminhante com memória $\tau = 1$ que partindo de um sítio A encontra um atrator $ABCD$ (ver a Figura 4.1). Embora o sítio E seja escolhido para o agrupamento (desde que $BE < AB$), nunca será visitado porque $BC < BE$ e $CD < CE$. Porém se o caminhante parte do sítio C , ele entra em um atrator de período 3 que inclui o sítio E (CBE).

Figura 4.1: Exemplo de superposição de atratores para $d = 2$ e $\tau = 1$. Partindo de um sítio A , obtemos um atrator de período 4 ($ABCD$), mas partindo de C encontramos um atrator de período 3 (CBE).

Esta degenerescência e superposição de atratores pode ser entendida observando-se que as trajetórias estão no espaço de configurações e não no espaço de fase. No espaço de fase, para $\tau = 1$ as coordenadas são a posição atual, a posição anterior $X(t), Y(t), X(t-1), Y(t-1)$ e as trajetórias nunca se cruzam.

Para $\tau > 0$, ao se enviar viajantes de diferentes posições iniciais percebe-se a coexistência de atratores de diferentes períodos. Fazendo isto para diferentes valores de memória τ observamos o aparecimento, desenvolvimento ou fusão dos atratores. A Figura 4.2 mostra este comportamento para uma paisagem com $N = 400$ sítios, onde foram usadas condições de contorno livres. Um aspecto interessante é que, fixando a paisagem, os atratores produzidos pelos valores

ímpares de τ são correlacionados. O mesmo ocorre para valores pares de τ (Figura 4.3), mas não existe uma correlação aparente entre os atratores oriundos de valores de memória ímpares e pares. A razão desse fato ainda nos é desconhecida.

Figura 4.2: Exemplos do aparecimento de diferentes atratores para $d = 2$ e $N = 400$, com $\tau = 1, 3, 5$ e 9 . A mesma paisagem foi usada para cada um destes gráficos, deste modo pode-se observar o crescimento de diferentes atratores com τ . Foram utilizadas condições livres de contorno. Os pontos que não pertencem a nenhum ciclo não foram mostrados.

Os padrões e correlações podem ser observados somente quando as medidas são realizadas em uma única configuração de sítios e são uma forte indicação que os atratores maiores são produto da fusão de atratores com períodos menores. O processo de fusão, por sua vez, sugere que o fato da memória ser par ou ímpar afeta os períodos dos atratores. Quando a média de alguma grandeza for considerada em N_R realizações, os padrões e correlações observados em uma realização individual desaparecem, para dar lugar a uma distribuição uniforme de atratores em todo o espaço.

Figura 4.3: Exemplos do aparecimento de diferentes atratores para $d = 2$ e $N = 400$, com $\tau = 2, 4, 6$ e 10 . A mesma paisagem da Figura 4.2 foi usada para cada um dos gráficos.

4.2 Efeito de tamanho finito

Nesses grafos aleatórios, ao privilegiar os sítios mais próximos por causa da regra determinística, foi feita uma orientação das conexões usando as distâncias como pesos para as ligações entre dois sítios. Os únicos critérios adotados pelo caminhante para dar um passo de tempo são a memória e a distância do sítio que ele se encontra dos sítios vizinhos. Quando $\tau = 0$ em uma população distribuída aleatoriamente e a distância entre dois destes sítios é pequena comparada com a distância relativa aos outros sítios da rede, existe uma ocorrência óbvia de formação de pares. No espaço Euclidiano, a relação de vizinhos reflexivos (casais) para esta população é bem definida (ver o Capítulo 2).

4.2.1 Estudo de tamanho finito para $\tau = 0$

Podemos utilizar os cálculos teóricos de Clark e Evans (21; 22) e Cox (24) para validar os resultados de nossas simulações no caso sem memória ($\tau = 0$). O resultado deste estudo analítico do número de casais no espaço Euclidiano é equivalente ao valor numérico da densidade de atratores de período 2 $[D_0(2)]$ no modelo apresentado. Portanto, em analogia com os resultados analíticos para duas dimensões (24), pode-se esperar que:

$$D_0(2) = \frac{P_{rn}(2)}{2} = \frac{1}{2} \frac{6\pi}{8\pi + 3^{3/2}} \approx \frac{0.6215}{2} \approx 0.31075. \quad (4.1)$$

Quando $\tau = 0$, as medidas numéricas de $D_0(2)$ em função de N^{-1} validam o programa utilizado na simulação. Mostramos que quando N cresce, $D_0(2)$ converge para o valor teórico e que o erro da estimativa é inversamente proporcional ao tamanho do sistema (ver a Figura 4.4).

Um comportamento similar também é observado em sistemas unidimensionais. Em $d = 2$ os resultados numéricos da Figura 4.4 podem ser descritos por uma lei de potência do tipo:

$$D_0(2) = 0,3107 + 1,0412N^{-1}. \quad (4.2)$$

4.3 Resultados para $\tau > 0$

A partir do estudo apresentado na Figura 4.4 concluímos que a imprecisão devida ao efeito de tamanho finito em $d = 2$ também é da ordem de N^{-1} . Diante disso, realizamos a análise estatística do resultado destas simulações das caminhadas determinísticas considerando um conjunto de $N = 10000$ pontos ao longo de $N_R = 1000$ realizações. As coordenadas dos sítios (“pontos”) em R^2 foram espalhados aleatoriamente e uniformemente, utilizando-se condições

Figura 4.4: $D_0(2)$ em função de N^{-1} . Varia-se N e o número realizações N_R mantendo constante o número de caminhos ($N.N_R = 640000$). Os valores utilizados foram: $N = 100$ e $N_R = 6400$; $N = 200$ e $N_R = 3200$; $N = 400$ e $N_R = 1600$; $N = 800$ e $N_R = 800$; $N = 1600$ e $N_R = 400$; $N = 3200$ e $N_R = 200$; $N = 6400$ e $N_R = 100$.

periódicas de contorno. Construiu-se a matriz de distâncias e a matriz de vizinhos de acordo com o algoritmo apresentado no Capítulo 2. O estudo foi realizado somente no limite de $\tau/N \ll 1$.

4.3.1 Distribuição de tempos de transiente

Como no caso unidimensional, a distribuição de tempos de transiente não sugere a existência de um estado de transiente com comportamento de lei de potência no intervalo de memória considerado $\tau \in [0, 10]$, porque $P_\tau(t)$ também é exponencial,

$$P_\tau(t) \propto \exp\left(\frac{-t}{\xi(\tau)}\right). \quad (4.3)$$

Este comportamento é mostrado na Figura 4.5.

Figura 4.5: Distribuição de tempos de transiente $P_\tau(t)$ em $d = 2$ para $\tau = 0, 2, 3, 5, 7$ e 10 (da esquerda para a direita). Os tempos de transiente em $d = 2$ são menores que em $d = 1$.

Estas distribuições de tempos de transiente diferem do caso unidimensional (ver a Figura 3.6) no tipo de tempo de decaimento característico $\xi(\tau)$ que em $d = 2$ tem uma dependência em τ (ver a Figura 4.6),

$$\xi(\tau) \propto \tau^\delta, \quad (4.4)$$

com $\delta = 1$. Ou seja, no caso bidimensional os tempos de transientes crescem muito mais lentamente com a memória τ , permitindo um estudo para valores maiores ($\tau \approx 10$).

4.3.2 Comportamento das bacias de atração

A bacia de atração de um atrator é formada por todos os sítios que fazem o caminhante convergir para um mesmo atrator. A probabilidade de um viajante

Figura 4.6: Tempo de decaimento característico $\xi(\tau)$ em função da memória τ . Diferentemente do caso unidimensional $\xi(\tau)$ tem um comportamento de lei de potência em função de τ .

que parte de um sítio escolhido aleatoriamente pertencer a bacia de atração dos atratores de período p é mostrada na Figura 4.7. Esta distribuição em $d = 2$ sugere a existência de um comportamento de lei de potência com um corte do tipo exponencial para a cauda da distribuição.

Os resultados numéricos obtidos variando τ de 1 até 10 mostraram que, para $p > 2(\tau + 2)$, a probabilidade $P_A(p)$ do viajante pertencer a bacia de atração dos atratores de período p é dada por uma expressão analítica do tipo:

$$P_A(p) = C'(\tau)p^{-\alpha'} \exp\left(-\frac{p}{p_0(\tau)}\right), \quad (4.5)$$

em que C' é uma constante. O período mínimo $p_{min} = \tau + 2$ se desloca com a memória (Figura 4.7). Existe aqui uma diferença significativa em relação a $d = 1$, que é o fato do expoente α ser independente da memória em $d = 2$.

Figura 4.7: Probabilidade de um viajante pertencer a bacia de atração dos atratores de período p para $\tau = 1$ (triângulo), $\tau = 2$ (estrelas), $\tau = 4$ (circunferências), $\tau = 10$ (retângulos).

Chama-se a atenção para o fato de que observamos $p_0(\tau) \approx \tau$. Este comportamento sugere a possibilidade de $P_A(p)$ escalar com a variável reduzida p/τ e que existe uma função que colapsa todos os dados em uma única curva (28).

Em $d = 2$ também aparece um decaimento inicial mais rápido dos ciclos de períodos típicos entre $(\tau + 2) < p \leq 2(\tau + 2)$. A Figura 4.8 mostra que este decaimento claramente não é exponencial e que torna-se gradativamente menos acentuado com o aumento da memória.

Um aspecto interessante é que a probabilidade do caminhante pertencer a bacia de atração de um atrator de período mínimo $p_{min} = \tau + 2$, para $\tau \gg 1$

Figura 4.8: Comportamento inicial das distribuições de bacias de atração dos atratores de período p para $\tau = 1$ (triângulo), $\tau = 2$ (estrelas), $\tau = 4$ (circunferências), $\tau = 10$ (retângulos).

tende a ser inversamente proporcional a este período (ver a Figura 4.9):

$$P_A(p_{min}) = \frac{1,7}{\tau + 2} . \quad (4.6)$$

Figura 4.9: Comportamento dos períodos mínimos para $\tau \gg 1$. A linha sólida corresponde à Equação 4.6 e os quadrados correspondem aos valores numéricos.

Colapso de dados

É interessante observar o caráter robusto destas curvas através do escalonamento dos dados referentes a diferentes valores de τ . O colapso de dados de

todas estas curvas dependentes de τ pode ser feito usando uma função do tipo:

$$F(p/\tau) = \tau P_A(p) , \quad (4.7)$$

como ilustrado na Figura 4.10. Uma lei similar de escala não se aplica ao caso $d = 1$ porque o expoente α depende fortemente da memória τ .

Figura 4.10: Colapso dos dados da distribuição de probabilidade $P_A(p)$ usando a Equação 4.7, para $\tau = 1$ (triângulo), $\tau = 2$ (estrelas), $\tau = 4$ (circunferências), $\tau = 10$ (retângulos).

4.3.3 Bacias de atração de período p

Os resultados numéricos mostram que o tamanho médio das bacias de atração $\langle S(p) \rangle$ obedece uma lei do tipo:

$$\langle S(p) \rangle \propto C(\tau) p^\mu, \quad (4.8)$$

onde $\mu = 0,75 \pm 0,05$ (ver a Figura 4.11). Este resultado sugere que a forma geométrica do atrator tem um efeito sobre sua bacia de atração.

Figura 4.11: Tamanho médio das bacias de atração de período p para $\tau = 1$ (estrela), $\tau = 2$ (circunferências), $\tau = 4$ (retângulos), $\tau = 10$ (triângulo).

4.3.4 Densidade de atratores de período p

O número de atratores é proporcional a N , ou seja, a densidade $D_\tau(p) = (\text{número de atratores de período } p)/N$ é constante. Exemplos dessas densidades são mostrados na Figura 4.12. Como no caso $d = 1$, observamos um bico para períodos pequenos (ciclos especiais) e uma calda na forma de lei de potência com um corte exponencial.

$$D_\tau(p) \propto p^{-\alpha(\tau)} \exp\left(-\frac{p}{p_0(\tau)}\right), \quad (4.9)$$

onde α depende da memória.

Figura 4.12: Exemplos de densidade de atratores $D_\tau(p)$: $\tau = 1$ (quadrado), $\tau = 4$ (estrelas) e $\tau = 10$ (circunferências).

Note que, para $\tau = 1$, os ciclos ímpares são mais raros e que o efeito é similar ao encontrado no caso $d = 1$.

4.3.5 Densidade total de atratores

A densidade total de atratores $D(\tau) = \sum_{p=2}^{p=t_f} D_\tau(p)$ decresce com uma lei de potência do tipo $D(\tau) \sim \tau^{-\delta'}$, onde $\delta' = 1,0$, como ilustrado na Figura 4.13).

Figura 4.13: Densidade total de atratores em função da memória τ . Existe um comportamento de lei de potência, que é função de τ , para $d = 2$.

O tempo de transiente médio é proporcional a distância média entre os atratores (que são inversamente proporcionais a densidade de atratores). Isso significa que:

$$D(\tau)\xi(\tau) \propto \tau^{-\delta'} \tau^\delta \approx C, \quad (4.10)$$

onde $C \sim 2,3$ deve ser constante, isto é, $\delta \approx \delta'$, como ilustrado na Figura 4.14.

Figura 4.14: O produto da constante de decaimento pela densidade total de atratores de períodos p mostra que as contribuições de lei de potência se cancelam.

4.3.6 Probabilidade de distribuição de atratores

A distribuição de atratores em $d = 2$ escala com uma lei de potência, que depende do número de vértices no grafo. A densidade de atratores de período p dividida pela densidade total, nos fornece a distribuição de probabilidade de atratores $P(p) = D_\tau(p)/D_\tau$, que como em $P_A(p)$, para $p > 2(\tau + 2)$ pode ser descrita por uma lei de potência limitada que escala com o inverso da memória no intervalo $\tau/N \ll 1$:

$$P_\tau(p) = C(\tau)p^{-\alpha} \exp\left(\frac{-p}{p_0(\tau)}\right), \quad (4.11)$$

onde $C(\tau)$ é uma constante e $\alpha \sim 2$, isto é mostrado na Figura 4.15.

Figura 4.15: Probabilidade de distribuição de atratores de período p para $\tau = 1$ (triângulo), $\tau = 2$ (estrela), $\tau = 4$ (circunferências), $\tau = 10$ (retângulos).

Expoente da distribuição de atratores

A distribuição de atratores $P_\tau(p)$ depende de τ e apresenta um expoente α , que ao contrário do observado em $d = 1$, pode-se mostrar que independe da memória (do mesmo modo que na seção 4.3.2). Os dados podem ser colapsados por uma função similar a que descreve $P_A(p)$ (ver equação 4.7),

$$G(p/\tau) = \tau P_\tau(p), \quad (4.12)$$

como ilustrado na Figura 4.16. Observe que o colapso se dá para $p > 2(\tau + 1)$, ou seja, para a cauda da distribuição.

Figura 4.16: A Equação 4.12 representa o colapso de todas as distribuições $P_\tau(p)$, como ilustrado em função de p/τ para $\tau = 1$ (triângulos), 2 (estrelas), 4 (circunferências) e 10 (retângulos).

4.3.7 Número médio de sítios por atrator

O número médio de sítios por atrator tem um comportamento de lei de potência que depende do período do conjunto de atratores diferentes, para p grande, como mostra a Figura 4.17. Quando $\tau = 1$ o número médio de sítios por atrator também pode ser descrito por uma equação muito parecida com a utilizada em $d = 1$ (ver Equação 3.1):

$$\langle n(p) \rangle \propto p/2. \quad (4.13)$$

Isso sugere que muitos atratores têm estrutura similar ao caso $d = 1$, o que explicaria a ausência relativa de ciclos ímpares (menor probabilidade em $\tau = 1$), como mostra a Figura 4.15.

Figura 4.17: Número médio de sítios por atrator: $\tau = 1$ (estrelas), $\tau = 2$ (triângulo), $\tau = 4$ (cruz), $\tau = 6$ (retângulos), $\tau = 10$ (circunferências).

4.4 Conclusão

Nesse capítulo mostramos que a dimensionalidade do sistema pode afetar todas as distribuições de probabilidades inerentes da dinâmica determinística em meios aleatórios. Identificamos também algumas características gerais deste efeito:

1. embora as distribuições de tempos de transiente $P_\tau(t)$ mantenham o comportamento exponencial como observado em $d = 1$, os tempos de decaimento ξ que cresciam exponencialmente com a memória τ , em $d = 2$ apresentaram um crescimento muito mais lento (lei de potência com expoente 1) com a memória;
2. as distribuições de probabilidade $P_A(p)$ e $P_\tau(p)$, que em $d = 1$ tem uma forte dependência com a memória τ , em $d = 2$ foram colapsadas em uma

- única curva independente de τ que pode ser descrita por uma lei de potência com um corte exponencial, como nas Equações 4.7 e 4.12, respectivamente;
3. a distribuição de tamanho médio das bacias de atração $\langle S(p) \rangle$ tem uma dependência do tipo lei de potência com o período $p^{0,75 \pm 0,05}$ (Equação 4.8);
 4. a densidade total de atratores $D(\tau)$ que em $d = 1$ decai exponencialmente, em $d = 2$ apresentou um comportamento $D(\tau) \propto 1/\tau$.

Capítulo 5

Efeito da dimensionalidade

A lei que rege os tempos de transiente e as distribuições de períodos de atratores (lei de potência com um corte exponencial) depende da dimensão d considerada. À medida que a dimensionalidade do sistema cresce, os nossos resultados se aproximam dos resultados do modelo de ligações aleatórias simétricas. Nesse último modelo, resultados analíticos indicam que o corte exponencial na densidade de probabilidade da distribuição de períodos de atratores depende do tamanho do sistema e que desaparece quando $N \rightarrow \infty$. A distribuição torna-se, então, uma lei de potência pura.

Para efeito de comparação inicialmente mostraremos resultados para o modelo com ligações aleatórias assimétricas, que é equivalente ao mapa aleatório estudado por Derrida e Flyvbjerg (13). Veremos que a principal diferença é que o número de atratores é proporcional a $\log(N)$ no caso assimétrico e proporcional a N no caso simétrico.

5.1 Espaço de ligações aleatórias assimétricas

As redes com ligações aleatórias assimétricas podem ser geradas considerando as distâncias entre dois sítios como variáveis aleatórias independentes, distribuídas uniformemente no intervalo entre $[0, 1]$. Este procedimento se caracteriza pela introdução, no grafo do turista, de ligações que não estão correlacionadas. Isto significa que a distância do sítio i ao j , não é necessariamente a distância do j ao i . Todos os caminhos são como avenidas de mão única, onde o caminho de ida geralmente difere do caminho de volta ($D(i, j) \neq D(j, i)$).

Como os valores $D_{i,j}$ são aleatórios, a matriz de vizinhos também o é. Por

exemplo, para $\tau = 0$, se o sítio no instante t é $C(t) = C_i$, então no passo seguinte devemos ter $C(t + 1) = V(i, 1)$. Ou seja, usamos um mapeamento aleatório:

$$F(C_i) = V(i, 1), \quad (5.1)$$

o que corresponde exatamente ao mapeamento aleatório de Derrida e Flyvbjerg.

Considerando os casos com $\tau > 0$ a situação é similar. Dado o sítio $C(t) = C_i$ e os τ últimos sítios visitados, fica definido um mapeamento para algum sítio escolhido aleatoriamente. Um caminhante seguindo a mesma regra determinística do turista em uma configuração aleatória de sítios com conexões assimétricas, também descreve uma trajetória composta de um transiente e um atrator no final. No caso do mapeamento aleatório, o grafo correspondente a todas as conexões de um sítio i com seus $\tau + 1$ vizinhos mais próximos (a tabela $V(i, j)$), perde sua interpretação espacial e torna-se aleatória.

