

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE FÍSICA

# Aquisição e otimização de mapas de navegação usando redes neurais.\*

**Marcelo Trevisan**

Dissertação realizada sob orientação do professor Dr. Marco A. P. Idiart e apresentada ao Instituto de Física da UFRGS em preenchimento parcial dos requisitos para a obtenção do título de Mestre em Física.

Porto Alegre

1<sup>o</sup> semestre de 2001.

---

\* Trabalho financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Este trabalho é dedicado

à minha família.

# Agradecimentos

★ Ao orientador prof. Marco A. P. Idiart, pela dedicação demonstrada e pelo estímulo constante oferecido e, acima de tudo, pela amizade e compreensão durante este período de trabalho.

★ Aos meus pais: Isaltino e Cladis e irmãos: Rafael, Patrícia e Gabriela pelo estímulo e ajuda permanente.

★ Aos meus tios: Orildo e Lourdes pelos churrascos e as cervejas de domingo.

★ A minha vó: Maria Delícia.

★ Aos meus colegas e amigos da sala M204 e das outras salas que contribuíram para a agradável convivência nestes dois anos no IF.

★ Aos professores e funcionários do IF, que de alguma forma contribuíram para a minha formação e para a realização deste trabalho.

*Muito Obrigado*

# Resumo

A capacidade de encontrar e aprender as melhores trajetórias que levam a um determinado objetivo proposto num ambiente é uma característica comum à maioria dos organismos que se movimentam. Dentre outras, essa é uma das capacidades que têm sido bastante estudadas nas últimas décadas. Uma consequência direta deste estudo é a sua aplicação em sistemas artificiais capazes de se movimentar de maneira inteligente nos mais variados tipos de ambientes.

Neste trabalho, realizamos uma abordagem múltipla do problema, onde procuramos estabelecer nexos entre modelos fisiológicos, baseados no conhecimento biológico disponível, e modelos de âmbito mais prático, como aqueles existentes na área da ciência da computação, mais especificamente da robótica. Os modelos estudados foram o aprendizado biológico baseado em células de posição e o método das funções potenciais para planejamento de trajetórias. O objetivo nosso era unificar as duas idéias num formalismo de redes neurais.

O processo de aprendizado de trajetórias pode ser simplificado e equacionado em um modelo matemático que pode ser utilizado no projeto de sistemas de navegação autônomos. Analisando o modelo de Blum e Abbott para navegação com células de posição, mostramos que o problema pode ser formulado como um problema de aprendizado não-supervisionado onde a estatística de movimentação no meio passa ser o ingrediente principal. Demonstramos também que a probabilidade de ocupação de um determinado ponto no ambiente pode ser visto como um potencial que tem a propriedade de não apresentar mínimos locais, o que o torna equivalente ao potencial usado em técnicas de robótica como a das funções potenciais. Formas de otimização do aprendizado no contexto deste modelo foram investigadas.

No âmbito do armazenamento de múltiplos mapas de navegação, mostramos que é possível projetar uma rede neural capaz de armazenar e recuperar mapas navegacionais para diferentes ambientes usando o fato que um mapa de navegação pode ser descrito como

o gradiente de uma função harmônica. A grande vantagem desta abordagem é que, apesar do baixo número de sinapses, o desempenho da rede é muito bom.

Finalmente, estudamos a forma de um potencial que minimiza o tempo necessário para alcançar um objetivo proposto no ambiente. Para isso propomos o problema de navegação de um robô como sendo uma partícula difundindo em uma superfície potencial com um único ponto de mínimo. O nível de erro deste sistema pode ser modelado como uma temperatura. Os resultados mostram que superfície potencial tem uma estrutura ramificada.

# Abstract

The capacity of finding and learning optimal paths is common to most mobile organisms. This characteristic has been largely studied in biology in the last decades. A consequence of this is the application of the resulting knowledge in the design of mobile artificial systems capable of navigation in a variety of environments.

In this work we perform a multiple approach of the navigation problem aiming to establish a nexus between physiological models, based in the current biological knowledge, and models of a more theoretical and practical fashion like the ones we find in computer sciences, specifically in robotics. The models studied were learning using biological realistic place-cells and path planning method using potential fields. Our goal was to unify both ideas in a neural network formalism.

The path learning process can be simplified and equated in a mathematical model that can be used to design autonomous navigation systems. Analyzing the Blum and Abbott model of place-cells learning we show that it can be formulated as an unsupervised learning problem where the statistics of motion in the environment plays an essential role. We also demonstrate that the occupancy probability for a given site in the environment can be seen as a potential that does not present navigational local minima and therefore is equivalent to the potential used in the context of robotics methods like the potential field method. We also investigate forms of optimization of place-cell learning.

In the context of multiple navigation maps we demonstrate that it is possible to design a neural network to store and to recover navigation maps to different environments using the fact that a navigation map can be described as a harmonic function gradient. The advantage of this approach is the good performance of the network, in spite of the use of a lower number of synapses.

Finally for the purpose to know how is the shape of the map that minimize the time

required to reach a goal in a environment, we propose the navigation problem as the problem of a particle diffusing on a potential surface that has a single minimum. The level of error was modeled as a temperature noise. The results show that potential surface has a branched structure.

# Conteúdo

<b>1. Introdução</b>	1
1.1 Redes Neurais Artificiais	2
1.2 Memória Espacial	6
<b>2. Aprendizado de mapas de navegação</b>	12
2.1 O modelo de Blum e Abbott	12
2.1.1 Campo receptivo de um neurônio de posição	13
2.1.2 Acoplamento sináptico e aprendizado	14
2.1.3 Codificação da posição	15
2.2 Análise do modelo	17
2.2.1 O alcance do aprendizado	18
2.2.2 Simulações	21
2.3 Conclusões	24
<b>3. Aprendizado não supervisionado de trajetórias</b>	26
3.1 Mapa de navegação	27
3.1.1 O alvo	28
3.2 Contexto e múltiplos mapas	31
3.3 Usando os mapas durante o aprendizado	33
3.4 Modelos de rede	38
3.5 Conclusões	40
<b>4. Utilizando redes neurais para armazenar diferentes mapas</b>	42
4.1 O método das funções harmônicas	43



---

4.2	O modelo da rede . . . . .	44
4.2.1	Rede 1 - Recuperação do contexto . . . . .	45
4.2.2	Rede 2 - Recuperação do ambiente . . . . .	46
4.2.3	Rede 3 - Geração do mapa de navegação . . . . .	47
4.3	Resultados . . . . .	48
4.4	Conclusões . . . . .	49
<b>5.</b>	<b>Potencial ótimo para navegação . . . . .</b>	<b>52</b>
5.1	Teoria geral . . . . .	53
5.2	Problema unidimensional . . . . .	54
5.3	Problema bidimensional . . . . .	57
5.4	Conclusões . . . . .	60
	<b>Conclusões e Perspectivas Futuras . . . . .</b>	<b>62</b>
	<b>A. Alteração do vetor população induzida pelo treinamento . . . . .</b>	<b>65</b>
	<b>B. Princípio do máximo das funções harmônicas . . . . .</b>	<b>67</b>
	<b>Referências Bibliográficas . . . . .</b>	<b>69</b>

## Capítulo 1

# Introdução

A observação e a compreensão dos sistemas biológicos existentes na natureza sempre preocuparam os cientistas de um modo geral. Nas últimas décadas, um grande número de pesquisas vêm sendo realizadas com o objetivo de entender como que os organismos vivos conseguem solucionar, de maneira eficaz, problemas considerados difíceis do ponto de vista de sistemas artificiais. A partir desses estudos é possível conceber sistemas artificiais dotados de capacidades e comportamento semelhantes.