5.1.1 Tempos de transiente

Para um valor de N , a distribuição de tempos de transiente é bem descrita por uma expressão do tipo:

$$P_\tau(t) = A(\tau) \exp \left[- \left(\frac{t}{\xi(N)} \right)^\gamma \right]. \quad (5.2)$$

Esta expressão é análoga a forma integrada nos períodos encontrada por Bastolla e Parisi, em um estudo analítico realizado para o mapa aleatório (29).

A análise do efeito de tamanho finito mostrou que os tempos característicos de transiente escalam com o tamanho do sistema na forma $\xi(N) = \xi_0 N^{1/2}$. Assim, todas as distribuições podem ser colapsados em uma mesma curva (ver Figura 5.1) por uma função do tipo:

$$F(t/N^{1/2}) = N^{1/2} P_\tau(t). \quad (5.3)$$

É interessante enfatizar que o escalonamento das curvas com N é o mesmo obtido na referencia (29) considerando um modelo de ligações assimétricas (mapeamento aleatório de Derrida-Flyvberg).

Figura 5.1: Os tempos de transiente de uma caminhada determinística sobre uma configuração aleatória de sítios com conexões assimétricas escalam com $N^{1/2}$. Mostra-se $N = 50, 100, 200, 400, 800$ e 1600 para $\tau = 0$ (memória nula). A linha cheia é a curva analítica dada pela Equação 5.2 com $\xi_0 = 0,94$; $A = 1, 2$ e $\gamma = 1, 6$.

5.1.2 Efeito de tamanho finito com $\tau = 0$

A caminhada do turista no modelo de ligações aleatórias assimétricas não é equivalente a uma caminhada no espaço Euclidiano no limite de $d \rightarrow \infty$. A Figura 5.2 mostra que a densidade total de atratores tende para zero quando N vai para infinito. Como foi explicado no seção 2.4 o esperado é que $D(\tau) = 0,25$ para $d \rightarrow \infty$. Além disso, no espaço Euclidiano são esperados somente ciclos-2 para $\tau = 0$ e no mapeamento aleatório assimétrico aparecem ciclos maiores que 2.

Para podermos comparar os resultados numéricos realizamos o estudo de tamanho finito variando o número de sítios na paisagem e o número de realizações. Iniciou-se a análise com $N = 50$ sítios, para $N_R = 20000$ realizações, e variou-se N até 1600 em conjunto com $N_R = 1000000/N$ mantendo constante o número

Figura 5.2: A densidade total de atratores $D(\tau) = \sum_{p=2}^{p=t_f} D_\tau(p)$ na caminhada do turista com $\tau = 0$, no modelo de ligações aleatórias assimétricas, tende para zero quando o tamanho do sistema aumenta. São mostrados $N = 50, 100, 200, 400, 800$ e 1600 .

total de trajetórias.

Uma característica particular da distribuição de atratores nesta caminhada com ligações aleatórias assimétricas é o aparecimento de ciclos maiores que 2, mesmo para $\tau = 0$. A densidade de ciclos tende para zero conforme N aumenta. Este comportamento pode ser descrito aproximadamente pela seguinte expressão analítica:

$$D_\tau(p) = C(p)N^{-1}, \quad (5.4)$$

onde $C(p)$ é uma constante que depende do período considerado. A Figura 5.3 apresenta o comportamento lei de potência das densidades de atratores de período p (conjunto de atratores), em função de N^{-1} .

A densidade total de atratores para $\tau = 0$, no modelo de ligações aleatórias

Figura 5.3: Densidade total de período p de atratores em função do inverso do tamanho do sistema para $N = 50, 100, 200, 400, 800$ e 1600 . Os resultados são referentes $\tau = 0$, no modelo de ligações aleatórias assimétricas. De cima para baixo o período dos ciclos mostrados são: $p = 2, 3, 4, 6, 8$ e 16 .

assimétricas, obedece o resultado de Harris (14),

$$N \cdot D_0(p) \propto \log(N). \quad (5.5)$$

Este comportamento é mostrado na Figura 5.4.

Figura 5.4: No modelo de ligações aleatórias assimétricas com $\tau = 0$ o número de atratores (produto da densidade total de atratores pelo número de sítios), é proporcional a $\log(N)$. São mostrados $N = 50, 100, 200, 400, 800$ e 1600 .

5.1.3 Distribuição de períodos de atratores

A Figura 5.5 mostra o resultado da análise de tamanho finito para $\tau = 0$.

A expressão que descreve o comportamento da distribuição de atratores para os

diferentes valores de N é:

$$P_\tau(p) = C(\tau)p^{-\alpha} \exp \left[- \left(\frac{p}{p_0(N)} \right)^\beta \right], \quad (5.6)$$

onde p é o período.

A natureza do colapso de dados mostrou que a distribuição de atratores com $\tau = 0$ no mapeamento aleatório assimétrico também escalam com $N^{1/2}$, ou seja, $p_0(N) = \tilde{p}_0 N^{1/2}$. Portanto as curvas referentes aos diferentes valores de N podem ser colapsadas em uma mesma curva pela função;

$$G_\tau(p/N^{1/2}) = N^{1/2} P_\tau(p), \quad (5.7)$$

que é independente do valor de N ($N > 50$).

Figura 5.5: Colapso das distribuições de atratores, referentes a caminhada do turista com $\tau = 0$, para $N = 100, 200, 400, 800$ e 1600 . A linha cheia é a curva analítica dada pela Equação 5.7 ajustada com $\tilde{p}_0 = 1,78$; $C = 0,35 \pm 0,05$; $\alpha = 1$ e $\beta = 2$.

O efeito do aumento da memória

Verificamos que tanto para $\tau = 0$ quanto para $\tau = 1$ (Figura 5.5 e 5.6), a curva escalonada pode ser descrita pela Equação 5.6. O colapso de dados foi realizado por uma função analítica descrita pela Equação 5.7. Na tabela 5.1 compara-se os valores dos ajustes para $\tau = 0$ e $\tau = 1$. Observa-se que a memória não muda os expoentes α e β , mas apenas $C(\tau)$ e $\tilde{p}_0(\tau)$.

Figura 5.6: Colapso das distribuições de atratores para $\tau = 1$ referentes a $N = 100, 200, 400, 800$ e 1600 no modelo de ligações assimétricas. A linha cheia é a curva analítica ajustada com $\tilde{p}_0 = 2,40$; $C = 0,35 \pm 0,05$; $\alpha = 1$ e $\beta = 2$.

5.2 Espaço de ligações aleatórias simétricas

Uma boa aproximação do caso limite em que $d \rightarrow \infty$, para o problema do turista, pode ser feita distribuindo as distâncias aleatoriamente e ao mesmo tempo mantendo a correlação entre o caminho de ida e o de volta ($D_{ij} = D_{ji}$). Isto significa dizer que para simular esta situação a matriz de distâncias deve ser simétrica. O efeito das variáveis aleatórias na dinâmica da caminhada determinística é anular a correlação espacial entre os sítios, estabelecida pelas distâncias Euclidianas. A correlação entre três ou mais sítios é completamente anulada, mas a correlação entre $D(i, j)$ e $D(j, i)$ permanece.

τ	$\tilde{p}_0$	C	α	β
0	1,79	0,35	1	2
1	2,40	0,40	1	2

Tabela 5.1: Expoente das distribuições de atratores com $\tau = 0$ e $\tau = 1$ para o modelo de ligações aleatórias assimétricas.

Estudo de tamanho finito para $\tau = 0$

Iniciou-se a análise com $N = 50$ sítios e $N_R = 20000$ realizações. Variou-se N até 1600 em conjunto com $N_R = 1000000/N$ mantendo constante o número total de trajetórias. As “distâncias” $D(i, j)$ foram obtidas de uma distribuição uniforme no intervalo $[0, 1]$ mantendo a distância de ida igual a de volta. Como discutido no Capítulo 2.4 a densidade de ciclos-2 é igual a densidade de vizinhos reflexivos.

A aproximação de ligações aleatórias simétricas para a caminhada do turista produz uma densidade de atratores $D_0(2)$ que tende para o valor previsto teoricamente por (21), (22) e (24) para $d \rightarrow \infty$ e $N \rightarrow \infty$. A Figura 5.7 mostra o resultado deste estudo e sugere que para $N \geq 1000$ esta aproximação bastante boa.

Figura 5.7: Efeito de tamanho finito sobre a densidade de atratores, na caminhada do turista com $\tau = 0$, no grafo construído a partir do modelo de ligações aleatórias simétricas. São mostrados $N = 50, 100, 200, 400, 800$ e 1600 .

5.2.1 Tempos de transiente com um passo de memória

De modo análogo ao descrito na subseção 5.1.1 deste capítulo, as distribuições de probabilidade de tempos de transientes, no modelo de ligações aleatórias simétricas, também são descritas pela Equação 5.2 e podem ser colapsadas em uma única curva pela Equação 5.3. Como não temos conhecimento de nenhum estudo analítico que possa prever este comportamento, vamos apresentar apenas os resultados numéricos com os respectivos valores usados para o ajuste das curvas pela mesma expressão analítica usada no modelo assimétrico (ver Figura 5.8).

Figura 5.8: Na aproximação de ligações aleatórias simétricas os tempos de transiente também escalam com $N^{1/2}$ como foi descrito na Equação 5.3, com $\xi = 1, 59$, $A = 0, 70$ e $\gamma = 1, 6$. Aqui estamos mostrando $N = 80, 160, 320, 400$ e 1600 com $\tau = 1$.

5.2.2 Distribuições de atratores

A maneira mais comum de se determinar numericamente a existência de um comportamento de lei de potência é verificando se a curvatura das distribuições de probabilidade se desloca, quando o tamanho do sistema aumenta. Os resultados numéricos do estudo de efeito tamanho finito mostraram que existe um deslocamento da distribuição de probabilidade $P_\tau(p)$ em função de N que indica

um comportamento em lei de potência para $P_\tau(p)$, no limite de $N \rightarrow \infty$ (ver a Figura 5.9).

Figura 5.9: Efeito de tamanho finito na distribuição de atratores para $\tau = 1$ no modelo de ligações aleatórias simétricas. Da esquerda para direita tem-se: $N = 10, 20, 40, 320$ e 1600 , respectivamente.

Observe que não foi possível realizar o escalonamento dos dados referentes a Figura 5.9 com $N^{1/2}$ como no caso das ligações aleatórias assimétricas mostrado na Figura 5.5. Possivelmente, uma das razões para isto deve-se ao fato da matriz de distâncias ser simétrica e somente metade das variáveis serem aleatórias, o que resulta em uma densidade de atratores muito maior que no caso simétrico.

5.3 Interpolação entre os dois modelos

Seja q a probabilidade de $D(i, j)$ ser independente de $D(j, i)$ e $1 - q$ a probabilidade de $D(i, j) = D(j, i)$. O procedimento para interpolar entre o modelo de ligações aleatórias simétrica e assimétricas consiste em inserir conexões assimétricas na matriz de vizinhos $V(i, k)$, construída a partir do modelo de ligações aleatórias simétrica, com probabilidade q . Variando-se $i = 1, \dots, N$ e $k = 1, \dots, \tau + 1$, com probabilidade q substitui-se cada conexão do vértice i , escolhendo aleatoriamente um sítio do grafo. A repetição deste procedimento para todas as k gerações de vizinhos do sítio i , resulta em paisagens constituídas de ligações simétricas e assimétricas. Quando $q = 0$ as conexões dos sítios são

simétrica e para $q = 1$ todas as conexões simétricas são substituídas por conexões assimétricas.

A Figura 5.10 mostra o efeito das ligações aleatórias assimétricas na densidade total de atratores $D(\tau) = \sum_{p=2}^{p=t_f} D_\tau(p)$. Observe que o comportamento tipo mapa aleatório só aparece quando $p \rightarrow 1$.

Figura 5.10: Densidade total de atratores em função da probabilidade de assimetria das conexões dos vértices do grafo, para $N = 1000$ e $N_R = 100$. São mostrados $\tau = 0$ e $\tau = 10$.

5.4 Efeito da dimensionalidade do espaço

Para analisarmos o efeito da dimensionalidade da paisagem na caminhada do turista escolhemos $\tau = 1$ porque nesse caso já aparece uma distribuição de atratores $P_\tau(p)$ não-trivial, ou seja, com períodos maiores que 2. Fixamos o número de pontos N em 1000 e N_R em 10000. Os sítios foram distribuídos uniformemente no espaço de $d = 1, 2, 4, 8, 16, 32, 64$ para o modelo de ligações aleatórias simétricas usando condições periódicas de contorno. Para o mapeamento aleatório, as distâncias entre os sítios foram distribuídas aleatoriamente.

Nesta seção vamos determinar numericamente as distribuições estatísticas da seção 2.2 do Capítulo 2 e mostrar como as distribuições de probabilidade são afetadas pela dimensionalidade, comparando com o caso em que a cami-

nhada é desenvolvida sobre uma configuração aleatória de ligações (simétricas e assimétricas).

5.4.1 Tempos de transiente

Os resultados numéricos mostram que as distribuições de tempos de transiente dependem da dimensão da paisagem. Estas probabilidades de tempo de transiente podem ser descritas pela mesma expressão analítica apresentada na subseção 5.1.1 (ver Equação 5.2). Observa-se que para todas as dimensões consideradas o comportamento é exponencial, mas agora γ depende da dimensão d . A Figura 5.11 mostra a convergência dos tempos de transiente $P_1(t)$ para o limite de dimensão infinita. À medida que se aumenta a dimensionalidade do sistema, o expoente do argumento da função exponencial varia no intervalo entre $\gamma = 1$ (baixa dimensionalidade) e $\gamma = 2$ (alta dimensionalidade).

Na Tabela 5.2 comparamos os resultados usados para o ajuste das distribuições de transiente no modelo de ligações aleatórias simétricas e assimétricas com a Equação 5.2. O que diferencia os resultados dos dois modelos é que na caminhada sobre uma configuração aleatória de sítios com conexões assimétricas a constante A é maior e ξ_0 é menor que na simétrica (ver a Tabela 5.2).

<i>modelo</i>	ξ_0	A	γ
<i>si</i>	65	0,02	1,6
<i>assi</i>	30	1,20	1,6

Tabela 5.2: Quando fazemos a comparação dos valores usados para ajustar os resultados numéricos referentes aos modelos de ligações aleatórias simétricas (si) e assimétricas (assi), observamos que o comportamento das distribuições de transiente é robusto.

Na Figura 5.12 mostramos a dependência do expoente γ , dos tempos de transientes, em função do inverso da dimensionalidade. Para $d \leq 4$ o expoente do argumento da exponencial se mantém praticamente constante.

5.4.2 Comportamento das bacias de atração

A probabilidade do caminhante partir de um sítio escolhido aleatoriamente e ser atraído para um atrator de período p no limite de $d \rightarrow \infty$, tanto em

Figura 5.11: Distribuições de tempos de transiente para caminhadas do turista com $\tau = 1$. Da esquerda para direita tem-se $d = 1, 2, 3, 4, 8, 16, 32$ e 64 . Observa-se a convergência no limite de $d \rightarrow \infty$, para o modelo de ligações aleatórias simétricas (circunferências vazias). Com a finalidade de comparação, os quadrados vazios correspondem ao modelo de ligações aleatórias assimétricas (mapeamento aleatório assimétrico), que não é um bom limite para $d \rightarrow \infty$.

um grafo com ligações aleatórias simétricas quanto com ligações aleatórias assimétricas pode ser descrita pela Equação 5.6 com $\alpha = 0$. A Figura 5.13 mostra que aumentando a dimensionalidade os resultados convergem para o modelo de ligações aleatórias simétricas.

5.4.3 Distribuição de atratores

A Figura 5.14 mostra a convergência lenta da probabilidade de distribuição de atratores $P_\tau(p)$ para o limite de $d \rightarrow \infty$. Embora tanto o caso de ligações simétricas como assimétricas apresentem os mesmos expoentes na Equação 5.6, vemos claramente uma convergência do caso $d \rightarrow \infty$ para o modelo de ligações aleatórias simétricas. A análise dos resultados numéricos com esta expressão indica que o comportamento é robusto, como mostra a Tabela 5.3.

Figura 5.12: Convergência do expoente da distribuição de atratores $P_{\tau=1}(t)$ para o limite de $d \rightarrow \infty$. Da direita para esquerda tem-se $d = 1, 2, 4, 8, 16, 32$ e 64 .

<i>modelo</i>	p_0	C	α	β
<i>si</i>	93,00	0.27	1	2
<i>assi</i>	53,50	0,35	1	2

Tabela 5.3: Comparação dos resultados numéricos das distribuições de atratores $P_\tau(p)$ referentes aos modelos de ligações simétricas (si) e assimétricas (assi). Para $\tau = 1$ observa-se que os expoentes são os mesmos para os dois modelos.

O resultado da análise dos resultados numéricos ilustrados pela Figura 5.15, utilizando a Equação 5.6, mostra como α varia com a dimensionalidade (ver Figura 5.15).

5.4.4 Número médio de sítios por atrator

Mesmo quando a correlação entre as distâncias é baixa (d grande) o número médio de sítios por atrator continua sendo proporcional ao período:

$$\langle n(p) \rangle \propto C p, \quad (5.8)$$

Figura 5.13: Probabilidade do caminhante pertencer a bacia de atração dos atratores de período p para $\tau = 1$. Da esquerda para direita tem-se: $d = 1$ (triângulo), $d = 2, 4, 8, 16$ e 32 . Observa-se uma convergência lenta para o modelo de ligações aleatórias simétricas (circunferências vazias). Os quadrados vazios correspondem ao modelo de ligações aleatórias assimétricas (mapeamento aleatório de Derrida-Flyvbjerg). As curvas analíticas são dadas pela Equação 5.6.

A constante C aparentemente não depende da dimensionalidade d . A análise de dados para $d = 1, 2, 4, 8, 16, 32$ e (ligações aleatórias) (ver a Figura 5.16) indica que a Equação 5.8 descreve corretamente $\langle n(p) \rangle$. Observa-se que a constante C tem uma variação pequena para valores intermediários de dimensionalidade.

5.5 Conclusão

O comportamento da distribuição de atratores para toda dimensão d pode ser descrito por uma lei de potência com um decaimento exponencial no final (ver equação 5.6) e este decaimento depende do tamanho do sistema N . Em

Figura 5.14: Probabilidade $P_\tau(p)$ de distribuição de atratores de período p para $\tau = 1$. Da esquerda para direita tem-se: $d = 1, 2, 4, 8, 16, 32$ e 64 . Observa-se a existência de uma convergência lenta para o modelo de ligações aleatórias simétricas (as circunferências vazias). Os quadrados vazios correspondem ao modelo de ligações aleatórias assimétricas (mapeamento aleatório de Derrida-Flyvbjerg).

contraste com o caso $d = 1$, os expoentes α' e α independem τ e um colapso de dados pode ser feito.

O principal resultado deste capítulo é a constatação que o mapeamento aleatório **não** fornece o limite correto $d \rightarrow \infty$ para as caminhadas do turista. O limite correto é dado pelo modelo de ligações aleatórias simétricas, que possui um número de atratores proporcional a N , em vez de $\log(N)$, como é o caso do mapeamento aleatório assimétrico.