Um dos problemas que tem sido extensivamente estudado diz respeito ao projeto de sistemas artificiais capazes de se movimentar de maneira eficaz e inteligente nos mais variados tipos de ambientes. Esses sistemas, chamados de sistemas de navegação autônomos, devem ser capazes de navegar contornando possíveis obstáculos existentes no meio de maneira autônoma, ou seja, sem a necessidade de interferência humana. Além disso, esses sistemas devem ter a capacidade de encontrar as melhores trajetórias que levam a um determinado objetivo proposto no ambiente.

Mas apesar do grande número de pesquisas realizadas na área, a tarefa de desenvolver sistemas de navegação autônomos ainda é um problema em aberto. Frequentemente, o projeto desses sistemas são baseados em mecanismos observados em organismos biológicos encontrados na natureza [1, 2, 3, 4, 5, 6]. Na verdade, o desenvolvimento de um sistema de navegação autônomo baseados em sistemas biológicos está diretamente relacionado à compreensão de como que ocorre a representação espacial e o aprendizado de trajetórias nestes sistemas. O'Keefe e Nadel [7] observaram que a medida que a complexidade e a capacidade de percepção dos organismos biológicos aumentam, uma representação espacial explícita pa-

rece ser utilizada como um suporte cognitivo para executar a navegação em ambientes mais complexos. Em geral, todas as espécies de animais parecem ter um sistema especializado em representação espacial mais ou menos bem desenvolvido e capaz de executar sofisticados cálculos [8], embora a estrutura neuro-fisiológica responsável por esta representação espacial possa variar de espécie para espécie. Em alguns trabalhos recentes [9, 10, 11] são apresentados modelos matemáticos de sistemas capazes de realizar o aprendizado espacial de trajetórias de maneira autônoma. Esses sistemas são redes neurais artificiais que constroem, de forma não supervisionada, uma representação espacial do ambiente durante o processo de exploração de maneira que, após este período, a representação possa ser utilizada no processo de navegação no ambiente.

Neste trabalho trataremos do aprendizado de trajetórias visto de ângulos diferentes. Inicialmente, faremos um estudo de como que as informações do mundo exterior afetam o comportamento neural de um sistema biológico, mais especificamente dos ratos. Veremos que, a partir disso, é possível criar um modelo matemático capaz de reproduzir algumas características obtidas experimentalmente. Veremos que as características associadas ao ambiente podem determinar a estatística do movimento. Esta estatística permite uma estruturação das informações necessárias para navegar no ambiente de uma forma que possam ser aprendidas por uma rede neural artificial. Posteriormente, faremos uma análise e uma otimização matemática do processo de exploração de um ambiente.

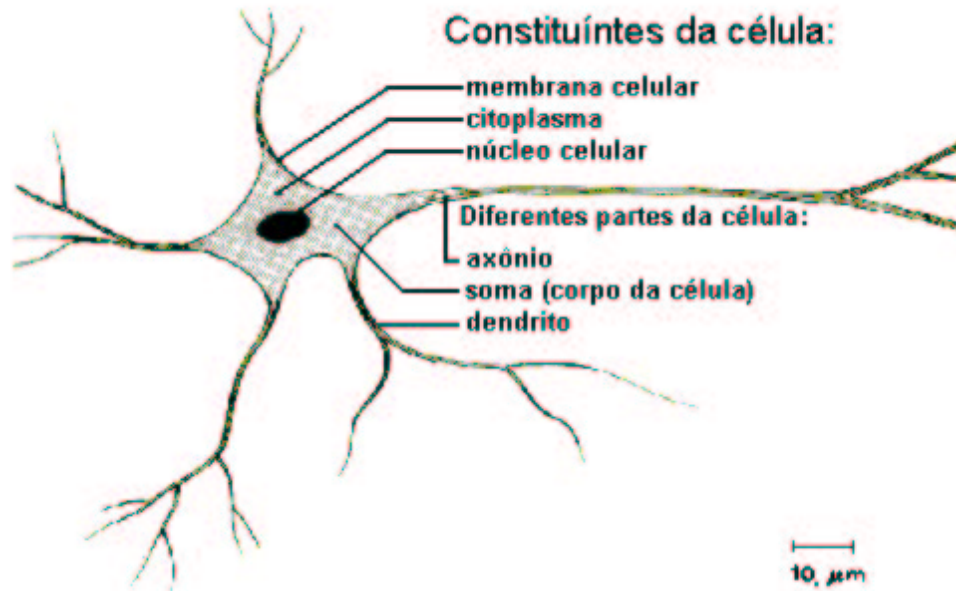
Primeiramente, faremos uma rápida descrição de alguns dos conceitos básicos que envolvem redes neurais artificiais.

## 1.1 Redes Neurais Artificiais

Redes neurais artificiais são técnicas computacionais que usam uma formulação matemática inspirada na estrutura neural de organismos inteligentes que adquirem conhecimento através da experiência. Uma rede neural artificial é composta de várias unidades de processamento chamadas de neurônios. Esta nomenclatura surge da analogia direta às células que compõem o sistema nervoso animal.

O neurônio biológico pode ser dividido em três partes fundamentais: o corpo celular, ou

soma, que é responsável pelo controle das funções celulares essenciais, os dendritos que são prolongamentos somáticos responsáveis pela coleta de informação vinda de outros neurônios e o axônio que é um prolongamento somático especializado no envio de informações para outros neurônios (veja figura 1.1).



**Fig. 1.1:** *Esquema dos constituintes de um neurônio.*

A informação neural é codificada em eventos elétricos chamados de potenciais de ação. Um potencial de ação é uma inversão local da polaridade (depolarização) da membrana que se origina no soma e se propaga pelo axônio. Um neurônio pode produzir potenciais de ação em um intervalo de frequência que vai desde poucos Hz até centenas de Hz. A atividade de um neurônio é caracterizada por esta frequência de disparos.

A passagem de informação entre dois neurônios se dá em regiões especializadas chamadas sinapses. Nas sinapses as membranas celulares dos neurônios pré-sináptico e pós-sináptico se aproximam muito, podendo ficar a distâncias da ordem de 5nm. Com esta proximidade, o neurônio pré-sináptico pode enviar para o neurônio pós-sináptico sinais que indicam seu estado de atividade. Isto pode se dar de duas formas: uma delas puramente elétrica nos chamados *gaps* de junção, onde as membranas efetivamente se tocam, e a outra forma é química, onde o potencial de ação, após percorrer o axônio e chegar à sinapse, causa a

liberação de substâncias químicas (os neurotransmissores) que atuam na membrana pós-sináptica, permitindo a passagem da informação. O neurotransmissor pode diminuir ou aumentar a polaridade da membrana pós-sináptica, inibindo ou excitando a geração dos pulsos no neurônio pós-sináptico. Este processo depende de vários fatores, como a geometria da sinapse e o tipo de neurotransmissor.

Uma rede neural artificial é composta por várias unidades de processamento (neurônios artificiais), cujo funcionamento é bastante simples. Essas unidades recebem informações por conexões sinápticas que estão associados a determinado peso. As unidades fazem operações apenas sobre seus dados locais. O comportamento inteligente de uma rede neural artificial vem das interações entre as unidades de processamento da rede.

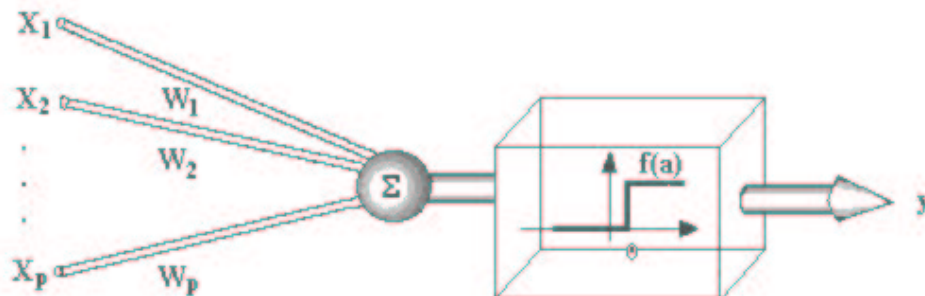
McCulloch e Pitts [12] propuseram um modelo matemático para representar um neurônio de maneira artificial. Este modelo estava baseado em dois pressupostos importantes: primeiro, que a informação neural é binária, ou seja, que o nível de atividade neural não é importante, segundo, que o neurônio processa uma combinação linear dos estímulos pré-sinápticos. Com isto eles propuseram que uma unidade de processamento (figura 1.2) que operava sobre um conjunto de sinais de entrada. Cada sinal era multiplicado por um fator, ou peso, que indicava a influência deste sinal na saída da unidade. Após a multiplicação era feita a soma ponderada dos sinais produzindo um determinado nível de atividade. Se este nível de atividade excedesse um certo limiar a unidade gerava uma resposta de saída. Escrevendo em uma forma matemática temos

$$y_i(t+1) = \Theta \left( \sum_j w_{ij} x_j(t) - \mu_i \right) \quad (1.1)$$

onde  $y_i$  é chamado de estado ou ativação de uma unidade  $i$ . A função  $\Theta(a)$  ou função Heaviside é dada por:

$$\Theta(a) = \begin{cases} 1 & \text{se } a \geq 0; \\ 0 & \text{se } a < 0. \end{cases} \quad (1.2)$$

Com isso  $y_i$  assume valores discretos 1 ou 0 e representa o estado do neurônio  $i$  (disparando ou não respectivamente). O peso  $w_{ij}$  representa a intensidade da conexão sináptica do neurônio  $j$  com o neurônio  $i$ . Esta conexão pode ser negativa ou positiva correspondendo a



**Fig. 1.2:** Diagrama esquemático de um neurônio de McCulloch-Pitts. A unidade dispara se o somatório  $\sum_i w_i x_i$  atingir ou exceder um certo limite.

uma sinapse inibitória ou excitatória. Caso a conexão for zero, significa que não há sinapse entre  $i$  e  $j$ . O parâmetro  $\mu_i$  é o limiar da unidade  $i$ . A soma dos pesos das entradas deve atingir ou exceder o limiar para o neurônio disparar.

Embora simples, o neurônio de McCulloch e Pitts é um poderoso dispositivo. Eles provaram que um conjunto sincronizado de tais neurônios são capazes do princípio universal da computação para apropriadas escolhas dos pesos  $w_{ij}$ . Isto significa que ele pode executar qualquer computação que um computador normal executa, embora não necessariamente tão rápido. Fazendo uma generalização dos neurônios de McCulloch-Pitts temos

$$y_i = f \left( \sum_j w_{ij} x_j - \mu_i \right) \quad (1.3)$$

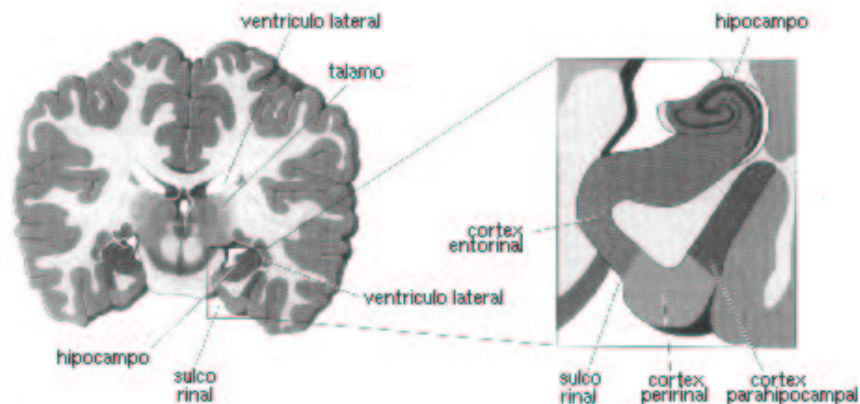
onde  $y_i$  assume valores contínuos e é chamado de estado ou ativação de uma unidade  $i$ . A função  $\Theta(x)$  foi substituída por uma função mais geral  $f(x)$  chamada de função de transferência.

Conforme mencionado na seção anterior, existem redes neurais capazes de construir e armazenar uma representação espacial de um determinado ambiente, de tal maneira que esta representação possa ser utilizada no processo de navegação neste ambiente. Na próxima seção veremos, de uma forma geral, como que ocorre a formação da memória espacial no cérebro dos mamíferos e sua relação com as redes neurais.