Figura 5.15: Expoente da distribuição de atratores $P_\tau(p)$ em função de $1/d$ para $\tau = 1$. Da direita para esquerda tem-se $d = 2, 4, 8, 16, 32$ e 64 .

Figura 5.16: Número médio de sítios por atrator para $\tau = 1$ e $d = 1, 2, 4, 8, 16$ e 32 . As maiores flutuações são para dimensionalidades intermediárias. Os retângulos correspondem ao modelo de ligações aleatórias assimétrica e a circunferência ao modelo de ligações simétricas.

Capítulo 6

Aplicação a grafos de dicionários

A caminhada do turista é capaz de modelar uma variedade ampla de problemas que podem ser formulados em termos vértices de um grafos com vizinhos ordenados. As palavras em um dicionário estão relacionadas umas as outras por seu significado. Aqui nós mostramos quantitativamente através de caminhadas parcialmente auto-repulsivas (caminhada do turista) que a estrutura topológica dos grafos de dicionários podem ser embebidas no espaço Euclidiano de baixa dimensionalidade como evidenciado por Steyvers e Tenenbaum (30) e ao contrário do previsto pelo modelo de análise de semântica latente LSA (17; 18).

As máquinas de tradução atuais apresentam vários problemas para converter textos de um idioma para o outro, e vale ressaltar, mesmo para humanos esta não é uma tarefa muito fácil. Por exemplo, se traduzimos uma frase e depois a convertemos para o idioma original, freqüentemente observamos flutuações no sentido original. Como exemplo, consideramos a tradução iterativa, de uma linguagem para outra, de frases pequenas (em geral provérbios). Escolhemos um provérbio a partir do português e depois traduzimos para o inglês. Depois traduzimos novamente do inglês para o português. Este procedimento foi interrompido quando as traduções não mais se modificavam. O método leva a um transiente e a um ciclo dois no final, onde quase sempre as sentenças finais não fazem mais sentido em nenhum dos idiomas. Nós verificamos isto com os seguintes provérbios:

Água mole em pedra dura tanto bate até que fura.

Soft water in stone lasts so much it beats until that holes.

Água macia em pedra dura tanto bate até que fura.

Soft water in stone lasts so much it beats until that it holes.

Em casa de ferreiro, espeto de pau.

In blacksmith's house, wood spit.

Na casa de ferreiro, cuspe de madeira.

In blacksmith's house wood sputum.

Santo de casa não faz milagre.

House Saint doesn't make miracle.

Casa São não faça milagre.

House Saint doesn't make miracle.

Deus ajuda quem cedo madruga.

God helps who early gets up early.

Deus ajuda para cima que cedo adquirem cedo.

God helps upward that early acquire early.

Deus ajuda para cima aquele cedo adquira cedo.

God helps upward that I give in he acquires early.

Deus ajuda para cima que eu o cedo adquira cedo.

Em terra de cego quem tem um olho é rei.

In blind man earth who has an eye he/she is a king.

Em terra de homem de cortina que têm um he/she de olho está um rei.

In curtain man's earth that has an eye he/she this is a king.

Um fenômeno similar ocorre em dicionários analógicos eletrônicos ou enciclopédias. Partindo de uma palavra aleatória, se for estabelecido um processo iterativo indo para o sinônimo mais próximo (i.e., a primeira palavra da lista dada pelo dicionário), depois de um tempo de transiente, quase sempre converge-se para um ciclo de dois. A seguir apresentamos alguns exemplos obtidos a partir da

iteração da primeira palavra na lista de sinônimos do dicionário Microsoft Word 98:

1. Trajetória com transiente nulo.

(a) **link** → connection → **link**

2. Trajetória com tempo de transiente de dois passos.

(a) translation → conversion → **change** → alter → **change**

3. Exemplo de trajetórias diferentes

(a) dicionário do Reino Unido (Britânico).

constitution → charter → contract → **agreement** → accord →
agreement

(b) dicionário americano (USA)

constitution → company → corporation → business → commerce →
trade → deal → contract → **agreement** → accord → **agreement**

6.1 Teste dos dicionários

Para estabelecermos uma ligação do problema do turista com o problema dos dicionários e das máquinas de tradução, realizamos caminhadas sobre o grafo definido por um dicionário com memória $\tau = 0$ e $\tau = 1$.

Uma palavra é escolhida aleatoriamente a partir de um dicionário padrão. A palavra é colocada em um dicionário eletrônico (o dicionário de inglês da Microsoft 98). No dicionário usado, os sinônimos são organizados de acordo com sua frequência de uso. Para uma caminhada com $\tau = 0$, o sinônimo de uma palavra mais freqüente é usado como uma entrada e o primeiro vizinho (sinônimo usado mais freqüente) é usado em um processo iterativo. Este procedimento leva a um transiente e geralmente a um ciclo de período 2, onde dois sinônimos reciprocamente mais próximos são encontrados.

Depois de algumas reflexões e testes em dicionários eletrônicos padrões, encontramos que a distância média entre dois sinônimos cresce com $\log(N)$, onde N é o número de vértices (palavras) no grafo definido por um dicionário. Também observamos que os sinônimos têm alto coeficiente de agrupamento. Isto significa que a probabilidade de dois sinônimos de uma palavra estarem também conectados é grande e sugere que os programas de tradução e talvez os dicionários, em geral apresentam fenômeno de Mundo-Pequeno (1). Isso foi confirmado recentemente por Steyvers e Tenenbaum (30).

A proporção de vizinhos reflexivos deve ser estimada pela medida da fração de palavras que já pertencem a um ciclo dois (como mostrado anteriormente para a palavra **link**), ou seja, palavras com transiente zero. A partir de 1000 palavras ($\sim 1\%$ do dicionário), 272 delas levam a ciclos 2. Este valor $P_{rn} \sim 0,27 \pm 0,03$, com um nível de confiança de $\gamma = 0,95$, está muito distante do intervalo esperado $[1/2, 2/3]$ para pontos distribuídos aleatoriamente no espaço Euclidiano de dimensão d . Entretanto, observando a distribuição de transientes da Figura 6.1 está claro que este número foi subestimado. Uma das razões para isto é que uma grande fração de ciclos 2 são compostos de expressões com duas palavras, como por exemplo: “**punctual** $\rightarrow$ on time $\rightarrow$ **punctual**”. Mas, nossas caminhadas foram iniciadas sempre a partir de palavras únicas. Deste modo o ciclo: “**on time** $\rightarrow$ **punctual** $\rightarrow$ **on time**” nunca seria encontrado. A extrapolação do decaimento exponencial no transiente para $t = 0$ conduz a uma melhor estimativa de $P(t = 0)$.

Comparando os pontos da Figura 6.1 com uma distribuição analítica do tipo;

$$P(t) = c.e^{-ct}, \quad (6.1)$$

o resultado leva a $P(0) = P_{rn} = 0,618$ (circunferência aberta), que significa dizer que existe um número de vizinhos reflexivos da ordem do previsto pela fórmula do Cox (24) para baixas dimensões: $P_{rn}(2) = 0,62$ e $P_{rn}(2) = 0,59$

Alguns ciclos de períodos maiores foram encontrados nos experimentos

Figura 6.1: Distribuição de tempos de transiente para o dicionário de inglês americano (MS98 USA) com $\tau = 0$ e $N = 1000$ palavras iniciais.

com dicionários: $(5, 2 \pm 1, 4)\%$ de ciclos 3 e $(0, 5 \pm 0, 4)\%$ de ciclos 4, com um nível de confiança $\gamma = 95\%$. Isto significa que no dicionário examinado pode ocorrer que B seja vizinho mais próximo de A , C seja vizinho mais próximo de B e A seja vizinho mais próximo de C . Este tipo de relação não ocorre na geometria Euclidiana (devido a desigualdade triangular), sugerindo que os grafos de dicionários não são simétricos (nós não sabemos se esta é uma propriedade desejada, ou apenas um artefato dos dicionários examinados).

6.1.1 Caminhadas com $\tau = 1$

A caminhada do turista no dicionário eletrônico foi realizada por Gisele Maranghetti Lourenço que anotou toda a sequência de palavras, a partir de 1000 palavras escolhidas ao acaso, obtidas em dicionário analógico.

Para caminhadas em dicionários com $\tau = 1$, encontramos que a proba-

bilidade $P_A(p)$ de uma palavra pertencer a uma bacia de atração de período p é muito similar àquela observada em caminhadas no espaço Euclidiano (ver a Figura 6.2). Este resultado é espantoso e sugere que os grafos de palavras podem ser embebidos no espaço Euclidiano de baixa dimensionalidade (mesmo com $d = 2$), preservando a maior parte de suas relações de vizinhanças. Quando comparamos o comportamento dos dicionários para $\tau = 1$, observamos uma diferença significativa em relação a distribuição $P_A(p)$ para um grafo com uma distribuição de vizinhos aleatória (que corresponde a usar grafos com ligações aleatórias assimétricas) e mesmo com relação a $P_A(p)$ para dimensões maiores que 4. Existe uma boa concordância com a distribuição de probabilidade $P_A(p)$ no espaço Euclidiano de baixa dimensionalidade, mas esta similaridade não explica as ligações assimétricas. Portanto, o mais provável é a existência de efeito de mundo pequeno com baixa probabilidade de religação, que permita algumas ligações assimétricas e mantenha as propriedades Euclidianas localmente.

Figura 6.2: Comparação entre a fração de sítios pertencentes a ciclos de período p $P_A(p)$ formados com $\tau = 1$ em um grafo Euclidiano aleatório e a distribuição de atratores de um grafo formado por $N = 700$ palavras iniciais do dicionário de inglês da MS98 com $d = 2$ (círculos). O resultado para grafos com ligações aleatórias assimétricas também é mostrado (quadrado).

6.2 Conclusão

Mostramos que uma simples caminhada sobre grafos com a vizinhança ordenada produz o aparecimento de ciclos com propriedades estatísticas interessantes. O estudo desta dinâmica se torna interessante para a compreensão de grafos cuja estrutura global não seja acessível (como por exemplo, a base de dados dos dicionários usados), por ser feito localmente através de um processo iterativo.

Ciclos refletem as propriedades de agrupamentos da rede. Estes resultados indicam que as propriedades de agrupamentos dos dicionários são similares às aquelas geradas por um grafo Euclidiano espacialmente aleatório com $d = 2$ (ou $d = 3$). Caminhadas sem memória em grafos de dicionários ($\tau = 0$) detectam facilmente uma ausência de simetria no grafo. Desde que a proporção de relações assimétricas seja pequena, também é possível que elas sejam eliminadas pelo uso de melhor ordenamento dos vizinhos.

Nossos resultados, sugerem que grafos de sinônimos podem ser embebidos em espaços de baixas dimensão, em discordância com os modelos de análise de semântica latente LSA (17; 18) que assumem que $d > 100$.

Capítulo 7

Caminhadas tendenciosas

O estudo de caminhadas aleatórias em meios desordenados é um assunto bastante desenvolvido e tem sido usado para modelar uma grande variedade de sistemas. Em geral a teoria de processos estocásticos (Markovianos) é simples de tratar devido a sua ergodicidade. O estudo de caminhadas determinísticas é também muito interessante, mas apresenta dificuldades analíticas comuns à área de sistemas dinâmicos complexos e tem sido bem menos investigada (3; 31; 32; 33; 34).

Nos capítulos anteriores vimos que uma regra de otimização local da distância percorrida sempre leva ao aparecimento de ciclos. Tais ciclos prendem o caminhante e impedem a exploração da paisagem total. Se for interessante maximizar a exploração espacial o caminhante precisa usar algum mecanismo aleatório. Interesses conflitantes como estes, que combinam determinismo com estocasticidade, têm levado os pesquisadores a adotar modelos híbridos. Algoritmos mais sofisticados apresentam um mecanismo (como o gradiente descendente determinístico) de busca local, em conjunto com um mecanismo de busca global que em geral é estocástico. Um exemplo são os algoritmos genéticos (GA) (35) e de recozimento simulado [*simulated annealing*] (SA) (36).

No estudo de caminhadas parcialmente alto repulsivas dos capítulos anteriores vimos que, de acordo com o modelo determinístico proposto, somente a memória não é suficiente para recuperar a ergodicidade. Aqui nós consideramos um caso muito simples da caminhada do turista, sem memória, mas com um parâmetro estocástico extra, o análogo da temperatura T . Abaixando a temperatura $T \rightarrow 0$ é possível estudar a quebra da ergodicidade em determinada

temperatura característica.

Apresentamos resultados analíticos em $d = 1$ e resultados numéricos em $d = 1$ e 2 . Usando uma definição natural de “eficiência” da caminhada (número de sítios novos visitados dividido pela distância percorrida) também mostramos que a eficiência é máxima na borda da transição. Neste Capítulo considera-se um modelo híbrido de caminhadas em meios aleatórios com a introdução de um parâmetro externo, a temperatura. O objetivo é estudar numericamente a transição do regime não ergódico para o ergódico nas caminhadas. A introdução desta dinâmica produz um aumento significativo do tempo das simulações numéricas e a introdução da memória exige computadores com maior capacidade de processamento do que dispomos no momento. Por esta razão vamos tratar somente as situações em que o viajante não apresenta memória, ou seja, $\tau = 0$. A apresentação dos resultados da análise estatística da caminhada do turista mostra que em altas temperaturas o sistema se torna ergódico (o viajante pode percorrer todo o espaço), enquanto que em baixas temperaturas o viajante fica aprisionado localmente (regime não-ergódico). A transição entre estes dois regimes é análoga a uma transição vítrea (19).

7.1 Modelo estocástico da dinâmica do turista

O modelo é definido por uma paisagem composta de N pontos cujas coordenadas são distribuídas aleatoriamente e uniformemente em um espaço de d dimensões no volume de $[0, 1]^d$ e em seguida é aplicado condições periódicas de contorno. Estes pontos podem ser imaginados como “cidades” ou “locais seguros”. De cada sítio parte um viajante com um tempo de exploração t_f (tempo de vida do viajante) que se desloca do sítio i para o sítio j com probabilidade $P_\beta(i \rightarrow j)$.

A fim de que no limite de $T \rightarrow \infty$ a dinâmica determinística do turista seja recuperada, iremos considerar um viajante que vai de um sítio para outro com uma probabilidade que decai exponencialmente. A caminhada do turista pode ser generalizada com a seguinte regra: **vá do sítio i para o sítio j que**

não tenha sido visitado nos últimos τ passos, com uma probabilidade $P_\beta(i \rightarrow j)$.

Note que a distância média entre os sítios $\langle D(i, j) \rangle$ diminui conforme N aumenta na forma $\langle D(i, j) \rangle \propto 1/N^{1/d}$. Assim é natural usarmos distâncias escaladas $\overline{D}(i, j) = N^{1/d} D(i, j)$.

A probabilidade de transição do sítio i ao j é definida como:

$$P_\beta(i \rightarrow j) = \frac{e^{-\beta \overline{D}(i, j)}}{Z_i}, \quad (7.1)$$

onde

$$Z_i = \sum_{j(\neq i)=1}^N e^{-\beta \overline{D}(i, j)} \quad (7.2)$$

é o fator de normalização. Enfatiza-se que o fator de normalização é diferente para cada sítio da rede e $\beta = 1/T$ é o inverso da temperatura.

Para $\beta \rightarrow \infty$, caso determinístico, o caminhante obedece o comportamento descrito nos capítulos anteriores. Para $\beta \rightarrow 0$, a probabilidade do caminhante ir de um sítio para outro é simplesmente $1/(N - 1)$, i.e., a caminhada é totalmente aleatória. Em valores intermediários de β , para cada passo de tempo, o caminhante sem memória ($\tau = 0$) pode ser aprisionado em ciclos de período 2, de onde eventualmente escapa, indo para outro ciclo de período 2, permanecendo por um tempo e escapando novamente (eventualmente voltando para o ciclo precedente). O comportamento do caminhante muda drasticamente de um regime estendido (ergódico) para outro localizado (não-ergódico) em β_c . O regime estendido favorece a maximização da exploração do espaço enquanto que o outro regime favorece a minimização da distância viajada.

Nesse modelo, o caminhante é impedido de permanecer no mesmo sítio, i.e., $P_\beta(i \rightarrow i) = 0$, ou seja, é como se tivéssemos $\overline{D}(i, i) \rightarrow \infty$.

Com a finalidade de estudar qualitativamente o comportamento do nosso modelo, vamos iniciar a discussão de modo que se tenha baixa temperatura, ou seja, um valor alto de β , comparado com algum valor característico ainda a ser determinado.

Um viajante inicia a sua caminhada em um determinado sítio, indo preferencialmente para os sítios mais próximos. Este viajante será aprisionado em ciclos de período 2 quando encontrar dois sítios que sejam vizinhos mutuamente mais próximos. É conveniente denominarmos de “casais” os pares (a, b) de sítios mutuamente mais próximos. Cada casal é caracterizado pela distância $D(a, b)$ que separa esses dois sítios.

Para baixas temperaturas e longas caminhadas o processo de difusão do turista basicamente é controlado pelo tempo de residência nos casais.

7.2 O sistema a altas temperaturas

Considere a probabilidade de transição

$$P_\beta(i, j) = \frac{1}{Z_i} e^{-\beta \bar{D}(i, j)} , \quad (7.3)$$

onde

$$Z_i = \sum_{j(\neq i)=1}^N e^{-\beta \bar{D}(i, j)} . \quad (7.4)$$

No limite de $(\beta \rightarrow 0)$ temos

$$P_\beta(i, j) = \frac{1}{N-1} , \quad (7.5)$$

ou seja, todos os sítios são igualmente acessíveis. Conforme β aumenta, transições para os sítios mais próximos são favorecidas. O alcance desse efeito é controlado por β .

7.3 O sistema a baixas temperaturas

A análise que vamos iniciar aplica-se ao regime de baixa temperatura para se obter a estabilidade dos atratores. No regime de baixa temperatura $(\beta \rightarrow \infty)$, o caminhante é aprisionado em atratores de período 2, compostos de dois pontos a e b que são vizinhos mutuamente mais próximos (também chamados de “vizinhos reflexivos” ou de “casais”) (21), (22). A probabilidade de um ponto ser vizinho reflexivo de outro ponto (de pertencer a casais) já foi obtida analiticamente (seção 2.4) e vale:

$$P_{rn}(d) = \frac{1}{1 + p_d} , \quad (7.6)$$

onde,

$$p_d = I_{1/4} \left(\frac{1}{2}, \frac{d+1}{2} \right), \quad (7.7)$$

com $I_z(a, b)$ sendo a função beta incompleta.

Para calcular o tempo de residência nos casais (vizinhos reflexivos mais próximos), consideremos que a e b são os sítios de um casal e $D_0 = D_{ab}$ a distância entre eles.

7.3.1 Distribuição de distâncias entre vizinhos reflexivos

A expressão analítica para a função densidade de probabilidade de D_0 foi obtida por Sebastian Risau-Gusman (25). Ela é uma simples exponencial:

$$P(x_0) = e^{-x_0}, \quad (7.8)$$

em termos da variável:

$$x_0 = \frac{\pi^{d/2}(1+p_d)}{\Gamma(d/2+1)} D_0^d, \quad (7.9)$$

onde $\Gamma(z)$ é a função gama e p_d é dado pela Equação 7.7.