## 1.2 Memória Espacial

Nos mamíferos o conhecimento e a noção de espaço são tarefas realizadas pelo cérebro. Estas tarefas incluem a organização de informações de diferentes tipos e origens em estruturas mentais, chamadas de mapas, que guardam a natureza geométrica do espaço e que correspondem a nossa percepção do mundo exterior. A partir de experimentos realizados em roedores [13, 14, 7, 15, 16, 17, 18, 19] e em macacos [20, 21, 22] surgiram evidências de que o hipocampo desempenha um importante papel no processamento espacial [7, 23, 24, 25, 26, 9, 27, 10]. Esta região cerebral está funcionalmente situada na extremidade final de todas as terminações sensoriais e tem, de um grosso modo, as mesmas características anatômicas em todos os mamíferos.

O hipocampo é uma estrutura em forma de dobras situado medialmente ao ventrículo lateral. Na lateral do hipocampo estão situadas três regiões corticais: o córtex entorhinal, o córtex perirhinal e o córtex parahipocampal, (veja figura 1.3).



**Fig. 1.3:** O cérebro está seccionado perpendicularmente ao seu comprimento para mostrar o hipocampo e o córtex de lobo médio temporal (adaptado de [28]).

Em uma série de experimentos em ratos John O'Keefe e colaboradores [13, 7] mostraram que muitos neurônios do hipocampo respondem seletivamente de acordo com localização espacial do rato. Suponha que temos um microeletrodo implantado no hipocampo de um rato enquanto ele anda dentro de uma caixa. O microeletrodo mede a atividade de uma célula piramidal nesta região. A medida em que o rato se move dentro da caixa é possível

observar que a célula somente responde quando o rato está em uma determinada região da caixa. As posições do espaço que geram uma maior atividade na célula neuronal são chamadas de campo receptivo e estes neurônios são chamados de células de posição. As células de posição podem ter seus campos receptivos ajustados as novas situações (figura 1.4). Esta propriedade chamamos de plasticidade.

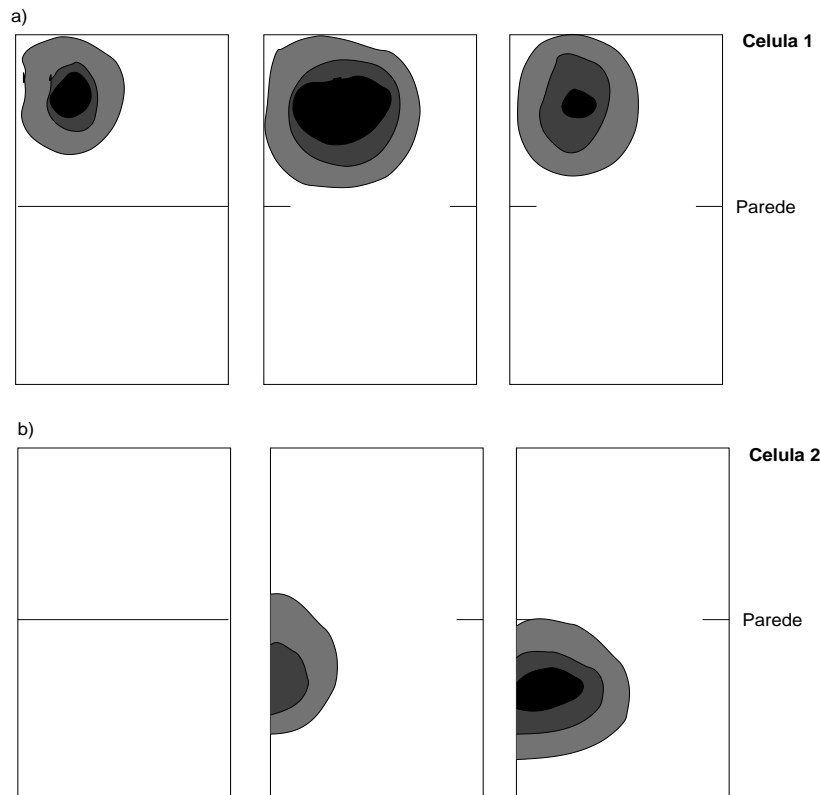
Uma teoria bastante aceita é a que as células de posição do hipocampo fornecem um mapa ambiental que é utilizado para a navegação em um determinado ambiente [7, 24, 25]. Este mapa é chamado de mapa de navegação. Mas apesar dessa teoria ser bastante aceita, existem duas questões fundamentais relacionadas ao processamento espacial que ocorre no hipocampo: como as células de posição surgem? E como as células de posição interagentes podem dar origem a um mapa de navegação? A solução destas questões está intimamente ligada ao entendimento dos mecanismos envolvidos na plasticidade dessas células.

O principal modelo neuro-biológico para as modificações sinápticas está baseado nos paradigmas experimentais eletrofisiológicos conhecidos como LTP (long-term potentiation) e LTD (long-term depression). A LTP é um mecanismo que está relacionado ao aumento da intensidade das sinapses induzidas por curtos estímulos. A LTP usualmente é medida como um aumento da resposta pós-sináptica a estimulação elétrica do neurônio pré-sináptico após breves seqüências de estímulos de alta freqüência [29] (veja figura 1.5b). Uma das propriedades da LTP é a associatividade na qual as entradas de um grupo de células apresentam LTP somente se o estímulo das mesmas estiver simultaneamente associado a um outro estímulo destas entradas, isto está ilustrado na figura 1.5c. A associatividade aparece apenas quando a ativação do neurônio pós-sináptico ultrapassar o limiar do deste neurônio.

A LTP é do tipo homosináptica quando as sinapses que aumentam de intensidade são as mesmas pelas quais o neurônio é ativado (figura 1.5b, entrada A). A LTP apresenta também a característica sinapse-específica, onde aquelas entradas que não são ativadas pelo estímulo não apresentam LTP mesmo se a célula for intensamente estimulada em outras entradas (figura 1.5b, entrada B). Nestas entradas a intensidade sináptica pode ser diminuída ou permanecer como está.