7.3.2 Tempos de residência nos casais

As probabilidades de transição do sítio a para o sítio b e vice-versa, são respectivamente:

$$p_a = e^{-\beta D_0} / Z_a \quad (7.10)$$

e

$$p_b = e^{-\beta D_0} / Z_b. \quad (7.11)$$

A probabilidade de escape do casal, i.e., de ocorrer uma transição do sítio a (b) para algum outro sítio $k \neq b$ (a), é:

$$q_a = 1 - p_a = \sum_{k(\neq a, b)=1}^N e^{-\beta D_{ak}} / Z_a \quad (7.12)$$

e

$$q_b = 1 - p_b = \sum_{k(\neq a, b)=1}^N e^{-\beta D_{bk}} / Z_b. \quad (7.13)$$

Suponha que, em $t = 0$, o viajante esteja no sítio a . A Tabela 7.1 mostra a probabilidade $P(t)$ do viajante permanecer no casal por t passos.

Passo de tempo t	Probabilidade	$P(t)$
1	q_a	q_a
2	$p_a q_b$	$p_a p_b q_b / p_b$
3	$p_a p_b q_a$	$p_a p_b q_a$
4	$p_a p_b p_a q_b$	$(p_a p_b)^2 q_b / p_b$
5	$p_a p_b p_a p_b q_a$	$(p_a p_b)^2 q_a$
6	$p_a p_b p_a p_b p_a q_b$	$(p_a p_b)^3 q_b / p_b$
7	$p_a p_b p_a p_b p_a p_b q_a$	$(p_a p_b)^3 q_a$
$\vdots$	$\vdots$	$\vdots$
t ímpar		$(p_a p_b)^{(t-1)/2} q_a$
t par		$(p_a p_b)^{t/2} q_b / p_b$

Tabela 7.1: Probabilidade do viajante permanecer no par de sítios a e b , esse exemplo inicia-se em a .

A probabilidade do viajante permanecer no casal depois de t passos de tempo, onde t é par vale:

$$P_e(t) = (p_a p_b)^{t/2} q_b / p_b , \quad (7.14)$$

e para valores de t ímpares:

$$P_o(t) = (p_a p_b)^{(t-1)/2} q_a . \quad (7.15)$$

Desde que o viajante possa partir do sítio a ou do sítio b , a probabilidade de que ele permaneça no ciclo é:

$$P_e(t) = (p_a p_b)^{t/2} \left(\frac{q_a}{p_a} + \frac{q_b}{p_b} \right) / 2 , \quad (7.16)$$

e

$$P_o(t) = (p_a p_b)^{(t-1)/2} (q_a + q_b) / 2 . \quad (7.17)$$

Chamamos a atenção para o fato de que $P(t)$ está normalizado:

$$\sum_{t=1}^{\infty} P(t) = \sum_{k=0}^{\infty} [P_o(2k+1) + P_e(2k+2)] = 1 . \quad (7.18)$$

7.3.3 Tempo de residência

Para um casal, o tempo de residência médio t_r é dado por:

$$\begin{aligned}
 t_r &= \sum_{t=1}^{\infty} P(t)t \\
 &= \sum_{k=0}^{\infty} [P_o(2k+1)(2k+1) + P_e(2k+2)(2k+2)] \\
 &= \frac{1 + (p_a + p_b)/2}{1 - p_a p_b} \\
 &= \frac{Z_a Z_b + e^{-\beta D_0} (Z_a + Z_b)/2}{Z_a Z_b - e^{-2\beta D_0}}. \tag{7.19}
 \end{aligned}$$

A partir de agora consideremos a situação em que $Z = Z_a = Z_b$. Como vimos anteriormente esta aproximação pode ser justificada no limite de temperaturas altas (aproximação de *annealing*). Nesse caso o tempo de residência vale:

$$t_r = \frac{1 + e^{-\beta D_0}/Z}{1 - (e^{-\beta D_0}/Z)^2} = \frac{1}{1 - e^{-\beta D_0}/Z} \tag{7.20}$$

No limite $\beta \rightarrow 0$, $Z = N - 1$ (ver Equação 7.2) substituindo os valores em $e^{\beta D_0}/Z = 1/(N - 1)$ temos que $t_r = (N - 1)/(N - 2)$ tende para a unidade (menor valor possível) no limite termodinâmico $N \gg 1$.

Para β arbitrário, $i = a$ ou b no fator de normalização Z_i , nós ordenamos os termos como segue:

$$Z_i = e^{-\beta D_0} + \sum_{k(\neq a, b)=1}^N e^{-\beta \overline{D}(i, k)}. \tag{7.21}$$

Seja $D_1 \equiv \min(D_{ak}, D_{bk'})$ com k e k' sendo o segundo vizinho mais próximo de a e b respectivamente (a e b são vizinhos mutuamente mais próximos). Chama-se de D_1 a distância do vizinho mais próximo ao casal e de D_2 a distância do segundo vizinho mais próximo ao casal. Note que D_2 não é necessariamente $D_2 \equiv \max(D_{ak}, D_{bk'})$. Em outras palavras, nós escrevemos Z_i como uma função do aumento da distância em relação ao casal considerado. A série de aproximações corretas deve considerar a forma de organização dos conjuntos D_{ak} e $D_{bk'}$. Estes conjuntos são organizado em ordem crescente de distância em relação ao casal $\{D_1, D_2, \dots, D_{(N-1)}\}$. Mas, pode envolver um outro sítio k . Deste modo,

usando-se a aproximação $Z_a = Z_b = Z$ pode-se escrever:

$$\frac{e^{-\beta D_0}}{Z} = \frac{1}{1 + e^{-\beta \Delta_1} + e^{-\beta \Delta_2} + \dots} \quad (7.22)$$

Para β pequeno, escrevemos;

$$\frac{e^{-\beta D_0}}{Z} = 1 - e^{-\beta \Delta_1} - e^{-\beta \Delta_2} + \dots \quad (7.23)$$

onde

$$\Delta_1 = D_1 - D_0 \quad (7.24)$$

e

$$\Delta_2 = D_2 - D_0 . \quad (7.25)$$

Chamaremos a grandeza Δ_1 de “estabilidade” do casal (ver a Figura 7.1).

Figura 7.1: Sistema unidimensional com $N = 8$. Observe que D_0 é a distância entre o casal, D_1 a distância do primeiro vizinho do casal e D_2 a distância do segundo vizinho do casal.

Casais podem ser classificados por sua estabilidade. Se pensarmos em Δ_1 como sendo barreiras de ativação, podemos fazer uma analogia entre este sistema e o modelo de vidro do Bouchaud (19). Neste modelo, Bouchaud assumiu que a distribuição de barreiras de ativação possui uma calda exponencial. Em nosso modelo a distribuição exponencial aparece naturalmente e é exata.

Este resultado leva a um tempo de residência dependente da estabilidade do casal:

$$t_r = e^{\beta \Delta_1} [1 - e^{-\beta(\Delta_2 - \Delta_1)} + \dots] . \quad (7.26)$$

Um aspecto importante nesta expressão é que não temos como justificar o fato que $e^{-\beta(\Delta_2-\Delta_1)} \ll 1$. Todavia, para ganharmos alguma experiência consideraremos que $e^{-\beta(\Delta_2-\Delta_1)}$ é desprezível em relação a unidade e assim podemos escrever simplesmente:

$$t_r = e^{\beta\Delta_1} . \quad (7.27)$$

O tempo de residência médio em um sistema com infinitos casais é:

$$\langle t_r \rangle = \int_0^\infty d\Delta_1 P_d(\Delta_1) e^{\beta\Delta_1} , \quad (7.28)$$

onde $P_d(\Delta_1)$ é a função de densidade de probabilidade de Δ_1 .

Até o momento, consideramos o problema com uma dimensionalidade genérica d . Para obtermos expressões explícitas devemos considerar o caso unidimensional.

Tempo de residência em sistemas unidimensionais

Em $d = 1$, Sebastian Riseau-Gusman (25) obteve analiticamente que:

$$P_d(\Delta_1) = 2e^{-2\Delta_1} , \quad (7.29)$$

que leva a $\langle \Delta_1 \rangle = 1/2$ e $\langle \Delta_1^2 \rangle = 1/2$.

Ele também obteve analiticamente:

$$P_d(\Delta_2) = 4e^{-2\Delta_2}(1 - e^{-2\Delta_2}), \quad (7.30)$$

que leva a $\langle \Delta_2 \rangle = 3/4$ e $\langle \Delta_1^2 \rangle = 7/8$. Estas expressões validam as distribuições obtidas numericamente como ilustrado na Figura 7.2.

É interessante observar a partir da Equação 7.28 que o tempo de residência diverge para $\beta \geq \beta_c = 2$. Para ver isso, observamos:

$$\langle t_r \rangle = 2 \int_0^\infty d\Delta_1 e^{-2\Delta_1} e^{\beta\Delta_1} , \quad (7.31)$$

fazendo $x = (2 - \beta)\Delta_1$ temos;

$$\langle t_r \rangle = \frac{2}{2 - \beta} \int_0^\infty e^{-x} dx. \quad (7.32)$$

Figura 7.2: Densidade de probabilidade de estabilidade para $d=1$, em função da distância dos primeiro e segundo vizinhos do casal. As linhas sólidas são os resultados analíticos das Equações 7.29 e 7.30

Essa integral diverge para, $x \leq 2$, ou seja, para $\beta \geq 2$.

Para $\beta < \beta_c$, $\langle t_r \rangle = 2(\beta_c - \beta)^{-1}$, tal que a temperatura crítica β_c separa a fase onde o movimento do viajante não é localizado no espaço (fase estendida) de uma fase que favorece a localização nos vizinhos mais próximos, como mostrado na Figura 7.3. Observe que $\langle t_r \rangle \propto (T - T_c)^{-1}$, com $T_c = 1/2$.

7.4 Experimentos numéricos

Depois de ter construído a matriz de distâncias Euclidianas $D(i, j)$ e definida a temperatura, foram calculadas as probabilidades de transição $P_\beta(i \rightarrow j)$. Em cada uma das realizações enviou-se um viajante com tempo de vida $t_f = 4N^2$. Observe que esta paisagem aleatória apresenta muitos casais ($\sim 2/3$ dos pontos pertencem a casais) que podem ser identificados como o equivalente aos mínimos locais em modelos de vidros. Partindo de um sítio i escolhido aleatoriamente, o viajante vai para o sítio j com probabilidade $P_\beta(i \rightarrow j)$. Uma vez que no sítio j , o movimento para outro sítio k é realizado com probabilidade $P_\beta(j \rightarrow k)$.

Figura 7.3: Tempo de residência médio para $\beta < \beta_c$. Dados obtidos com $N = 1000$, número de realizações $N_R = 200$ e $\beta_c = 2$. A linha sólida corresponde a $2/(2 - \beta)$.

k) e assim, sucessivamente, esta regra é aplicada ao longo de toda a caminhada. A estatística é desenvolvida por amostragem de Monte Carlo medindo as variáveis definidas abaixo.

A dinâmica desta caminhada é determinada basicamente pelo tempo de vida do viajante (tempo experimental t_f). Embora o caminhante possa residir por muito tempo em um atrator, se for esperado tempo longo o suficiente, existe a possibilidade de ele escapar. Para caracterizar numericamente a dinâmica do modelo foram usadas as seguintes grandezas estatísticas:

1. Define-se a taxa de visitação $R(\beta)$ como sendo a razão entre o número de sítios novos visitados $n_s(\beta)$ e o número total de sítios N na rede.

$$R(\beta) = \frac{n_s(\beta)}{N}, \quad (7.33)$$

Em N_R realizações é interessante trabalhar com a média de $n_s(\beta)$.

2. A probabilidade do viajante ir para o sítio vizinho mais próximo é estimada como a razão entre o número de vezes que o caminhante foi para o sítio mais próximo $nn(\beta)$ e o número total de passos dados na trajetória t_f :

$$P_{NN}(\beta) = \frac{nn(\beta)}{t_f} , \quad (7.34)$$

Em N_R realizações é interessante trabalhar com a média de $nn(\beta)$

3. A variância s_R^2 de $R(\beta)$ foi calculada numericamente usando a mesma metodologia descrita no Capítulo 2, considerando diferentes realizações.
4. O deslocamento médio é estimado pelo desvio-padrão, acumulando depois de cada caminhada a diferença entre a posição final do caminhante $\vec{x}(t_f)$ e a posição inicial $\vec{x}(0)$, na temperatura $T = 1/\beta$:

$$\langle r(\beta) \rangle = \sum_{i=1}^{N_v} \frac{|\vec{x}(t_f) - \vec{x}(0)|}{N_v N_R} . \quad (7.35)$$

onde N_v é o número de viajantes.

5. A eficiência foi estimada pelo valor médio da razão entre o número de sítios novos visitados $n_s(\beta)$ e a distância percorrida acumulada $dc(\beta)$, para a temperatura $T = 1/\beta$:

$$\eta(\beta) = \left\langle \frac{n_s(\beta)}{dc(\beta)} \right\rangle , \quad (7.36)$$

Esta definição implica que a temperatura de maior eficiência na caminhada é aquela em que o viajante visita o maior número de sítios novos percorrendo a menor distância acumulada.

7.4.1 Sistemas unidimensionais

Número de sítios novos visitados

Conforme o previsto pela teoria, observamos numericamente uma mudança de comportamento na dinâmica do sistema como ilustrado Figura 7.4. No regime de alta temperatura $\langle R(\beta) \rangle \sim 1$ e a variância $s_R^2(\beta)$ é muito pequena.

A medida que a temperatura se aproxima do seu valor crítico ($\beta_c \sim 2 \pm 0.2$), a variância $s_R^2(\beta)$ de $\langle R(\beta) \rangle$ aumenta e passa por um máximo nessa região. Aumentando o valor de β (a temperatura diminui $T = 1/\beta$), $\langle R(\beta) \rangle$ e a variância $s_R^2(\beta)$ cai rapidamente. O deslocamento para a esquerda das curvas referentes a taxa de visitação média $\langle R(\beta) \rangle$ pode ser explicado pelo fato do sistema ser pequeno, pois para $N \geq 400$ as curvas tendem a se sobrepor.

Figura 7.4: Medidas da taxa de visitação $\langle R(\beta) \rangle$ e da probabilidade do caminhante ir para o sítio mais próximo $\langle P_{nn}(\beta) \rangle$. A variância $s_R^2(\beta)$ em $d = 1$ para $\langle R(\beta) \rangle$ em diferentes valores de temperatura $T = 1/\beta$. Da direita (estrelas) para a esquerda (circunferências) $N = 100, 200, 400$ e 800 a média foi obtida ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$), enviando um viajante por realização e usando número de passos $t_f = 4N^2$.

Observando detalhadamente a grandeza $\langle P_{nn}(\beta) \rangle$ na Figura 7.4 percebe-se que no regime de alta temperatura a probabilidade do viajante visitar os sítios mais próximos é baixa e isto equivale a uma situação em que o caminhante vai de um sítio para o outro “voando” isotropicamente no espaço. A medida que a temperatura se aproxima do seu valor crítico ($\beta_c \sim 2$), a probabilidade do viajante ir para o sítio mais próximo se aproxima de $1/2$ ($\langle P_{nn}(\beta) \rangle \sim 0,5$) e a variância $s_R^2(\beta)$ aumenta até passar por um máximo nessa região. Isto indica que

a probabilidade do caminhante permanecer no casal é igual à probabilidade de escapar. Ou seja, a dinâmica começa a ser regida pelos casais.

Deslocamento médio do viajante

Uma questão natural é saber como a temperatura afeta o deslocamento do caminhante na paisagem. No regime de alta temperatura o deslocamento médio é da ordem de 0,35. Para baixas temperaturas $\langle r(\beta) \rangle$ tende para zero. Os resultados numéricos da Figura 7.5 mostram que a mudança do regime estendido para o localizado começa em $\beta = 2$.

Figura 7.5: O deslocamento médio $\langle r(\beta) \rangle$ em $d = 1$ para diferentes valores de temperatura $T = 1/\beta$. São mostrados $N = 400$ (triângulo) e 800 (circunferência). As médias foram obtidas ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$), enviando um viajante por realização e usando número de passos $t_f = 4N^2$.

A eficiência da caminhada em sistemas unidimensionais

Quando se pensa em estratégia de otimização é intuitivo imaginar que ao caminhante convém visitar o maior número de sítios novos andando o mínimo possível. Percorrer distâncias longas tem um custo alto (energia gasta) que pode

ser compensado pelo ganho no número de sítios novos visitados. A Figura 7.6 mostra que a eficiência de uma caminhada, que obedece esta dinâmica, depende do tamanho do sistema e que para $N \geq 400$ a máxima eficiência se obtém em $\beta_\eta = 1,3 \pm 0.2$. Quando observamos o número total de sítios novos visitados nesta temperatura (ver também a Figura 7.4), verificamos que existe uma probabilidade de que toda a paisagem seja visitada, $\langle R(\beta_\eta) \rangle \approx 1$.

Figura 7.6: Medidas de eficiência $\langle \eta(\beta) \rangle$ em $d = 1$ para diferentes valores de temperatura β . Da direita (estrelas) para a esquerda (circunferências) $N = 100, 200, 400$, e 800 . Os resultados foram obtidos ao longo de $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$) e usando um número de passos $t_f = 4N^2$.

7.4.2 Sistemas bidimensionais

Número de sítios novos visitados

Da mesma forma que em $d = 1$, a Figura 7.7 mostra o comportamento de $\langle R(\beta) \rangle$, $s_R^2(\beta)$ e $\langle P_{nn}(\beta) \rangle$ para $d = 2$. Da esquerda para a direita, o valor de $\langle R(\beta) \rangle = 1$ para temperaturas altas indica que todos os sítios da paisagem foram visitados. O número de vezes que o viajante vai para o sítio mais próximo $\langle P_{nn}(\beta) \rangle$

vai lentamente para 1 e a variância $s_R^2(\beta)$ é máxima na região de transição.

Figura 7.7: Medidas da taxa de visitação $\langle R(\beta) \rangle$ e da probabilidade do caminhante ir para o sítio mais próximo $\langle P_{nn}(\beta) \rangle$. A variância $s_R^2(\beta)$ em $d = 2$ para $\langle R(\beta) \rangle$ em diferentes valores de temperatura $T = 1/\beta$. Da esquerda (triângulos) para a direita (circunferências) $N = 200, 400, 800$ e 1600 . As médias foram obtidas ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$) enviando um viajante por realização e usando número de passos $t_f = 4N^2$.

Para $\beta > 0$ tem-se dois regimes: se $\beta < \beta_t$ predomina o regime estendido (ergódico) e os atratores de período 2 são menos estáveis. Isto pode ser visto pelo número de vezes que a caminhante vai para o sítio vizinho mais próximo (ver $\langle P_{nn}(\beta) \rangle$ na Figura 7.7). Quando $\beta > \beta_t$ os atratores de período 2 ficam mais estáveis e predomina o regime localizado, de tal modo que $\langle P_{nn}(\beta) \rangle \rightarrow 1$, quando $\beta \rightarrow \infty$. Como os atratores não são totalmente estáveis o caminhante pode ser aprisionado e depois de algum tempo escapar, seguindo para outro atrator de período 2 ou voltando para o atrator precedente, onde será aprisionado e novamente escapará. Próximo da região de transição, a variância no número de sítios novos visitados $s_R^2(\beta)$ aumenta muito rapidamente (quando comparado com a variância $s_R^2(\beta)$ de temperaturas mais altas ou mais baixas). A variância $s_R^2(\beta)$ em $d = 2$ mostra que a região de transição se encontra ao redor de $\beta_t \sim 12 \pm 2$ e

que como em $d = 1$ seu ponto de máximo diminui conforme o tamanho do sistema aumenta.

Deslocamento médio do viajante

A Figura 7.8 mostra que, para $\beta \leq 10$, em média o viajante explora toda a paisagem. Para valores maiores de β , o deslocamento tende para zero ($\beta \rightarrow \infty$), sendo que as curvas referentes a $N = 400, 800$ e 1600 , se cruzam quando $\beta \sim 12 \pm 2$ (ver Figura 7.7).

Figura 7.8: Medidas do deslocamento médio $\langle r(\beta) \rangle$ em $d = 2$ para diferentes valores de temperatura β . Da esquerda (estrelas) para a direita (circunferências) $N = 400, 800$ e 1600 . As médias foram obtidas ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$), enviando um viajante por realização com um tempo de vida $t_f = 4N^2$.

Eficiência da caminhada

Em $d = 2$, a eficiência é máxima para $\beta_\eta \sim 9 \pm 1$. Isto significa que como em $d = 1$ a temperatura de máxima eficiência é maior que a temperatura de transição β_t (ver Figura 7.9).

Na região de máxima eficiência ($\beta \sim 9 \pm 1$) o percentual de vezes que o caminhante vai para o sítio mais próximo é de $\sim 75\%$. Este valor pode ser comparado aos resultados experimentais de Gross et. al. (37). Eles estudaram

as caminhadas de ovelhas em pastos onde vários montes de alfafa foram espalhados aleatoriamente. O melhor modelo para explicar tais caminhadas foi aquele construindo com base na regra do vizinho mais próximo, de forma que aproximadamente 75% de todos os movimentos fossem para as moitas mais próxima. Se este fosse um modelo de comportamento animal, por analogia, significaria dizer que as ovelhas otimizam suas caminhadas na borda da transição, entre o regime estendido e o localizado.

Figura 7.9: Medidas da eficiência $\langle \eta(\beta) \rangle$ em $d = 2$ para diferentes valores de temperatura β . Os valores de N são: 200 (triângulos), 400 (estrelas), 800 (circunferência) e 1600 (retângulos). As médias foram obtidas ao longo $N_R = 100$ realizações na caminhada sem memória ($\tau = 0$) e usando um número de passos $t_f = 4N^2$.

7.5 Conclusão

Nesse capítulo consideramos uma possível generalização para o problema da caminhada do turista. Permite-se que o turista não vá necessariamente para o ponto turístico mais próximo. Isto é realizado através da introdução de um parâmetro extra ao problema original, que controla a estocasticidade da caminhada. Com a finalidade de simplificar, consideramos somente as caminhadas sem memória ($\tau = 0$).

Para sistemas unidimensionais obtivemos resultados analíticos aproxima-

dos. Estes resultados analíticos validaram os experimentos numéricos. Para sistemas bidimensionais, somente resultados numéricos foram apresentados.

Esta caminhada é caracterizada por dois regimes, um estendido e outro localizado. Estes dois regimes são separados por uma temperatura característica bem definida. Em uma dimensão mostramos que em β_c realmente pode existir uma transição de fase. Em duas dimensões, a temperatura β_t , que separa estes dois regimes é tipicamente uma ordem de grandeza maior do que em sistemas unidimensionais e é análoga a temperatura fenomenológica T_g na teoria de vidros (38), onde a verdadeira temperatura de transição se situa em $T = 0$ ($\beta_t \rightarrow \infty$).

Mostramos que a eficiência máxima entre interesses conflitantes ocorre em temperaturas na borda da transição. Em suma, existe uma mudança drástica de comportamento de um regime estendido (ergódico) para um regime localizado (não-ergódico) em uma temperatura característica cujas propriedades são:

- Para $d = 1$ esta transição parece ocorrer em $\beta_c = 2$ e para $d = 2$ em $\beta_t = 12$. Este deslocamento de β é um efeito da diminuição de correlação entre os sítios com o aumento da dimensionalidade.
- Na temperatura de transição, a probabilidade do caminhante ir para o sítio mais próximo $\langle P_{nn}(\beta) \rangle$ é aproximadamente 50% para $d = 1$ e 80% para $d = 2$.
- Na borda da transição a eficiência é maximizada e o percentual de vezes que o caminhante vai para o sítio mais próximo é de $\sim 75\%$ em sistemas bidimensionais. Sugestivamente, este valor é igual aos resultados experimentais de Gross et. al. (37), para comportamento de herbívoros em pastos.
- o valor diferente da temperatura de transição β_c para $d = 1$ e 2 está indicando que a dimensionalidade afeta a estabilidade dos ciclos, aumentando o grau de liberdade do caminhante em $d = 2$.

Capítulo 8

Conclusão

No nosso trabalho, a abordagem computacional se revelou importante para nos ajudar a compreender processos dinâmicos em uma rede, com uma riqueza de detalhe dificilmente alcançada se o experimento fosse real. Esse tipo de estudo também tem sido muito usado em experimentos envolvendo dinâmica social de propagação de doenças, dinâmica de sistemas biológicos e propagação de sinais. Experimentos como estes são muito difíceis de serem realizados, até porque em muitos casos não é possível controlar todas as variáveis externas e em outros casos envolvem questões éticas.

Nesse trabalho, nós estudamos as propriedades estatísticas de uma dinâmica determinística em uma paisagem aleatória. A dinâmica foi denominada de caminhada do turista por tratar-se de um processo de otimização local, parecido com o problema de um viajante mochileiro que depende de carona para visitar as cidades turísticas. Em todas as situações consideramos que a memória das cidades visitadas é pequena comparada com o número de cidades existentes, que supusemos ser grande.

No Capítulo 3, mostramos que os sistemas unidimensionais são importantes por apresentar certas patologias que diferem drasticamente dos sistemas de dimensionalidade mais elevadas. Mostramos por exemplo, que alguns atratores são proibidos enquanto que outros são raros por causa de restrições geométricas.

No Capítulo 4, mostramos que os sistemas bidimensionais já apresentam aspectos mais gerais, como por exemplo, o fato de todas as distribuições de atratores poderem ser descritas por uma única função, uma lei de potência com um corte exponencial. As distribuições de tempos de transiente continuaram a

depende da memória, enquanto o número médio de sítios por atrator tornou-se independente dela.

No Capítulo 5, estudamos o modelo no limite de alta dimensionalidade e em dimensionalidade intermediária. O caso ($d \rightarrow \infty$) foi simulado considerando o modelo de ligações aleatórias simétricas estudado por Percus (15). Observamos que os tempos de transiente são afetados pelo tamanho do sistema, mas podem ser colapsados em uma única curva. Na distribuição de atratores verificou-se que o corte exponencial se torna menos importante à medida que o número de sítios aumenta. Com a finalidade de comparação, implementamos a mesma dinâmica para o mapeamento aleatório de Derrida-Flyvbjerg (13) (ligações aleatórias assimétricas). Neste modelo, as distribuições de tempos de transiente e de atratores dependem do tamanho do sistema, mas podem ser escalonadas em termos de variáveis reduzidas. Como no caso de ligações simétricas, verificou-se que a memória não afeta o expoente da distribuição de atratores. Também mostramos que somente o modelo de ligações aleatórias simétricas é equivalente ao modelo Euclidiano em dimensão infinita. Verificamos que, à medida que a dimensionalidade aumenta, as distribuições apresentadas na seção 2.4 convergem para o modelo de ligações aleatórias simétricas. Conjeturamos que o comportamento de lei de potência é predominante no limite termodinâmico (número de sítios muito grande).

No Capítulo 6, aplicamos a caminhada do turista em um dicionário de sinônimos. Mostramos que o grafo de tal dicionário pode ser embebido em um sistema de coordenadas Euclidianas de baixa dimensão, de acordo com a teoria de Steyvers e Tenenbaum (30) e contrariando ao preconizado pelo modelo de Análise de Semântica Latente (LSA) (17; 18).

No Capítulo 7, fizemos uma generalização do problema do turista com $\tau = 0$, para uma dinâmica estocástica cuja aleatoriedade é controlada por um parâmetro externo análogo a temperatura. Mostramos que existe uma temperatura crítica que separa o regime estendido do localizado. Para $d = 1$ este efeito pode ser estudado analiticamente. Quando $d > 1$, resultados analíticos sugerem que a temperatura crítica tende para zero, mas a mudança de regime continua sendo drástica. Esta temperatura é análoga à temperatura T_g encontrada na teoria de

vidros. Uma analogia é estabelecida com o modelo de vidros de Bouchaud (19).

Nesta tese examinamos um modelo original e apresentamos uma série de comportamentos interessantes presentes nele. Por ser um primeiro estudo sabemos que todos os nossos resultados ainda podem ser muito melhorados e refinados. Isso deverá ser feito em nosso trabalho futuro.

Finalmente, acreditamos que o modelo de otimização local proposto possa descrever várias situações que não foram abordadas neste trabalho. Nos comentários sobre o nosso trabalho na revista *Nature*, Stanley e Buldyrev(27) enfatizaram que o problema do turista proporciona uma nova abordagem para modelar comportamento migratório de insetos, pássaros e mamíferos. Suas conclusões sugerem que alguns padrões de comportamento, usualmente associados ao livre arbítrio, podem ser resultado da aplicação de regras determinísticas em um meio ambiente aleatório.

Referências Bibliográficas

- [1] D. J. Watts, S. H. Strogatz, Collective dynamics of “small world” networks, Nature 393 (1998) 440–442.
- [2] R. Lewin, Complexidade: A vida no limite do caos, Editora Rocco, Rio de Janeiro, 1994.
- [3] H. Freund, P. Grassberger, The Red Queen’s walk, Physica A 190 (3-4) (1992) 218–237.
- [4] T. Tomé, M. J. de Oliveira, Dinâmica estocástica e reversibilidade, Editora da Universidade de São Paulo, São Paulo, 2001.
- [5] E. Arashiro, J. R. D. de Felício, Damage spreading in one-dimensional Hinrichsen-Domany model, Braz. J. Phys. 30 (4) (2000) 677–679.
- [6] T. Tomé, J. R. D. de Felício, Probabilistic cellular automaton describing a biological immune system, Phys. Rev. E 53 (1996) 3976.
- [7] H. Flyvbjerg, N. Kjaer, Exact solution of Kauffman’s model with connectivity one, J. Physics A: Math. and Gen. 21 (7) (1988) 1695–1718.
- [8] S. N. Coppersmith, L. P. Kadanoff, Z. Zhang, Reversible boolean networks i: Distribution of cycle lengths, Physica D 149 (1-2) (2001) 11–29.
- [9] S. N. Coppersmith, L. P. Kadanoff, Z. Zhang, Reversible boolean networks ii. phase transitions, oscillations, and local structures, Physica D 157 (1-2) (2001) 54–74.
- [10] A. Bhattacharjya, S. Liang, Power-law distributions in some random boolean networks, Phys. Rev. Lett. 77 (8) (1996) 1644–1647.

- [11] P. Bak, *How Nature works*, Oxford University Press, Oxford, 1997.
- [12] Z. Somogyvári, S. Payrits, Length of state cycles of random boolean networks: an analytic study, *J. Phys. A: Math Gen.* 33 (2000) 6699–6706.
- [13] B. Derrida, H. Flyvbjerg, The random map model - a disordered model with deterministic dynamics, *Journal de Physique (Paris)* 48 (6) (1997) 971–978.
- [14] B. Harris, Probability-distributions related to random mappings, *Ann. Math. Stat.* 31 (4) (1960) 1045–1062.
- [15] A. G. Percus, O. C. Martin, The stochastic traveling salesman problem: finite size scaling and the cavity prediction, *J. Stat. Phys.* 94 (1999) 739–758.
- [16] A. Percus, *Voyageur de commerce et problèmes stochastiques associés*, Ph.D. thesis, Université Paris 6, Paris (Septembre 1997).
- [17] T. Landauer, P. W. Foltz, D. Laham, Introduction to latent semantic analysis. discourse processes, *Psychological Review* 25 (1998) 259–284.
- [18] T. K. Landauer, S. Dumais, A solution to plato’s problem: The latent semantic analysis theory of acquisition, introduction, and representation of knowledge, *Psychological Review* 104 (1997) 211–240.
- [19] J. P. Bouchaud, Weak ergodicity breaking and aging in disordered-systems, *J. Phys. I* 2 (9) (1992) 1705–1717.
- [20] C. Monthus, J. P. Bouchaud, Models of traps and glass phenomenology, *J. Phys. A: Math. Gen.* 29 (14) (1996) 3847–3869.
- [21] P. J. Clark, F. C. Evans, On some aspects of spatial pattern in biological populations, *Science* 121 (1955) 397–398.
- [22] J. P. Clark, Grouping in spatial distributions, *Science* 123 (1956) 373–374.
- [23] M. F. Dacey, Proportion of reflexive n^{th} -order neighbours in a spatial distribution, *Geographical Analysis* 1 (1969) 385–388.

- [24] T. F. Cox, Reflexive nearest neighbours, *Biometrics* 37 (1981) 367–369.
- [25] S. Risau-Gusman, Comunicação privada (2001).
- [26] D. Treit, J. Menard, C. Royan, Anxiogenic stimuli in the elevated plus-maze, *Pharmacology, Biochemistry and Behavior* 44 (2) (1993) 463–469.
- [27] H. E. Stanley, S. V. Buldyrev, Statistical physics - the salesman and the tourist, *Nature* 413 (6854) (2001) 373–374.
- [28] S. M. Bhattacharjee, F. Seno, A measure of data collapse for scaling, *J. Phys. A* 34 (2001) 6375–6380.
- [29] U. Bastolla, G. Parisi, Attractors in fully asymmetric neural networks, *J. Phys. A:Math. Gen.* 30 (1997) 5613–5631.
- [30] M. Steyvers, J. B. Tenenbaum, The large-scale structure of semantic networks: statistical analyses and a model for semantic growth, *cond-Mat/0110012* (2001).
- [31] L. A. Bunimovich, S. E. Troubetzkoy, Recurrence properties of Lorentz lattice gas cellular automata, *J. Stat. Phys.* 67 (1-2) (1992) 289–302.
- [32] D. Gale, J. Propp, S. Sutherland, S. Troubetzkoy, Further travels with my ant, *Math. Intelligencer* 17 (3) (1995) 48–56.
- [33] C. Salum, S. Morato, A. C. Roque-da Silva, Anxiety-like behavior in rats: a computational model, *Neural Networks* 13 (1) (2000) 21–29.
- [34] G. F. Lima, A. S. Martinez, O. Kinouchi, Deterministic walks in random media, *Phys. Rev. Lett.* 87 (1) (2001) 010603.
- [35] D. E. Goldberg, *Genetic Algorithms*, Addison-Wesley Publishing, Reading, 1989.
- [36] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by simulated annealing, *Science* 220 (1983) 671–680.

- [37] J. E. Gross, C. Zank, N. T. Hobbs, D. E. Spalinger, Movement rules for herbivores in spatially heterogeneous environments: responses to small scale pattern, *Landscape Ecology* 10 (1995) 209–217.
- [38] M. Mézard, Statistical physics of the glass phase, cond-Mat/0110363 (2001).

Produção Científica

Participação em manuscritos publicados no decorrer do curso:

1. G. F. Lima, A. S. Martinez and O. Kinouchi. Deterministic walk in random media. *Phys. Rev. Lett.*, **87**, 010603 (2001).
2. S. M. Kuva, G. F. Lima, O. Kinouchi, M. H. R. Tragtenberg and A. C. Roque. A minimal model for excitable and bursting elements. *Neurocomputing*, **38-40**, 255-261 (2001).

Manuscrito submetido:

1. O. Kinouchi, A. S. Martinez, G. F. Lima, G. M. Lourenço and S. Risau-Gusman. Deterministic walks in random networks: an application to The-saurus graphs. *Physica A*.

Alguns comentários relativos ao tema da tese:

1. H. E. Stanley and S. Buldyrev. Statistical Physics - The salesman and the tourist. *Nature*, **413**, 373-374 (27 September 2001).
2. *Agência USP de notícias* do Campus de Ribeirão Preto, número **735**, página 07, em 03/12/2001.
3. *Jornal a Cidade*, de Ribeirão Preto - São Paulo. Caderno Especial, página 1, número **277**, em 02/12/2001.
4. *Agência USP de notícias*, dia 22/02/2002.
<http://www.usp.br/agenciausp/repags/2002/pags/038.htm> .
5. *Ciência-Hoje on-line*, dia 04/04/2002.
<http://www.uol.com.br/cienciahoje/chdia/n570.htm> .

Deterministic Walks in Random Media

Gilson F. Lima,^{1,2} Alexandre S. Martinez,¹ and Osame Kinouchi¹

¹*Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Avenida Bandeirantes 3900, CEP 14040-901, Ribeirão Preto, SP, Brazil*

²*Escola Técnica Federal de Mato Grosso, R. Zulmira Canavarros 95, CEP 78005-390, Cuiabá, MT, Brazil*

(Received 5 May 2000; published 19 June 2001)

Deterministic walks over a random set of N points in one and two dimensions ($d = 1, 2$) are considered. Points (“cities”) are randomly scattered in $\mathbf{R}^d$ following a uniform distribution. A walker (“tourist”), at each time step, goes to the nearest neighbor city that has not been visited in the past τ steps. Each initial city leads to a different trajectory composed of a transient part and a final p -cycle attractor. Transient times (for $d = 1, 2$) follow an exponential law with a τ -dependent decay time but the density of p cycles can be approximately described by $D(p) \propto p^{-\alpha(\tau)}$. For $\tau \gg 1$ and $\tau/N \ll 1$, the exponent is independent of τ . Some analytical results are given for the $d = 1$ case.

DOI: 10.1103/PhysRevLett.87.010603

PACS numbers: 05.40.Fb, 05.45.Df, 05.90.+m, 87.15.Aa

The study of random walks has been very fruitful in physics and mathematics, and the theory of such stochastic processes is a well-developed subject. The study of deterministic walks is also an interesting subject, but presents the analytical difficulties common to the area of nonlinear dynamical systems and has been less investigated [1]. Here we propose a simple and intriguing problem: a deterministic walk over a random graph with N nodes that is also an example of a local (“on-line”) optimization dynamics. It may be called the “local traveling salesman problem” or perhaps the “tourist problem” for short. The deterministic dynamics produces a division of the system phase space in a $\mathcal{O}(N)$ number of attractor basins which trap the walker (ergodicity is broken). The problem is reminiscent of walks in rugged landscapes or zero-temperature spin-glass dynamics but the equivalent of “local minima” are cycles instead of point attractors, as in Kauffman networks [2].

The model is defined as follows: N points are randomly distributed with a uniform density ρ in $\mathbf{R}^d$, where d is the dimensionality of the space. These points may be thought of as “cities” and they may be viewed as vertices of a random graph. At each time step, the “tourist” follows the deterministic rule: *Go to the nearest city that has not been visited in the past τ time steps*. Notice that the tourist wants to minimize only the distance to the next city (a local optimization procedure), not the sum of all distances along the trajectory as in the traveling salesman problem.

Starting from a random city, the tourist performs a trajectory composed of a transient part and a final p -cycle attractor. In this Letter, we report the statistics for some relevant quantities similar to those measured in Kauffman networks [2]: (a) the probability $P_\tau(t)$ for obtaining a transient of size t , defined as the number of steps before the tourist enters some attractor; for large t , it is exponential $P_\tau(t) \propto \exp[-t/\xi(\tau)]$ with the decay time $\xi(\tau)$ growing exponentially for $d = 1$ and linearly for $d = 2$; (b) the density of p cycles $D_\tau(p)$, for $d = 1$, seems to be neither

exponential nor a power law, while for $d = 2$, $D_\tau(p)$ follows a power law for nonextreme values $D_\tau(p) \propto p^{-\alpha(\tau)}$ for $\tau \gg 1$, with α independent of τ for $\tau/N \ll 1$; (c) the total density of attractors, $\mathcal{D}(\tau) = \sum_p D_\tau(p)$, is the number of different attractors per city for a given memory, decays exponentially for $d = 1$ and, for $d = 2$, as τ^{-1} ; for $d = 1$, $D\xi$ is linear in τ , and for $d = 2$ it is a constant ($\tau \gg 1$ and $\tau/N \ll 1$); and (d) for $d = 1$, the average number $\langle n(p) \rangle$ of cities present in a p cycle which is compared to an analytical result where we show that some cycles are prohibited as odd cycles (apart 3 cycle) and 6 cycle for $\tau = 1$.