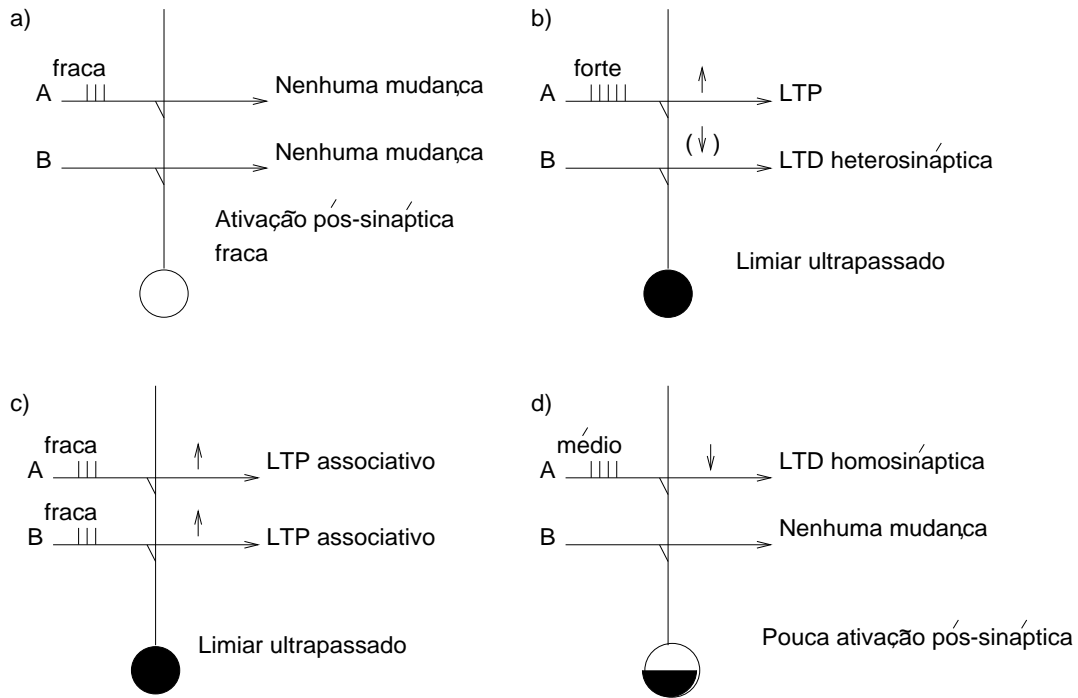
A regra básica da LTP é que as sinapses que conectam dois neurônios tornam-se mais intensas se existir uma correlação temporal da atividade pré-sináptica e da atividade pós-





**Fig. 1.4:** *Células de posição no hipocampo. Um rato explora uma parte da caixa durante 10 minutos (figuras da esquerda), então a parede é removida para que o rato possa explorar toda a caixa (figuras do centro e da direita). a) Escala de cinza indicando a região da caixa onde uma célula de posição no hipocampo responde: preto = resposta intensa, cinza escuro = resposta moderada, cinza claro = resposta fraca e branco = nenhuma resposta. Esta célula tem o campo receptivo na parte superior da caixa, quando a parede é removida, o campo permanece na mesma posição. b) Neste caso, um eletrodo está próximo a uma célula no hipocampo que não responde quando o animal está na parte superior da caixa (esquerda). Após 20 minutos, um campo receptivo desenvolve-se na outra parte da caixa (centro e direita), (adaptado de [28]).*

sináptica. A condição temporal necessária para a produção da LTP é de que o estímulo das entradas do neurônio ocorra próximo e antes do momento de disparo deste neurônio (dentro de aproximadamente 500 ms). Esta regra para a modificação sináptica é conhecida como



**Fig. 1.5:** Esquema de como ocorre a modificação sináptica. Existem duas entradas para o neurônio: A e B. (a) Um fraco estímulo na entrada A (indicado pelas 3 barras) produz pouca ativação pós-sináptica, não ocorrendo nenhuma mudança na intensidade desta sinapse. (b) Um forte estímulo na entrada A (indicado pelas 5 barras) gera uma forte ativação pós-sináptica e causa um aumento da intensidade da sinapses ativas. (c) Duas fracas entradas estimuladas simultaneamente em A e B somam-se para produzir uma forte ativação pós-sináptica, com isso ambas sinapses apresentam LTP. (d) Um estímulo de intensidade intermediária em A gera pouca ativação do neurônio pós-sináptico. As sinapses ativas são enfraquecidas pelo mecanismo de LTD (adaptado de [29]).

regra de Hebb [30].

A LTD (long-term depression), que também pode acontecer, está relacionada a diminuição da intensidade das sinapses. A LTD pode ser associativa ou não associativa. Na LTD associativa, a alteração da intensidade sináptica depende das atividades pré- e pós-sináptica. Existem dois tipos de LTD: a LTD heterosináptica onde a sinapse enfraquecida é diferente daquela pela qual o neurônio pós-sináptico é ativado e acontece somente quando o neurônio

pós-sináptico está fortemente ativo e existe uma baixa atividade do neurônio pré-sináptico (veja figura 1.5b, entrada B) e a LTD homosináptica onde a sinapse enfraquecida é a mesma pela qual o neurônio é ativado e ocorre somente quando o neurônio pré-sináptico está fortemente ativo e o neurônio pós-sináptico com baixa atividade (figura 1.5d).

Além do entendimento do análogo biológico, o projeto de redes neurais artificiais envolve a compreensão e a especificação de alguns ingredientes importantes. Um deles é o mecanismo de modificação sináptica, conhecido como regra de aprendizado. É através desta regra que a rede neural se auto organiza para realizar a tarefa para qual foi concebida. Outro ingrediente, igualmente importante, é a arquitetura da rede. Este item compreende o tipo de neurônio (se binário ou contínuo, se linear ou não linear) e a sua conectividade. Por último, um ingrediente bastante essencial está relacionado a estatística dos padrões de entrada à qual a rede está submetida. Portanto, sempre que for necessário, evidenciaremos quais dos três ingredientes estamos referindo: a regra de aprendizado, a arquitetura da rede ou o espaço dos padrões de entrada.

No capítulo 2 apresentaremos o modelo proposto por Blum e Abbott [10, 31] de como um mapa bidimensional do espaço pode ser criado no hipocampo de ratos através do mecanismo de LTP em uma rede de células de posição. Neste modelo, a seqüência de disparos das células de posição durante a exploração do ambiente induz um processo de LTP que afeta os disparos subseqüentes das células de posição. Eles se baseiam no fato de que a localização do rato é dada pela atividade de um conjunto de células de posição [18] e na necessidade de uma assimetria temporal nas condições para a produção da LTP e LTD sináptica [32, 33, 34, 35, 36]. Com isso, o rato pode se movimentar a partir da posição atual até o local correspondente ao maior nível de atividade das células de posição. Blum e Abbott sugerem que a posição codificada pela atividade do conjunto de células de posição é deslocada na direção e ao longo de uma trajetória específica que o animal está realizando caso o processo de LTP ocorra enquanto o animal estiver sobre esta trajetória. Eles aplicaram estas idéias na simulação computacional do labirinto de Morris, posteriormente Gerstner e Abbott [37] simularam o movimento de um rato fictício em um ambiente bidimensional.