Our model can be of general interest for optimization theory with local constraints, nonadditive cost functions, and studies of deterministic dynamical systems with quenched disorder. However, we would like to suggest some more specific motivations for considering this class of problem. The model can be viewed as an example of local foraging strategies [3]. It is arguable that for biological agents (and biologically inspired robots) it could be sometimes more important to minimize the distance traveled in each movement between two safe places instead of optimizing some global cost function [4]. In another spatial scale, cycles could be related to stable migratory routes on environments with localized resources. Local optimization appears due to short range sensorial/cognitive capacities which is a determinant factor in unfamiliar or hostile landscapes [5].

The only nontrivial parameter of the model is the memory window τ . Self-avoidance is limited to this window and trajectories can intersect outside this range. If $\tau = 0$ (no memory) the tourist goes simply to the nearest city until two cities, which are reciprocally nearest neighbors, are found, entering a 2-cycle. For $\tau = N - 2$ the trajectory is totally self-avoiding and one has a kind of traveling salesman problem nearest-neighbor algorithm [4]. The interesting cases are the intermediate ones. For example, if $\tau = 1$, the last visited city cannot be revisited,

and only p cycles with $p \geq 3$ can exist. For all τ , the relation $p \geq \tau + 2$ holds.

We stress that this problem is not simply related to geometrical properties since the cycles appear only due to the introduced dynamics. Naively, one could think that a p cycle is a geometrical object, for example a cluster where the distances between the points are smaller than any distance outside the cluster. This indeed is a sufficient but not necessary condition to obtain a p attractor. For example, for $d = 2$ (Fig. 1a), a walker with memory $\tau = 1$ starts from city A and finds the 4-cycle $ABCD$. Although city E is close to the cluster (since $BE < AB$), it is never visited because $BC < BE$ and $CD < CE$. However, if the tourist starts from city C , one gets a 3-cycle that includes city E . This degeneracy and superposition of attractors can be understood noticing that the trajectories of Fig. 1 are represented in configuration space, not in phase space. In phase space, points correspond to $\tau + 1$ -tuples $(\mathbf{X}_t, \dots, \mathbf{X}_{t-\tau})$ where $\mathbf{X}_t$ is the position (x, y) of the tourist at time t and trajectories never intersect. Only for $\tau = 0$ the configuration space is equivalent to the phase space.

In the numerical experiments, N points (with d spatial coordinates) are randomly scattered following a uniform distribution in the interval $[0, 1]^d$. The cities are arbitrarily labeled as $i = 1, \dots, N$ and one constructs the Euclidean distance matrix $\mathbf{D}$ [6]. Starting from some city, the tourist accesses a transient trajectory until getting trapped in some periodic attractor. The number of steps before the tourist enters the cycle defines the transient length t . The period p and the number of different cities n that pertain to the attractor are also determined. The same city can be visited more than once, thus $n \leq p$.

A finite size study has showed that the behavior of the system is smooth as a function of $1/N$, so we have used $N = 10\,000$ as a reasonable number for our simulations. Periodic boundary conditions have been considered for $d = 1$ and $d = 2$ and do not cause significant differences compared to free boundaries. Since each city is used as a starting point, a landscape with N cities produces N differ-

ent transients. The statistics over $N_R = 1000$ realizations of cities sets ("maps") are collected. The fluctuation in Figs. 2, 3 and 5 are mainly due to the sampling (N_R), not N . We have $NN_R = 10^7$ trajectories treated for each dimensionality d .

A natural question is about the existence of some critical τ that produces a phase transition (in the thermodynamical limit $N \gg 1$), for example the emergence of an untrapped (percolating) transient state. The distribution of transient times does not suggest this possibility because it is exponential, $P_\tau(t) \propto \exp[-t/\xi(\tau)]$. This is shown in Figs. 2 and 3 for $d = 1$ and $d = 2$, respectively. The memory can be very large ($\tau \gg 1$) but $\tau/N \ll 1$. The exponential decay may be understood as follows. The accumulative of the transient times distribution is the probability of a walker not being trapped in a cycle at t (irrespective to the cycle period p). Consider $q(\tau)$ the probability of the walker not entering a cycle after a movement. The movement of the walker is generated deterministically, but $q(\tau)$ is the same on every movement. The probability of a walker making t steps and not entering a cycle is $P(t) \propto e^{-t/\xi(\tau)}$, where $\xi(\tau) = -1/\ln q(\tau)$. The probability that a walker enters a cycle in a given movement is $\tilde{p} = 1 - q$. Assuming $\tilde{p} \ll 1$, leads to $\tilde{p} = 1/\xi(\tau)$.

For $d = 1$, the characteristic times grow as $\xi(\tau) \propto \exp(\gamma\tau)$ (inset Fig. 2) and the total density decays as $\mathcal{D}(\tau) \propto \exp(-\gamma'\tau)$ (Fig. 4a). For $d = 2$, one observes the linear dependence $\xi(\tau) \propto \tau^\delta$, $\delta = 1.0$ (inset Fig. 3); the total density decays as a power law $\mathcal{D}(\tau) \propto \tau^{-\delta'}$ with $\delta' = 1.0$ (Fig. 4b). The probability $\tilde{p}$ is to a first approximation proportional to $\mathcal{D}(\tau)$. This means that $\mathcal{D}(\tau)\xi(\tau)$ should be constant, that is, $\gamma = \gamma'$ and $\delta = \delta'$. This is indeed the case for large values of τ in $d = 2$ (see Fig. 4c). For $d = 1$, although the exponential terms cancel, a linear factor remains. A better expression for $d = 1$ decay time is $\xi(\tau) = c\tau \exp(\gamma\tau)$. The linear prefactor occurs in both $d = 1$ and $d = 2$, the exponential dependence on τ (in $\mathcal{D}$ Fig. 4a and ξ inset of Fig. 2) is a fingerprint

FIG. 1. (a) Example of superposition of attractors for $d = 2$ and $\tau = 1$: starting from A one obtains the 4-cycle $ABCD$, but starting on C one gets the 3-cycle CBE ; (b) the 3-cycle base block for $\tau = 1$; (c) the 4-cycle for $\tau = 1$; (d) next permissible cycle for $\tau = 1$: an 8-cycle made of two base blocks (see also Fig. 5a); and (e) an example of odd ($p = 13$) cycle for $\tau = 2$ which is possible because of an internal loop.

FIG. 2. Distribution of transient times $P_\tau(t)$ for $d = 1$. From left to right: $\tau = 0, 2, 3, 4$, and 5 . Inset: Decay time $\xi(\tau)$.

FIG. 3. Distribution $P_\tau(t)$ of transient times for $d = 2$. From left to right: $\tau = 0, 1, 2, 3, 5$, and 10 . Inset: Decay time $\xi(\tau)$.

of $d = 1$ and is due to the probability of the occurrence of barriers [$P(d_{i+1} > d_i + d_{i-1} + \dots + d_{i-\tau})$, where d_i is the distance traveled in step i] which trap the dynamics of a walker into cycles. On the basis of central limit theorem, it is possible to show that this probability decays exponentially as a function of τ .

A property of natural interest is the density $D_\tau(p)$ of p cycles, estimated as the number of *different* p cycles divided by N , in the limit of very large systems ($N \gg 1$).

FIG. 4. Total density of attractors $\mathcal{D}(\tau)$: (a) $d = 1$; (b) $d = 2$; and (c) $\mathcal{D}(\tau)\xi(\tau)$ for $d = 1$ (squares) and $d = 2$ (circles), error bars smaller than symbol size.

Evaluating this quantity requires careful enumeration because, when starting from all the possible initial states, one must not count the same attractor twice. For $d = 1$, some attractors are forbidden as the odd (apart $p = 3$) and $p = 6$ cycles for $\tau = 1$, the presence of long cycles indicates $D_\tau(p)$ is nonexponential, although the evidence for a power law is weak (Fig. 5a). For $d = 2$, one observes, for nonextreme values, a power law $D_\tau(p) \propto p^{-\alpha(\tau)}$ (Fig. 5b) behavior [7–9]. For $\tau \gg 1$ and $\tau/N \ll 1$, the exponent $\alpha = 2.7 \pm 0.2$ is independent of τ (inset Fig. 5b).

The average number of cities $\langle n(p) \rangle$ pertaining to cycles of period p has been studied (Fig. 6). For $d = 1$ there is almost no dispersion in the number of cities per attractor. A p cycle has $n(p)$ cities. We also have found that, for $n > 2(\tau + 2)$, the following relation holds for even cycles: $n(p) = p/2 + \tau + 1$. To see how this relation emerges, notice that for each τ there are configurations of points that constitute barriers to the displacement of the tourist. For example, for $\tau = 1$, a barrier to right displacement appears when $d_{i-1} + d_i < d_{i+1}$. Normally, a tourist will perform a constant direction movement (say, to the right) until a barrier is encountered, then the tourist turns around and moves until another barrier (to left displacements) is found. A cyclic attractor stabilizes, the attractor set being composed by the points between the two barriers.

FIG. 5. Examples of attractor densities $\mathcal{D}_\tau(p)$: (a) $d = 1$, $\tau = 1$ (squares), $\tau = 3$ (triangles), and $\tau = 6$ (circles); (b) $d = 2$, $\tau = 1$ (squares), $\tau = 4$ (stars), and $\tau = 10$ (circles), $N = 10\,000$ and $N_R = 1000$. Inset: exponent $\alpha(\tau)$.

FIG. 6. Number of cities per attractor $\langle n(p) \rangle$. (a) $d = 1$, $\tau = 1$ (filled circles) and $\tau = 6$ (empty circles), theoretical curves (solid) $\langle n(p) \rangle = p/2 + \tau + 1$.

For each τ there is a minimum cycle of period $p_\tau = \tau + 2$, which we call a *base block* (Fig. 1b). A base block is composed of $n_\tau = \tau + 2$ cities. The next cycles follow specific constructions (Fig. 1c). But when n is large, geometrical constraints impose that the most common p cycles are made of two base blocks (one in each attractor extremity) joined by n_I *intermediate cities*; see Fig. 1d. An attractor with n cities thus has $n_I = n - 2n_\tau$ intermediate points. These intermediate cities contribute to the total period with $p_I = 2n_I + 2$ steps (since, for $n_I = 0$, the joining of the base blocks contributes with two steps). Thus, the total period is $p = 2 \times p_\tau + p_I = 2(n - \tau - 1)$, which leads to $n(p) = p/2 + \tau + 1$ (Fig. 6).

This relation holds for cycles with $n \geq 2n_\tau = 2(\tau + 2)$, because only these cycles can incorporate two independent base blocks. For $\tau = 1$, this is the unique conceivable manner of constructing cycles, meaning that odd cycles are prohibited (and also $p = 6$ cycles; see Fig. 5a). For $\tau > 1$, it is possible to construct odd cycles by using internal loops and very specific initial conditions (an example with $\tau = 2$ is given in Fig. 1e).

In $d = 2$, the attractors are polygons with different forms and shapes so that this strict relation $n(p)$ between periods and cities does not hold, although $\langle n \rangle$ also scales linearly with p (not shown). For $\tau = 1$, one finds that odd cycles are less probable than even cycles (Fig. 5b), which is reminiscent of the $d = 1$ behavior. Indeed, this occurs because elongated, quasilinear, odd attractors in two-dimensional space are prohibited by the same geometrical constraints present in the one-dimensional case.

Finally, another analytical result for the $d = 1$ and $\tau = 0$ case can be obtained. It is not hard to see that the distribution of interval sizes follows an exponential distribution $P(d_i) = Ne^{-Nd_i}$ and only 2-cycle attractors exist, corresponding to pairs of reciprocal nearest neighbors. Since the interval sizes are independent variables, one can show

that the probability to have mutually nearest neighbors is $P_2 = P(d_{i-1} > d_i \text{ and } d_{i+1} > d_i) = 1/3$, that is, on average, one third of the sequences of four points leads to reciprocal nearest neighbors and so to 2-cycles. Since the number of sequences of four points is, in the large N limit, equal to the number of points, one obtains $D_0(2) = 1/3$. This has been fully confirmed by our numerical simulations [10].

Deterministic walks, which are partially self-avoiding, have been presented in continuous space in one and two dimensions. This simply stated problem presents a rich and highly nontrivial dynamics, where some particular and general aspects have been presented. Also, we have shown that although memory favors space exploitation, it does not lead to an ergodic behavior. Ergodicity can be obtained by a generalization of the proposed model by introducing a stochastic component which will be fully presented and explored in the near future.

O. Kinouchi acknowledges support from FAPESP. The authors thank S. Risau-Gusman, N. Caticha, R. Vicente, N. Alves, and U.P.C. Neves for their useful suggestions and P. Stadler for suggesting the name “tourist problem.”

-
- [1] H. Freund and P. Grassberger, *Physica* (Amsterdam) **190A**, 218 (1992); L. A. Bunimovich and S. Troubetzkoy, *J. Stat. Phys.* **67**, 289 (1992); D. Gale, J. Propp, S. Sutherland, and S. Troubetzkoy, *Math. Intelligencer* **17**, 48 (1995); C. Salum, S. Morato, and A.C. Roque-da-Silva, *Neural Netw.* **13**, 21 (2000).
 - [2] A. Bhattacharjya and S. Liang, *Phys. Rev. Lett.* **77**, 1644 (1996).
 - [3] G.M. Viswanathan, S. Buldyrev, S. Havlin, M.G.E. da Luz, E.P. Raposo, and H.E. Stanley, *Nature* (London) **401**, 911 (1999).
 - [4] J.E. Gross, C. Zank, N.T. Hobbs, and D.E. Spalinger, *Landscape Ecol.* **10**, 209 (1995).
 - [5] P.A. Zollner and S.L. Lima, *OIKOS* **80**, 51 (1997).
 - [6] From the $\mathbf{D}$ matrix another matrix V_{ik} is created where $k = 1, \dots, \tau + 1$ are the (ranked) nearest neighbors of city i . The dynamics can be performed over the entries of the $\mathbf{V}$ matrix instead of $\mathbf{D}$.
 - [7] A finite-size study has been performed and is not conclusive since the cutoff lies in the region affected by the statistical fluctuation. A stretched exponential [8] does not fit well large p values. The expression $D(p) \propto p^{-a} e^{-ap}$ with $a = 0.06$ for $\tau = 2$ and $a = 0.01$ for $\tau = 10$ fits well intermediate and high p values. Since the argument of the exponential function is small and a pure power-law fitting, all over the p range, leads to $R^2 > 0.99$ we write simply $D(p) \propto p^a$.
 - [8] J. Laherrere and D. Sornette, *Eur. Phys. J.* **B2**, 525 (1998).
 - [9] O. Kinouchi and C.P.C. Prado, *Phys. Rev. E* **59**, 4964 (1999).
 - [10] For $\tau = 0$ only 2-cycles exist: $D_0(2) = 0.333 \pm 0.001$ for $d = 1$, and $D_0(2) = 0.31 \pm 0.01$ for $d = 2$.

version of normal (if premature) ageing than does dyskeratosis congenita. In people with dyskeratosis congenita and in telomerase-deficient mice, it is tissues that normally express telomerase that one would predict to suffer most from its loss, and this proves to be the case.

People with dyskeratosis congenita, as well as late-generation telomerase-deficient mice, also suffer from a higher rate of cancer. This can likewise be explained by the lack of telomerase, which results in unstable chromosomes — in dyskeratosis congenita sufferers and the mutant mice, many chromosomes fuse end-to-end^{8,9}, probably because their telomeres are terminally eroded (discussed in ref. 10).

In short, the symptoms of this disease provide a glimpse of the effects of telomerase defects on the maintenance of human tissues. Cells in rapidly dividing tissues, with progenitors that usually express telomerase, are more strongly affected. It will be inter-

esting to see whether Werner's syndrome and related premature-ageing disorders affect the maintenance of telomeres in tissues that do not have telomerase-expressing progenitors.

Robert Marciniak is in the Departments of Medicine and of Cellular and Structural Biology, University of Texas Health Science Center, San Antonio, Texas 78229, USA.

Leonard Guarente is in the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

e-mails: marciniak@oncology.uthscsa.edu
leng@mit.edu

1. Heiss, N. S. *et al.* *Nature Genet.* **19**, 32–38 (1998).
2. Mitchell, J. R., Wood, E. & Collins, K. *Nature* **492**, 551–555 (1999).
3. Vulliamy, T. *et al.* *Nature* **413**, 432–435 (2001).
4. Bodnar, A. *et al.* *Science* **279**, 349–352 (1998).
5. Blasco, M. A. *et al.* *Cell* **91**, 25–34 (1997).
6. Rudolph, K. L. *et al.* *Cell* **96**, 701–712 (1999).
7. Herrera, E. *et al.* *EMBO J.* **18**, 2950–2960 (1999).
8. Artandi, S. E. *et al.* *Nature* **406**, 641–645 (2000).
9. Dokal, I. *Br. J. Haematol.* **110**, 768–779 (2000).
10. de Lange, T. & Jacks, T. *Cell* **98**, 273–275 (1999).

Statistical physics

The salesman and the tourist

H. Eugene Stanley and Sergey V. Buldyrev

Solutions to optimization problems, such as that faced by the travelling salesman, have many practical applications. Might a related problem offer insight into the behaviour of foraging animals?

The travelling salesman problem is of fundamental interest to mathematicians and physicists, and has a number of practical applications, such as computer design. At first sight, the related travelling tourist problem might seem to be of little interest to anyone except perhaps quarrelling travelling companions.

What are these two problems? The travelling salesman problem is the question of what is the shortest possible route followed by a salesman assigned to visit N clients spread across a huge territory; it is a classic example of a class of problems called global optimization problems. Because the salesman's next step depends on his knowledge of the location of every client yet to be visited, the solution to the travelling salesman problem is far from trivial, and has attracted some of the best minds^{1–3}.

In *Physical Review Letters*, Lima *et al.*⁴ introduce a new example from a class of problems known as local optimization problems, and name it the travelling tourist problem. This is the question of what is the best path to be followed by a tourist not constrained to only one visit per city, but constrained — say by his budget — to visit the nearest city. Note that the tourist wants to minimize only the distance to the next city, a local optimization, not the sum of all the distances along the trajectory as in the

travelling salesman problem. So, at first sight, the travelling tourist problem would seem to be trivial.

Nonetheless, even a trivial-sounding problem can yield rich behaviour. For example, the random-walk problem considers a walker limited to the vertices of a square grid. At each step the walker moves to the neighbouring north, south, east or west vertex depending on the throw of a four-sided die. Clearly the walker does not need to know more than his immediate environment. Even so, his behaviour cannot be predicted in many cases, such as when he

cannot revisit a vertex, or if the grid has some randomly scattered sites that can trap the walker.

Similarly, the travelling tourist problem exhibits some surprising behaviour. For example, after visiting a few cities the tourist will visit city A, whose nearest neighbour, city B, has no other neighbour closer than A. So as soon as the tourist visits either A or B, he will thereafter simply oscillate between them, in what is called a two-city limit cycle. Such two-cycles act as effective 'tourist traps' with the tourist going back and forth, like Sisyphus, for eternity.

Lima *et al.*⁴ enrich the travelling tourist problem by assuming that the idealized tourist is more like a real tourist and does not wish to visit the city he went to last, so they introduce a rule that prevents this. The same sort of logic applies as before, except that now the tourist can be trapped in a three-cycle (three cities that form a triangle A-B-C such that a tourist starting at A will pass to B and then to C, before returning to A). Surprisingly, in addition to three-cycle traps, there are also p -cycle traps, where p is as large as desired.

This simple rule can be further generalized if the tourist cannot visit a city already visited in the previous V visits (in the two previous cases, $V=0$ and 1, respectively). Remarkably, after a trip exploring several cities, the tourist is always trapped into repeatedly revisiting a subset of them. Indeed, the entire set of N cities can be partitioned into a large number of 'free cities' that the tourist can enter and then leave, and a number of tourist traps, in which he is condemned to revisit repeatedly the same subset of p cities (Fig. 1).

So, at first sight, the travelling tourist confronts a landscape not so different from that confronted in the random-walk problem with traps. But in the random-walk problem, the traps are typically assumed to have sizes that follow a normal distribution centred around a typical size. In the travelling tourist problem, the p -cycle traps can have a range of p -values, so the number of

Figure 1 The travelling tourist problem. An example in which the tourist is allowed to visit any one of 1,600 random 'cities', subject only to the condition that he does not revisit a city visited in the previous V visits, where $V=10$. As well as a great number of free-standing cities to which the tourist can come and go, there are 378 cities that belong to one or more p -cycles, which serve as 'tourist traps' by constraining the tourist to visit the same set of cities over and over again. The 29 traps shown here occur in a range of sizes and shapes. Lima *et al.*⁴ show that the distribution of these p -cycles follows a power-law function. Such a distribution is normally associated with random processes, rather than the completely deterministic walk followed by the tourist.

Figure 2 A wandering albatross looking for food. This bird may follow a stable migratory route — it visits the same sites (typically feeding areas) over and over again in much the same way as the hapless tourist in Fig. 1. (Reproduced with permission of the British Antarctic Survey.)

traps of a given size might be expected to follow a normal distribution. But Lima *et al.*⁴ find that the number of p -cycles follows a monotonically decreasing power-law function of p , with an exponent of about 2.7. This is a surprising result because power laws are typically associated with systems dominated by random processes, but here the walk is completely deterministic. The power law must somehow arise from the initial random distribution of the cities, so it is intriguing to think there may be deep similarities between this problem and typical power-law systems⁵.

There are many examples of trivial-sounding problems having unexpected applications; for example, random-walk problems can describe a rich range of diffusive motion⁶, and are even used as a crude model of price movements on the stock market⁷. Accordingly, the travelling tourist problem could apply to some interesting questions. Some animals such as birds follow migratory routes in which they revisit an identical set of feeding spots in the same sequence (Fig. 2). These animals may not be altogether unlike the tourist, choosing deterministically the closest feeding ground to where they are, and conditioned by the fact that if they revisit the same feeding area without waiting a certain time the food supply will not have had the chance to rejuvenate itself. So the travelling tourist might provide new insights into the foraging behaviour of insects, mammals and birds, which can follow power-law functions^{8,9}.

Many humans tend, after some time, to fall into a rut — to revisit the same sequence of cities on their travels, to revisit the same sequence of restaurants in their social life, and even for some scientists to revisit the same set of ideas in their scholarship — all actions seemingly occurring as if by magic when in fact there may be a simple underlying reason. Indeed, this type of behaviour is more like the travelling tourist problem than the travelling salesman problem. Normally we think of our behaviour as emerging

from free will. It is possible that some behaviour patterns are the result of applying deterministic rules to a stochastic (random) envi-

Plant hormones

Transporters on the move

Mark Estelle

The hormone auxin is moved across plant cells by transporters called efflux carriers. It now looks as if these carriers behave much more dynamically than had been thought.

If you're a plant you cannot get by without auxin, for it regulates various essential aspects of growth and development. In their paper on page 425 of this issue¹, Geldner *et al.* provide a surprise for biologists studying auxin transport. But their results will also be of wider interest, for they bear more generally on the questions of how cell polarity is established and maintained, and on protein trafficking in plant cells.

Auxin is a hormone that operates through its effects on cell division and elongation. It is transported through files of cells by a process that is thought to depend on the asymmetric distribution of auxin 'efflux carriers' acting at a cell's plasma membrane. The importance of this transport system has been amply demonstrated using synthetic inhibitors of polar auxin transport, which were believed to interfere specifically with the action of the efflux carrier. Geldner *et al.* now show that the efflux carriers cycle rapidly between the plasma membrane and an as-yet-unspecified cellular compartment, and that the inhibitors prevent this cycling. Further, they find that auxin-transport inhibitors are not specific: they have a general effect on the transport of membrane proteins.

Auxin is produced in young organs, primarily near the plant apex. It is then transported to other parts of the plant by influx and efflux carriers, located at the plas-

ma membrane, perhaps not altogether unlike the rules governing the journey of the travelling tourist.

H. Eugene Stanley and Sergey V. Buldyrev are at the Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA.

e-mails: hes@bu.edu

buldyrev@bu.edu

1. Kirkpatrick, S., Gelatt, G. D. & Vecchi, M. P. *Science* **220**, 671–680 (1983).
2. Baskaran, G., Fu, Y. & Anderson, P. W. *J. Stat. Phys.* **45**, 1–25 (1986).
3. Mézard, M. & Parisi, G. *Europhys. Lett.* **2**, 913–918 (1986).
4. Lima, G. F., Martinez, A. S. & Kinouchi, O. *Phys. Rev. Lett.* **87**, 010603 (2001).
5. Stanley, H. E. *Rev. Mod. Phys.* **71**, S358–S366 (1999).
6. Ben Avraham, D. & Havlin, S. *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge Univ. Press, 2000).
7. Mantegna, R. N. & Stanley, H. E. *Introduction to Econophysics: Correlations & Complexity in Finance* (Cambridge Univ. Press, 2000).
8. Viswanathan, G. M. *et al. Nature* **401**, 911–914 (1999).
9. Viswanathan, G. M. *et al. Nature* **381**, 413–415 (1996).

ma membrane, which move the hormone through files of cells by successive rounds of transport into and out of these cells². The key to this mechanism is the asymmetric distribution of the efflux carrier. For example, auxin in the stem is transported in a highly directional fashion from the top towards the base of the plant. The PIN1 protein is a member of a large family of putative efflux carriers, and in the model plant *Arabidopsis* (and by implication all plants) it is required for this polar transport of auxins. Various studies in different plant organs found, in each case, that the efflux carrier is located on the downstream side of the transporting cell^{3–5}.

Our understanding of auxin transport comes largely from studies using artificial inhibitor compounds^{2,6}. Two of the best-characterized are called TIBA and NPA. These compounds reduce polar auxin transport in stem segments by blocking cellular auxin efflux, implying that they specifically interfere with the efflux carriers. The biological effects of transport inhibition are diverse and dramatic. Treatment of developing embryos with NPA causes defects in pattern formation and behaviour of the meristems, the groups of stem cells that give rise to all postembryonic structures⁷. Later in development, the inhibitors prevent proper growth of roots and shoots in response to environmental cues such as light or gravity⁸. Bio-

Deterministic walks in random networks: an application to thesaurus graphs

O. Kinouchi A. S. Martinez G. F. Lima G. M. Lourenço

Departamento de Física e Matemática, FFCLRP, Universidade de São Paulo, Av. Bandeirantes 3900, CEP 14040-901, Ribeirão Preto, SP, Brazil.

S. Risau-Gusman

Zentrum für Interdisziplinäre Forschung, Wellenberg 1, D-33615 Bielefeld, Germany

Abstract

In a landscape composed of N randomly distributed sites in Euclidean space, a walker (“tourist”) goes to the nearest one that has not been visited in the last τ steps. This procedure leads to trajectories composed of a transient part and a final cyclic attractor of period p . The tourist walk presents a simple scaling with respect to τ and can be performed in a wide range of networks that can be viewed as ordinal neighborhood graphs. As an example, we show that graphs defined by thesaurus dictionaries share some of the statistical properties of low dimensional ($d = 2$) Euclidean graphs and are easily distinguished from random link networks which correspond to the $d \rightarrow \infty$ limit. This approach furnishes complementary information to the usual clustering coefficient and mean minimum separation length.

1 Introduction

We live in a world formed by networks: biological, social, linguistic and technological. The study and characterization of such networks have boosted recently by the emergence of new ideas, increasing network databases and available computational power to test models and link them to data [1].

* *Permanent Address:* Escola Técnica Federal de Mato Grosso, R. Zulmira Canavarros 95, CEP 78005-390, Cuiabá, MT, Brazil.

Email addresses: osame@dfm.ffclrp.usp.br (O. Kinouchi), amartinez@ffclrp.usp.br (A. S. Martinez), sebastian.risau-gusman@uni-bielefeld.de (S. Risau-Gusman).

Networks form the substrate where dynamical processes can occur, as the spreading of diseases or diffusion of information. Such processes are usually studied by using stochastic dynamics and random walks. Navigation processes and exploratory behavior [2,3] can also be modeled by walks inside graphs. Of course, navigation processes are not purely random nor purely deterministic. However, as a first step, it would be interesting to know the generic properties of deterministic navigation inside networks [4–6]. If these networks are disordered, deterministic walks will lead to some statistics of trajectories which could provide information about the network topology and kind of disorder. Consider the following examples of deterministic walks in the context of translation machines and thesaurus graphs.

Automatic translation software presents various problems due to the difficult task, even to humans, of converting phrases from one language to another. For example, it seems a desirable feature that if a translated sentence is translated anew to the original language one should get the original phrase. However, the iteration of the translation process in standard translator software frequently produces a drift from the original sense. After a transient, this drift achieves a two-cycle that constitutes a pair of usually nonsensical sentences, which are the same when translated back and forth.

At the word level, a similar phenomenon occurs in analogical dictionaries or thesaurus. Starting from a random word, if one iterates the process going to the nearest synonymous in some sense (say, for example, the first word in the given list), one achieves also cycles of period two as shown in Table 1.

1) link → connection → link
2) translation → conversion → change → alter → change
3) constitution → charter → contract → agreement → accord → agreement
4) constitution → establishment → organization → association → friendship → companionship → company → corporation → business → commerce → trade → deal → contract → agreement → accord → agreement

Table 1

Examples of two-cycle obtained from iteration to the first word in synonymous list (Microsoft Word 98): 1) zero transient trajectory; 2) trajectory with two transient steps; notice that trajectories may differ in 3) United Kingdom English and 4) USA English thesaurus.

This iterative procedure converges, after some (sometimes large) transient time, to a two-cycle. Two-cycles are the universal attractors of this iteration process. This can be easily understood if one thinks about words in a thesaurus as nodes of a graph. This graph is not purely random but their nodes are semantically or etymologically linked defining some ordinal relationship in the node neighborhood (first neighbor, second neighbor etc.). Two-cycles appear

when two mutually next neighbors are finally found.

An interesting statistical question concerns the average degree of separation between two words in a thesaurus. After some experimentation in standard electronic thesaurus, one finds that the average degree of separation is of order $\log N$, where N is the number of nodes (words) in the graph defined by a thesaurus. One also observes that synonymous have large clustering coefficient. This means that the probability that two synonymous of a word are also synonymous of each other is large. This suggests that a thesaurus is an instance of the so-called Small World networks [7]. This fact has been confirmed recently through exhaustive quantitative measures in the Merrian-Webster dictionary [8], Roget's thesaurus [9] in the Wordnet Database [9,10] and even for free word associations database [9].

Here we suggest a different point of view to probe the structure of these highly non-trivial networks, through the study of the statistics of a simple deterministic walk on them: the "tourist walk" [6]. Starting from an initial site, a walker (tourist) moves at each time step according to the simple deterministic rule: go to the nearest site that has not been visited in the preceeding τ time steps. The tourist performs a deterministic partially self-avoiding walk with memory window of τ steps. This iterative dynamics always leads to a trajectory that is composed of an initial transient part and a final attractor (a cycle of period $p \geq \tau + 2$) that traps the tourist. Here, τ is the limited tourist memory range or, alternatively, it can be thought as a refractory time of the sites. For $\tau = 0$, only cycles with period 2 appear, which correspond to a pair of mutually nearest neighbors points, corresponding to the two-cycle phenomena described earlier.

The main result presented in this paper is that deterministic walks in spatial graphs (graphs whose nodes lie in Euclidean d -dimensional space) present some interesting statistical regularities. These regularities could serve as a benchmark when studying real world networks. In Section 2 we show that the tail of the distributions $P_\tau(p)$ of cycle period p , i.e., the number of different p -cycles divided by the total number of cycles, can be described in the whole τ -range by a limited power-law. In this way all curves $P_\tau(p)$ can be collapsed into a single universal curve independent of τ . This is an extension and improvement to the description given in Ref. [6]. Also, for $\tau = 0$, an analytical support, given by Cox's formula [11–14], is presented along the text. In addition to the quantities studied in Ref. [6], another quantity is examined: the probability $P_A(p)$ of a random site to belong to a p -cycle attractor basin. This probability is interesting since it shares the scaling properties of $P_\tau(p)$ and is easily measurable from experimental data. These quantities have been studied as a function of the dimensionality d of the system where we show a slow convergence to the random link network. The random link network corresponds to the limit $d \rightarrow \infty$ and, for $\tau = 1$, the cycle distribution behaves as $P_\tau(p) \propto p^{-1}$.

In Section 3, we compare our results, obtained in low dimensional Euclidean random graphs, to those obtained in the Small World graphs of thesaurus and find surprising similarities between them and huge differences from the $d \rightarrow \infty$ limit. This suggests that the short range connectivity structure of thesaurus graphs could be embedded in a low dimensional space, with the Small World behavior given by a small fraction of long range connections. This contrasts to standard models which represent words in high dimensional Euclidean space [9,15]. Finally in Section 4 concluding remarks are presented.

2 Deterministic Walks

The construction of the d -dimensional Euclidean random graphs starts by randomly distributing the coordinates of $j = 1, 2, \dots, N$ sites with a uniform density ρ in the interval $[0, 1]$ in d dimensions. The Euclidean matrix of distances D_{ij} is used for ordering the neighbors and to create a neighborhood table $V_{ik} = j(k)$, where $j(k)$ is the site that is the k^{th} nearest neighbor of site i .

Notice that a random collection of points in Euclidean space does not define a graph because a set of links is also necessary. If all the points are connected by links, we would have a totally connected graph. We will represent this graph by G_{N-1} , where $N - 1$ is the number of outgoing links. But with a memory window τ , the walker dynamics is well defined if one knows what are the $\tau + 1$ nearest neighbors. Moves to the other neighbors never occur. This means that the walk is done inside a subgraph $G_{\tau+1}$ formed by linking directionally each site i to its $\tau + 1$ nearest neighbors. The tourist walk is always performed inside this directed subgraph that only presents an ordered set of neighbors V_{ik} ($k = 1, \dots, \tau + 1$) for each node. For example, in the G_1 subgraph each node is connected only to its nearest neighbor: indeed this graph is composed by several disconnected parts, one for each pair of mutual nearest neighbor. In the G_2 subgraph, each node is linked to its two nearest neighbors. Such graph has a giant component (percolating cluster) containing more than 98% of the nodes [16]. Notice that G_n is always a subgraph of G_{n+1} .

We stress that the Euclidean metric has been introduced in our model only as a simple mode of ranking the sites neighbors. As pointed out in Ref. [6], the dynamics is performed on the neighborhood table V_{ik} , not in the distance matrix D_{ij} .

2.1 Walks without memory

The density of cycles $D_\tau(p; d)$ is the number of different p -cycles divided by N for a given memory τ and dimension d . As observed before, in the no memory situation $\tau = 0$, only 2-cycles appear. Nearest neighbor statistics in Poisson processes on Euclidean spaces are well known and have been used more intensively after the work of Clark and Evans [11], followed by Refs. [12–14]. They have defined “reflexive nearest neighbors” as two sites that are the nearest neighbors of each other. These reflexive neighbors are precisely the attractors in the $\tau = 0$ tourist walk.

Applying Cox’s formula [14], the density $D_0(2; d)$ of 2-cycles in dimension d is given by:

$$D_0(2, d) = \frac{P_{rn}(d)}{2} = \frac{1}{2(1 + p_d)} , \quad (1)$$

$$p_d = 2 \frac{\Gamma(d/2 + 1)}{\sqrt{\pi}\Gamma[(d + 1)/2]} \int_0^{1/2} dx (1 - x^2)^{(d-1)/2} , \quad (2)$$

with $P_{rn}(d)$ being the fraction of reflexive neighbors and $\Gamma(z)$ the gamma function.

These formula can be worked out to obtain standard functions. Observing that:

$$\frac{\Gamma(d/2 + 1)}{\sqrt{\pi}\Gamma[(d + 1)/2]} = \frac{1}{B[1/2, (d + 1)/2]} , \quad (3)$$

where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$ is the beta function, one obtains:

$$p_d = I_{1/4}\left(\frac{1}{2}, \frac{d + 1}{2}\right) , \quad (4)$$

with $I_z(a, b)$ being the incomplete beta function. Some special values of these quantities are shown in Table 2.

2.2 Walks in $d = 2$

We consider now the dependence on the memory τ in the two dimensional case. For $\tau > 0$, different period cycles coexist. Examples of cycles appearing for a landscape with $N = 400$ sites and several values of memory τ are given

d	p_d	$D_0(2, d)$
1	1/2	1/3
2	$(2\pi + 3^{3/2})/(6\pi)$	$3\pi/(2\pi + 3^{3/2})$
3	11/16	8/27
$\vdots$	$\vdots$	$\vdots$
∞	1	1/4

Table 2

Some values of p_d and $D_0(2, d)$ as a function of d .

in Fig. 1, where open boundary condition has been used here. As τ grows, one observes the progressive disappearance, development or coalescence of cycles.

A natural question concerns the distribution of cycle periods

$$P_\tau(p) = \frac{D_\tau(p)}{\sum_p D_\tau(p)} , \quad (5)$$

and also if there is a value for τ where a percolation phenomenon occurs, i.e., the appearance of a cycle with diameter comparable to the system. However, the detailed study of this percolating regime demands strong computational effort, which is far from our resources. Here we only report results for the also interesting regime $\tau = O(\log N)$, which is of the order of the mean minimum separation length in Small Word networks. This regime gives information about the clustering properties of the network at scale of $O(\tau)$ cities.

Numerical simulations, using periodic boundary conditions, have been performed for $d = 2$ and τ varying from 1 to 10 with $N = 10^4$ and averaging over 100 landscapes, see Fig. 2a. We have found that, for $p > 2p_{min} = 2(\tau + 2)$, the probability of appearance of a p -cycle $P_\tau(p)$ can be fitted by the expression:

$$P_\tau(p) = C(\tau) p^{-\alpha} e^{-[p/p_0(\tau)]^2} , \quad (6)$$

The fitting function gives $\alpha \sim 2.6 \pm 0.1$ independent of τ . We observed that the cutoff length scales as $p_0 \propto \tau$, suggesting that all these curves could be scaled into a single universal function which depends only on the ratio p/τ . As the areas under the $P_\tau(p)$ curves are always equal to one, we have found the data collapse on the function:

$$G(p/\tau) = \tau P_\tau(p) , \quad (7)$$

as illustrated in Fig. 2b. It is worthwhile mentioning that other scalings as $(p - p_{min})/p_{min}$ or p/p_{min} with $p_{min} = \tau + 2$ have been tested and do not work

Fig. 1. Examples of emergent cycles for $d = 2$ and $N = 400$. The memory values are: $\tau = 1, 3, 5$ and 9 . The same landscape has been used for each plot with open boundary condition. The growing, distabilization and fusion of cycles can be observed as a function of the memory τ . Points not belonging to a cycle are not displayed.

as well as the one proposed above. A similar scaling does not hold for $d = 1$ since in this case the exponent α depends strongly on τ , see Ref. [6].