No capítulo 3 discutiremos a relação entre a exploração do ambiente e o surgimento de um mapa de navegação em uma rede de células de posição como a que foi descrita por

Blum e Abbott [10, 31]. Particularmente, mostramos que para uma exploração aleatória a rede captura o gradiente de uma função potencial do tipo harmônica, que é um conhecido método utilizado no planejamento de trajetórias em sistemas artificiais. Para um exploração direcionada, onde as trajetórias anteriormente aprendidas afetam o movimento atual, o mapa deixa de ser o gradiente de um potencial harmônico. Neste caso, dependendo dos parâmetros utilizados, o campo vetorial das direções preferenciais do movimento pode apresentar rotacionais não nulos, fazendo com que apareçam trajetórias fechadas (*loops*) não desejáveis. Apesar disso, mostramos que o aprendizado direcionado quando otimizado pode ser um método mais rápido no aprendizado de trajetórias que o aprendizado aleatório.

No capítulo 4 apresentaremos uma maneira de armazenar e recuperar, em uma rede neural, diferentes mapas de navegação de diferentes ambientes. O modelo é construído baseado na idéia de que um mapa de navegação pode ser escrito como a solução da equação de Laplace com as condições de contorno adequadamente aplicadas aos obstáculos e aos alvos presentes no ambiente. A compressão de informação gerada nos permite obter um bom desempenho com um número reduzido de conexões sinápticas.

No capítulo 5 vamos discutir os resultados obtidos com a simulação em um ambiente bidimensional da superfície potencial ótima para a difusão de partículas em direção a uma posição alvo (absorvedouro). Veremos que esta superfície tem uma estrutura ramificada parecida com a estrutura de um sistema de rios.

No último capítulo apresentamos as conclusões e as perspectivas futuras do nosso trabalho.

## Capítulo 2

# Aprendizado de mapas de navegação

O mapa de navegação de um determinado ambiente é uma estrutura mental que deve fornecer toda a informação necessária para ir de qualquer ponto no ambiente até uma posição alvo específica. Se existirem obstáculos entre alguns pontos iniciais e o objetivo, o mapa deve incluir estratégias para contorná-los. Um mapa de navegação não é um mapa que fornece a representação da estrutura topográfica de um ambiente, mas é um mapa que descreve a relação entre o ambiente e um alvo específico localizado neste ambiente fornecendo a direção do movimento que deve ser realizado para atingir o alvo a partir de qualquer ponto do ambiente. Esta direção pode guiar o movimento ao longo da linha reta que une a localização atual e o alvo ou ao longo de uma curva mais complexa que contorna um obstáculo.

Utilizando resultados experimentais relacionados a LTP (long-term potentiation) e as células de posições do hipocampo [13, 32, 38, 33, 39, 18, 35, 40] Blum e Abbott [10, 31] propuseram um modelo matemático de como um mapa de navegação do ambiente pode ser formado no hipocampo dos ratos. Este modelo está relacionado diretamente como o nosso trabalho, portanto é essencial que se faça uma análise detalhada do mesmo.

## 2.1 O modelo de Blum e Abbott

O modelo proposto por Blum e Abbott [10, 31] é baseado na hipótese que a atividade das células de posição do hipocampo dos ratos fornecem um código para a localização espacial [18]. Isto significa que as taxas de disparos de um grande número de células de posição podem ser combinadas para indicar um posicionamento espacial que é chamado de posição codificada. Neste modelo, a posição codificada é idêntica a posição real do

animal no ambiente. Entretanto, a LTP das sinápses entre as células de posição que ocorre durante a exploração do ambiente causa uma leve mudança na posição codificada. Após um determinado período de aprendizado, a posição codificada pelo conjunto de atividades das células de posição não mais correspondem a localização precisa do animal. Ao invés disso, se o aprendizado envolve um alvo específico, a atividade das células de posição representam a direção na qual o animal deve realizar o próximo movimento para ir da atual posição em direção ao alvo. A diferença entre a posição codificada pelas atividades das células de posição e a posição real que gera estas atividades fornece um mapa de navegação da direção que o animal deve seguir para atingir o alvo a partir de qualquer posição no ambiente.

Os elementos básicos do modelo são: uma rede de células de posição com as suas atividades definidas por uma função Gaussiana da posição do animal, sinapses ajustáveis entre estas células de localização, uma regra para a indução da LTP e LTD e uma forma de identificar a posição codificada pelas atividades das células de posição.

### 2.1.1 Campo receptivo de um neurônio de posição

O modelo de Blum e Abbott é composto por uma rede de  $N$  neurônios respondendo a uma localização espacial de um animal em um ambiente bi-dimensional. Os neurônios  $i = 1, 2, \dots, N$  da rede têm um campo receptivo localizado. A atividade de um dado neurônio  $i$  é dada por

$$s_i = f \left[ \sum_{j \neq i} w_{ij} s_j + A_i G(|\mathbf{x} - \mathbf{x}_i|) \right] \quad (2.1)$$

Na ausência de conexões ( $w_{ij} = 0$ ) os neurônios da rede têm um campo receptivo localizado. Antes do aprendizado pode-se escrever<sup>1</sup>

$$s_i = f [A_i G(|\mathbf{x} - \mathbf{x}_i|)] \quad (2.2)$$

---

<sup>1</sup> Observe que normalmente representamos uma dinâmica neural como atividade  $s_i = f(\text{atividade})$ . Aqui quando escrevemos a atividade  $s_i = f(\text{atividade}, \text{posição})$  existe a noção subjacente que a informação sobre a posição é comunicada ao neurônio  $i$  de forma não ambígua pela atividade de um outro conjunto de neurônios que não nos interessa modelar.

onde  $\mathbf{x}$  é um vetor bi-dimensional que representa a localização do animal,  $\mathbf{x}_i$  é o centro do campo receptivo do neurônio  $i$ ,  $A_i$  é o fator de amplitude e a modulação com a posição é representada por uma função Gaussiana

$$G(|\mathbf{x} - \mathbf{x}_i|) = \exp\left(-\frac{|\mathbf{x} - \mathbf{x}_i|^2}{2\sigma_f^2}\right) \quad (2.3)$$

onde  $\sigma_f$  é a largura do campo receptivo.