2.3 High Dimensional Walks

For $d > 1$, the $P_\tau(p)$ curve slowly converges to the $d \rightarrow \infty$ behavior, which is also the behavior of the random link network (Fig. 3a), which neglects the correlations among the distances between sites in a d -dimensional Euclidean

Fig. 2. a) Distribution of cycles periods $P_\tau(p)$ for $d = 2$ and different values of τ (from left to right, $\tau = 1, 2, 3, 4, 5, 6, 10$). Notice that $P_\tau(p)$ is defined only for integer p (as in the $\tau = 1$ curve) nevertheless lines have been used for better visualization. b) Data collapse using Eq. 7, $G(p/\tau) = \tau P_\tau(p)$ versus p/τ for $\tau = 1, 2, 3, 4, 5, 6, 10$ (from right to left).

space [17–19]. The distribution tail has the form

$$P_\tau(p) \propto p^{-\alpha(d)} \phi(p, \tau, N) , \quad (8)$$

with $\alpha(d)$ varying from $\alpha(2) \approx 2.6$ to $\alpha(\infty) = 1$ and $\phi(p, \tau, N)$ being a cutoff function.

2.3.1 Convergence to the random link behavior

The random link network has been generated numerically considering the distances between two points as uniform random variables in the interval $[0, 1]$. The distribution $P_\tau(p)$ for the random link network can be described by a power law p^{-1} (see Fig. 3a) with a cutoff which grows with the system size N , since it is a version of a random map model [20]. Notice, however, that we have studied a symmetrical situation ($D_{ij} = D_{ji}$) with $\tau = 1$. The Derrida random map [21] corresponds to an assymetrical random link case ($D_{ij} \neq D_{ji}$), it presents cycle periods $p \geq 2$ even for $\tau = 0$ and the total number of attractors is of order $\log N$. For $\tau = 0$, the symmetric case presents only two-cycles and is the correct approximation for the limit $d \rightarrow \infty$, with the number of attractors scaling as $0.25N$.

The slow convergence to the $d \rightarrow \infty$ behavior also appears in other statistical quantities as, for example, the probability $P_A(p)$ that a random initial site leads to a cycle- p (Fig. 3b). When studying deterministic walks in real world networks, the quantity $P_A(p)$ is more convenient from an experimental point of view. One can obtain better statistics for $P_A(p)$ since each different initial condition gives valid data and it is not necessary to keep track of only different cycles, as stated by the definition of $P_\tau(p)$.

3 Thesaurus graphs

To exemplify the use of the tourist walk in ordinal ranking problems, we have considered $\tau = 0$ and $\tau = 1$ walks in a graph defined by a thesaurus. In this situation a word is chosen at random and uniformly from a standard dictionary. The chosen word is then placed in an electronic thesaurus (the USA English Microsoft 98 thesaurus). In the considered thesaurus, the synonyms are ranked according to their use frequency.

3.1 No memory

For a $\tau = 0$ walk, the more frequent used synonym of a given word (the first word in the synonymous list or “first neighbor”) is then chosen as a new input and, again, its first neighbor is chosen in an iterative process, as illustrated in Table 1. This procedure leads to a transient and (generally) to a cycle of

Fig. 3. a) Distribution of cycle periods $P_\tau(p)$ where the dotline represents the $P_\tau(p) \propto p^{-1}$ fit. b) Distribution of cycle periods $P_A(p)$. Curves are for $\tau = 1$ and from downwards to upwards with: $d = 2, 4, 8, 16, 32$ and symmetric random link. Other values are: $N = 10^3$ points averaged for 10^4 landscapes with periodic boundary conditions.

period 2, where two nearest reflexive synonyms are found. The distribution of transient times presents an exponential decay (Fig. 4).

The proportion of reflexive neighbors could be estimated in principle by measuring the fraction of initial words that are already in a two-cycle (like the

word “link” in Table 1). From 1000 initial words ($\sim 2\%$ of the whole dictionary), 272 of them are found to belong to two-cycle. This value leads to $P_{rn} = 0.272 \pm 0.028$, with a confidence level $\gamma = 95\%$, which is very far from the range $[1/2, 2/3]$ expected for random points in d -dimensional Euclidean space. However, it is clear from the distribution of the transient times (Fig. 4) that this number has been underestimated. One of the reasons for this underestimation is that a large fraction of two-cycles are composed by cycles containing two word expressions as: “**punctual** - on time - **punctual**”. Since we have started always from single word expressions, the cycle: “**on time** - **punctual** - **on time**” will never be caught as a transient zero exemplar. Another source for this discrepancy may be the presence of a small fraction of random links (see below). Extrapolating the exponential decay in transient to $t = 0$ leads to a better estimate $P(t = 0) = P_{rn} = 0.618$, which is compatible with a low dimensional $d = 2$ (or maybe $d = 3$) estimate by Cox’s formula $P_{rn}(2) = 0.62$ ($P_{rn}(3) = 0.59$).

We shall also mention that some cycles of higher period have been found in the dictionary experiment: $(5.2 \pm 1.4)\%$ of three-cycles and $(0.5 \pm 0.4)\%$ of four-cycles with confidence level $\gamma = 95\%$. Thus, the examined dictionary may present links where B is the nearest synonymous of A , C is the nearest synonymous of B and A is the nearest synonymous of C . This kind of relation cannot be embedded in a metric space (due to the triangle inequality), suggesting that this thesaurus graph is slightly non-metric. We do not know if this is either a significant or desirable property, or an artifact from the specific dictionary examined.

3.2 One step memory

For $\tau = 1$ walks, we have found that the probability $P_A(p)$ of a word to belong to the basin of a p -cycle is very similar to that from walks in low dimensional Euclidean spaces, as shown in Fig. 5. This result is amazing and suggests that the lexicon graph could be embedded in a low dimensional Euclidean space preserving most of its neighbor order.

Notice that we have worked only with the G_1 ($\tau = 0$) and G_2 ($\tau = 1$) subgraphs, that is, graphs containing only one or two nearest neighbors of each point. So, we are examining the local structure of the network, not the large distance links which give the Small World behavior or power law degree of connectivity [9,10]. We have found that these subgraphs are very similar to those generated by a random collection of points in Euclidean space with $d = 2$ (or $d = 3$). It must be emphasized that higher values for d are excluded because this would be reflected already in the G_1 and G_2 structure. The $d = 1$ case is also excluded once cycles with periods 5, 6 and 7 have been measured and are

Fig. 4. Distribution of transient times for MS98 USA English thesaurus with $\tau = 0$ and $N = 1000$ initial words. Fitting the points with $P(t) = ce^{-ct}$ leads to the extrapolation $P(0) = c = 0.618$ (open circle), which can be compared to $P_{rn}(2) = 0.6215$.

Fig. 5. Comparison between fraction of sites $P_A(p)$ belonging to p -cycle attractor distribution of cycles for $\tau = 1$ in random Euclidean graphs with $d = 2$ (circles) for $N = 1000$ initial words of MS98 USA English thesaurus (triangles) and random link graphs (squares).

forbidden by the model [6].

This result contrasts to standard models which represent words in high dimen-

sional ($d > 100$) Euclidean space, for instance, the Latent Semantic Analysis (LSA) model [9,15]. If the present thesaurus were represented as points in such high dimensional spaces, the tourist walk behavior should be similar to the random link behavior. Instead, our findings suggest that, for best representing its local behavior, the nodes (words) from the studied thesaurus should be embedded in a low Euclidean dimension, with a small fraction of assymmetric links (to produce cycles with $p > 2$ for $\tau = 0$). Of course, as indicated by other studies [9,10], there are also long range links that give the Small World/Scale Free character of the network.

We have repeated this kind of study with the Portuguese Microsoft Word98 thesaurus. However, the existence of a large fraction of dead ends, i. e., nodes without outgoing links (reflecting a low quality database) prevented further analysis in this dictionary.

4 Conclusion

We have shown that simple deterministic walks inside graphs with ordered neighborhood produce the emergence of cycles with interesting statistical properties. The spectrum of stable cycles $P_\tau(p)$ and the probability $P_A(p)$ of falling in a p -cycle from a random start give complementary information to measures like the average minimal length L and clustering coefficient C . Networks with the same L and C could be distinguished by different spectrum $P_\tau(p)$ and $P_A(p)$.

Walks with no memory ($\tau = 0$) easily detect absence of metrics in the graph. The curves $P_\tau(p)$ and $P_A(p)$ given here define class behaviors for the less informative spatial distribution of nodes (the Poisson case in Euclidean space) that can be used as a benchmark for results from walks in other networks. Since graphs with ranked neighborhood, but random structure, abound in the real world (from neural and ecological to social networks [8]), it could be interesting to find if deterministic walks on them belong, or not, to the same class defined by the Euclidean random graphs studied here. A surprising result is that the graph G_2 generated by a thesaurus is best represented by random points in low dimensional, even $d = 2$, Euclidean space.

Acknowledgements

The authors thank M. Idiart for sharing preliminary results with us, J. F. Fontanari for calling our attention to the random map model, Nestor Caticha,

A. C. Roque da Silva , N. A. Alves and R. S. Gonzalez for fruitful discussions.
O. Kinouchi acknowledges support from FAPESP.

References

- [1] S. H. Strogatz, Exploring complex networks, *Nature* 410 (2001) 268–276.
- [2] G. M. Viswanathan, V. Afanasyev, S. Buldyrev, E. J. Murphy, P. A. Prince, H. E. Stanley, Levy flight search patterns of wandering albatrosses, *Nature* 381 (6581) (1996) 413–415.
- [3] G. M. Viswanathan, S. Buldyrev, S. Havlin, M. G. E. da Luz, E. P. Raposol, H. E. Stanley, Optimizing the success of random searches, *Nature* 401 (6756) (1999) 911–914.
- [4] H. Freund, P. Grassberger, The Red Queens walk, *Physica A* 190 (3-4) (1992) 218–237.
- [5] D. Gale, J. Propp, S. Sutherland, S. Troubetzkoy, Further travels with my ant, *Math. Intelligencer* 17 (3) (1995) 48–56.
- [6] G. F. Lima, A. S. Martinez, O. Kinouchi, Deterministic walks in random media, *Phys. Rev. Lett.* 87 (1) (2001) 010603.
- [7] D. J. Watts, S. H. Strogatz, Collective dynamics of “small world” networks, *Nature* 393 (1998) 440–442.
- [8] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, *cond-mat/0106096* (June 2001).
- [9] M. Steyvers, J. B. Tenenbaum, The large-scale structure of semantic network: statistical analysis and a model of semantic growth, *cond-mat/0110012* (2001).
- [10] M. Sigman, G. A. Gecchi, Global organization of the lexicon, *cond-Mat/0106509* (2001).
- [11] P. J. Clark, F. C. Evans, On some aspects of spatial pattern in biological populations, *Science* 121 (1955) 397–398.
- [12] P. J. Clark, Grouping in spatial distributions, *Science* 123 (1956) 373–374.
- [13] M. F. Dacey, Proportion of reflexive nth-order neighbours in a spatial distribution, *Geographical Analysis* 1 (1969) 385–388.
- [14] T. F. Cox, Reflexive nearest neighbours, *Biometrics* 37 (1981) 367–369.
- [15] T. K. Landauer, S. Dumais, A solution to plate’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge, *Psychol. Rev.* 104 (2) (1997) 211–240.
- [16] M. A. Idiart, private communication (2001).

- [17] J. Vannimenus, M. Mézard, On the statistical mechanics of the optimization problems of the travelling salesman type, J. Phys. Lett. France 45 (1984) L1145–L1153.
- [18] A. Percus, Voyageur de commerce et problèmes stochastiques associés, Ph.D. thesis, Université Paris 6, Paris (Septembre 1997).
- [19] A. G. Percus, O. C. Martin, The stochastic traveling salesman problem: finite size scaling and the cavity prediction, J. Stat. Phys. 94 (1999) 739–758.
- [20] U. Bastolla, G. Parisi, Relaxation, closing probabilities and transition from oscillatory to chaotic attractors in asymmetric neural networks, J. Phys. A: Math. Gen. 31 (20) (1998) 4583–4602.
- [21] B. Derrida, H. Flyvbjerg, The random map model - a disordered model with deterministic dynamics, Journal de Physique (Paris) 48 (6) (1997) 971–978.

Busca por: lima turista
Resultados encontrados: 2

Modelo ajuda a entender o ciclo animal
2002-02-22 08:58:00

Agência USP - Um novo modelo matemático computacional dá pistas para compreensão do comportamento de animais, mesmo os de baixa capacidade cognitiva na procura de alimentos em longas rotas migratórias, como borboletas, peixes e pássaros. O trabalho é resultado de uma pesquisa de doutorado de Gilson **lima**, do Departamento de Física e Matemática da Faculdade de Filosofia, Ciências e Letras da USP de Ribeirão Preto.

Segundo Alexandre S. Martinez e Osame Kinouchi, orientadores de **lima**, o estudo considera problemas de otimização, como os encontrados por um "hipotético **turista** mochileiro", na qual a rota de viagem não é predeterminada, mas obtida minimizando a distância para a próxima cidade a ser visitada. A questão não é descobrir o menor caminho que une todas as cidades a serem visitadas mas saber quais serão visitadas.

Mais informações: (0XX16) 602-3720 ou 602-3719 ou pelo e-mail amartinez@ffclrp.usp.br

Site: <http://www.usp.br/agenciausp/repgs/2002/pags/038.htm>

Turismo e conservação de cavernas
2001-10-08 08:44:00

Agência USP - As cavernas são, cada vez mais, alvo do interesse dos chamados "**turistas** de aventura". O problema é que, em muitos casos, essa busca por mistério e aventura causa danos que chegam a ser irreversíveis, prejudicando sensivelmente os estudos geológicos realizados no interior da terra. "De dez anos para cá, o fluxo desses **turistas** nas cavernas aumentou muito e sem que houvesse controle ou preocupação com o estado de conservação", alerta o geólogo e professor da USP, Ivo Karmann.

Karmann foi um dos palestrantes do evento Cavernas, Turismo e Conservação, organizado pelo Museu de Geociências da USP. Além de palestras, aconteceram mostras de filmes e exposições de materiais colhidos em cavernas. Tudo evidenciando a importância da preservação. "Através das cavernas é possível estudarmos paleoclimas — climas de períodos pré-históricos", atesta Karmann, que ainda cita formações de relevo e fontes de água subterrâneas como elementos relacionados às pesquisas espeleológicas (aquelas realizadas no interior de cavernas).

Entretanto, não se pode afirmar que o turismo de cavernas resulta apenas em prejuízos. "A situação sócio-econômica de muitos moradores do Vale do Ribeira, onde fica o Petar, melhorou bastante", pondera o geólogo. A região é carente e tem sua economia baseada na agricultura de subsistência. "Isso significa que o turismo em cavernas tem de ser cuidadosamente pensado, uma vez que existem realmente benefícios importantes", recomenda. "É um grande número de pessoas desprovidas de algumas condições básicas de sobrevivência que vêem nas cavernas uma oportunidade da qual elas não querem abrir mão", justifica.

Mais informações: (0XX11) 3091-3943 ou e-mail: ikarmann@usp.br

Site: <http://www.usp.br/agenciausp/repgs/2001/pags/076.htm>

[Início](#)

[Conheça o Projeto](#)

[CH](#)

[Revista](#)

[Ciência Hoje](#)

[Ciência Hoje](#)

[das Crianças](#)

[Ciência Hoje](#)

[na Escola](#)

[Ciência em Dia](#)

[Fale conosco!](#)

'Modelo da caminhada do turista' permite otimizar problemas

Algoritmo de físicos da USP descreve padrão em situações aparentemente aleatórias

É possível conhecer as capitais do Brasil fazendo sempre o trajeto mais curto e evitando a repetição de cidades? Um modelo matemático criado na USP talvez traga a resposta a problemas como esse

Parece improvável que haja qualquer relação entre situações tão distintas como as vivenciadas por turistas mochileiros, pássaros migradores ou mesmo internautas navegando pela rede. Embora não pareçam obedecer a nenhum padrão, essas situações apresentam um determinismo que resulta na ocorrência sistemática de ciclos viciosos.

Um modelo matemático capaz de descrever o padrão cíclico em situações aparentemente aleatórias como as descritas acima foi elaborado por Gilson Lima, Alexandre Martinez e Osame Kinouchi, físicos da Universidade de São Paulo (USP) em Ribeirão Preto. O objetivo dos pesquisadores é aplicar à resolução desses problemas as chamadas regras de otimização local para, assim, encontrar a melhor maneira de transpor os ciclos viciosos e reduzir ao máximo custos, imprevistos etc. O estudo foi publicado na *Physical Review Letters* em junho de 2001 e posteriormente comentado na revista *Nature*.

Imagine um turista mochileiro em excursão pelo Brasil, que pretende partir para a capital mais próxima de onde estiver. É importante que ele visite o máximo de capitais e que o trajeto entre duas cidades seja o menor possível, independentemente da rota final. Em um momento, o turista vê-se diante de um impasse: a capital mais próxima é também a última que visitou. Formou-se um ciclo vicioso. Nosso turista -- um ser pensante -- é capaz de evitar esses ciclos, pois leva em conta todas as últimas M capitais visitadas. Porém, em situações similares, é fatal que se caia em ciclos viciosos, independentemente do valor de M.

Os casos estudados pelos pesquisadores podem ser classificados como sistemas com dinâmica determinística. Porém, o modelo resulta em funções matemáticas que definem sistemas aleatórios (leis de potência). "A estatística obtida para a distribuição dos períodos nos ciclos viciosos é praticamente uma lei de potência", conta Martinez. "Esse fato é bastante surpreendente e precisa ser melhor explicado."

Chamada de 'modelo da caminhada do turista', o algoritmo elaborado pelos físicos aponta a periodicidade dos ciclos viciosos, a distribuição dos períodos e a probabilidade de eles ocorrerem. "O problema é que toda otimização determinística leva a ciclos", pondera Martinez. Os pesquisadores estudam agora uma possibilidade de quebra dos ciclos.

"Observamos que em um pasto fornido ao acaso de montes de alimentação, 75% das ovelhas se dirigem ao monte mais próximo e 25% a um outro monte qualquer. Isso significa que as ovelhas quebraram naturalmente o ciclo. Conseguimos fazer com que o modelo previsse essa quebra de modo natural, na proporção observada experimentalmente", comemora. É preciso saber se foi uma coincidência ou se o modelo realmente é capaz de prever fenômenos como esses.

Martinez ainda não tem idéia dos benefícios que o novo algoritmo possa trazer. "O modelo é como os personagens do dramaturgo italiano Pirandello, que estão em busca de uma história depois que seu autor recusou-se a continuá-la. Estamos em busca de um problema que possa ser resolvido somente pelo modelo", brinca.

Paula Americano
Ciência Hoje on-line
01/04/02

pg. 1/1