### 2.1.2 Acoplamento sináptico e aprendizado

Todos os neurônios na rede estão acoplados através de sinapses que são ajustadas durante o período de aprendizagem. Antes do aprendizado a eficácia  $w_{ij}$  das conexões dos neurônios pré-sinápticos  $j$  para os neurônios pós-sinápticos  $i$  é zero. Durante o aprendizado a eficiência sináptica varia de acordo com a regra de Hebb [30]. Isto significa que a mudança na intensidade de uma dada sinapse é proporcional a correlação entre a atividade dos neurônios pré- e pós-sináptico. Além disso, o aumento na eficiência sináptica depende do tempo relativo entre as atividades pré- e pós-sináptica. É necessário que o disparo do neurônio pré-sináptico ocorra simultaneamente ou antes que o disparo do neurônio pós-sináptico, caso contrário, a intensidade da sinapse entre esses neurônios é diminuída [32, 33, 35]. Então, se a atividade pré-sináptica ocorrer em um tempo  $t'$  e a atividade pós-sináptica em um tempo  $t$ , a mudança na intensidade da sinapse é proporcional ao fator  $g(t - t')$  dado por

$$g(t) = \begin{cases} \tau^{-1} \exp(-t/\tau^{-1}) & \text{para } t \geq 0, \\ -\zeta\tau^{-1} \exp(-t/\tau^{-1}) & \text{para } t < 0 \end{cases} \quad (2.4)$$

que é chamada de ‘função janela’. Esta função simula a assimetria temporal da LTP e LTD. O parâmetro  $\zeta$  controla a diminuição na intensidade da sinapse (LTD). Após um período  $T$  de aprendizado o peso da sinapse entre o neurônio  $j$  e o neurônio  $i$  é dado por [31]

$$w_{ij} = \int_0^T \int_0^T s_i(t)g(t - t')s_j(t')dt dt'. \quad (2.5)$$

Devido ao acoplamento que surge durante o aprendizado, a atividade  $s_i$  do neurônio  $i$  é influenciada pela atividade dos outros neurônios. Neste modelo, Blum e Abbott consideraram uma função de transferência do tipo  $f(x) = x$  e o conjunto de pesos sinápticos

pequenos, com isso, após o período de aprendizado, a atividade do neurônio  $i$  será dado por

$$s_i = A_i G(|\mathbf{x} - \mathbf{x}_i|) + \sum_j w_{ij} A_j G(|\mathbf{x} - \mathbf{x}_j|). \quad (2.6)$$

O segundo termo gera a alteração dependente do aprendizado que produz o mapa de navegação.

### 2.1.3 Codificação da posição

Para codificação da posição Blum e Abbott supuseram que, devido ao grande número de neurônios, os respectivos campos receptivos destes neurônios estão superpostos. Portanto, quando o animal estiver em uma posição  $\mathbf{x}$  muitos neurônios com o centro do campo receptivo próximo à  $\mathbf{x}$  estarão ativos ao mesmo tempo. Para interpretar o padrão de atividade resultante, eles usaram uma simples esquema de decodificação das atividades neurais onde a posição codificada é a média das atividades [38, 40]

$$\mathbf{p} = \frac{\sum_i \mathbf{x}_i s_i}{\sum_i s_i} \quad (2.7)$$

o vetor  $\mathbf{p}$  é chamado de vetor população. Antes do aprendizado, o vetor população representa a localização do animal. Assumindo que as posições centrais dos campos locais, representadas por  $\mathbf{x}_i$  para a célula  $i$  são uniformemente espaçadas e que elas cobrem densamente o ambiente então:

$$\mathbf{p} = \frac{\sum_i \mathbf{x}_i A_i G(|\mathbf{x} - \mathbf{x}_i|)}{\sum_i A_i G(|\mathbf{x} - \mathbf{x}_i|)} \approx \mathbf{x}. \quad (2.8)$$

Com isso é possível concluir que antes do aprendizado  $\mathbf{p} = \mathbf{x}$ . Entretanto, o vetor população muda com o aprendizado induzido pelas alterações na intensidade das sinapses. Em geral, após o aprendizado o vetor população é

$$\mathbf{p}(\mathbf{x}; \{X\}) = \mathbf{x} + \Delta\mathbf{p}(\mathbf{x}; \{X\}) \quad (2.9)$$

onde  $\{X\}$  representa o conjunto de trajetórias realizadas durante o processo de treinamento.

Após o treinamento as alterações no vetor população fazem com que a posição codificada não represente mais a posição real do animal, porém uma posição próxima a esta. A diferença entre a posição codificada e a posição real age como um mapa cognitivo capaz



de guiar o animal através do ambiente na direção do objetivo. Em outras palavras,  $\Delta\mathbf{p}(\mathbf{x})$  fornece a direção que o animal deve seguir a partir de  $\mathbf{x}$  para chegar até o alvo [10, 37]. Para ordem linear de  $w_{ij}$  (veja apêndice A) a alteração  $\Delta\mathbf{p}(\mathbf{x})$  do vetor população induzida após o treinamento é

$$\Delta\mathbf{p}(\mathbf{x}) = \frac{\sum_{ij}(\mathbf{x}_i - \mathbf{x})w_{ij}s_j(\mathbf{x})}{\sum_i s_i(\mathbf{x})}. \quad (2.10)$$

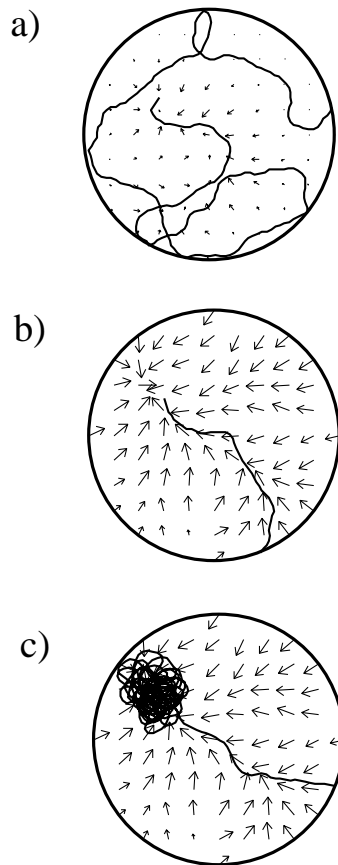
A equação (2.10) significa que o animal anda em direção a posição codificada pela atividade das células de posição. Abbott e Blum [10] aplicaram estas idéias para mostrar como um mapa de navegação surge e como ele poder ser usado para guiar o movimento no labirinto de Morris<sup>2</sup> (veja figura 2.1).

Na figura 2.1 as setas representam  $\Delta\mathbf{p}(\mathbf{x})$ . Esta quantidade é definida para todo o valor de  $\mathbf{x}$  num ambiente bidimensional que está representado de forma discreta na figura. Em todos os pontos da grade, a base da seta representa a posição  $\mathbf{x}$  do ambiente que serve como entrada para a rede neural. A extremidade superior representa o valor correspondente do vetor população decodificado a partir das atividades, geradas pela entrada  $\mathbf{x}$ , das células de posição.

O mapa de navegação que guia o movimento nestas simulações evolui de acordo com as trajetórias seguidas. As setas na figura 2.1 mostram os vários estágios da simulação. Inicialmente todas as setas têm comprimento zero, pois não existe nenhum tipo de aprendizado. Após algumas tentativas, as setas mais próximas do objetivo (plataforma) começam a apontar na direção do mesmo. A medida que novas tentativas vão sendo feitas as setas mais distantes também começam a apontar na direção do objetivo. Na figura 2.1c todas as setas apontam na direção da plataforma indicando que as atividades das células de posição podem fornecer as informações necessárias para encontrar a plataforma a partir de qualquer posição inicial.

---

<sup>2</sup> No experimento de Morris um rato é colocado em uma piscina cheia de água e leite. Uma pequena plataforma está submersa logo abaixo da superfície em uma determinada localização dentro da piscina. O rato colocado dentro da piscina nada até bater na plataforma e subir na mesma. Em geral, os ratos aprendem rapidamente onde está localizada a plataforma. Depois de encontrar a plataforma, o rato é retirado da piscina e colocado em uma caixa que é rotacionada para confundir o senso de direção do animal. Após este procedimento o rato é colocado novamente na borda da piscina em uma nova posição [16].



**Fig. 2.1:** Simulação da trajetória seguindo um mapa de navegação em um labirinto de Morris. As posições iniciais são escolhidas aleatoriamente ao longo do perímetro de 1m do “tanque”. O procedimento é executado até a trajetória interceptar a plataforma de 10cm de diâmetro ou até ultrapassar 100s. As setas mostram o mapa de navegação. (a) A segunda trajetória e o mapa de navegação para uma execução da simulação. (b) A 20<sup>a</sup> trajetória e o mapa de navegação para a correspondente simulação. (c) A 21<sup>a</sup> trajetória e o mapa de navegação para a correspondente simulação quando a plataforma é removida.

## 2.2 Análise do modelo

Nesta seção é feita uma análise de como um mapa útil para a navegação pode surgir a partir do ajuste das sinapses entre as células de posição que ocorre durante a exploração aleatória do ambiente. Considerando o efeito do aprendizado na presença de um único alvo

na posição  $\mathbf{u}$  e analisando o vetor população  $\mathbf{p}(\mathbf{x})$  que surge quando o animal está na posição  $\mathbf{x}$  é possível observar que antes do aprendizado  $\mathbf{p}(\mathbf{x}) = \mathbf{x}$  e após  $\mathbf{p}(\mathbf{x}) = \mathbf{x} + \Delta\mathbf{p}(\mathbf{x})$ . As mudanças  $\Delta\mathbf{p}$  na posição codificada pelas atividades das células de posição que são causadas pelas alterações sinápticas constituem um mapa útil para a navegação. Este mapa inclui as direções que apontam para o alvo, para longe das paredes e que contornam os obstáculos.

As alterações no vetor população podem ser divididas em dois termos com diferentes características e que é denominado de componentes estática e dinâmica. Se o animal está parado em uma posição  $\mathbf{x}_1$ , os neurônios com o centro do campo receptivo próximo a  $\mathbf{x}_1$  serão ativados e as sinapses entre eles serão fortalecidas devido aos seus disparos correlacionados. Se, mais tarde, o animal retornar para a vizinhança do ponto  $\mathbf{x}_1$ , os neurônios com o campo receptivo próximos a  $\mathbf{x}_1$  serão mais fortemente ativados do que aqueles com o campo receptivo mais longe de  $\mathbf{x}_1$  porque eles receberam um estímulo sináptico maior através das sinapses previamente fortalecidas. Como resultado, o vetor população para as posições próximas é mudado em direção à  $\mathbf{x}_1$ . Este aprendizado estático não depende da assimetria temporal da função janela (2.4) e tem um efeito local, pois para distâncias maiores que  $\sigma_f$  este efeito é nulo.

Além disso, existe o termo dinâmico. Se o animal move-se ao longo de uma linha a partir de  $\mathbf{x}_1$  até um ponto  $\mathbf{x}_2$  próximo, as sinapses entre neurônios com campos receptivos próximos a  $\mathbf{x}_1$  e os neurônios com campos próximos a  $\mathbf{x}_2$  serão fortalecidas se eles forem seqüencialmente ativados dentro de um período de tempo coberto pela função janela. Entretanto, sinapses entre os neurônios com campos receptivos próximos a  $\mathbf{x}_2$  e neurônios com campos receptivos próximos a  $\mathbf{x}_1$  serão enfraquecidas porque seus disparos são em ordem contrária. Em geral, desde que a LTP exija que o neurônio pré-sináptico esteja ativo antes ou durante a ativação do neurônio pós-sináptico, apenas as sinapses na direção do movimento e ao longo da trajetória é que serão fortalecidas. O efeito do aprendizado dinâmico dá origem a uma componente de longo alcance do mapa de navegação.

### 2.2.1 O alcance do aprendizado

Para o caso de um único alvo em um ambiente simples é possível fazer uma análise dos efeitos da LTP e LTD sobre o vetor população durante a exploração aleatória deste ambiente. A

alteração do vetor população dada por  $\Delta p_1(\mathbf{x}; \{X\})$  é gerada quando o animal está em uma posição de referência  $\mathbf{x}$  e após um período de aprendizado, que envolve o movimento ao longo de uma trajetória  $X(t)$ , atinge o alvo fixado no ponto  $\mathbf{u}$ . Considerando  $X(t)$  como sendo uma linha ao longo da qual o animal se move com uma velocidade constante  $\mathbf{v}$  parametrizada da seguinte forma

$$X(t') = \mathbf{x} + D(\mathbf{x})\mathbf{v}_\perp + (t' - t_0)\mathbf{v} \quad (2.11)$$

onde  $\mathbf{v}_\perp$  é o vetor unitário perpendicular a direção de movimento do animal,  $|D(\mathbf{x})|$  é a distância perpendicular entre a posição de referência  $\mathbf{x}$  e a trajetória e  $D(\mathbf{x})$  podendo ser negativo ou positivo dependendo da direção do movimento ao longo da trajetória, Blum e Abbott [31] mostraram que o aprendizado de uma trajetória  $X(t')$  com os parâmetros  $D(x)$  e  $\mathbf{v}$  causa uma alteração do vetor população dada por

$$\Delta \mathbf{p}_1(\mathbf{x}; D, \mathbf{v}) = \gamma \sigma_f \left[ g_0 \frac{\mathbf{v}_\perp}{|\mathbf{v}|} D(\mathbf{x}) + g_1 \frac{\mathbf{v}}{|\mathbf{v}|} \right] \exp \left( -\frac{D^2(\mathbf{x})}{4\sigma_f^2} \right) \quad (2.12)$$

onde  $\gamma$  é uma constante a  $g_0 = \int dt g(t)$  e  $g_1 = \int dt g(t)t$  são o momento médio e o primeiro momento da função janela. O primeiro termo (entre colchetes) da equação (2.12) produz uma mudança no vetor população que aponta em direção à trajetória, independente da direção do movimento ao longo da trajetória. Este é o termo de aprendizado estático que depende de  $\sigma_f/|\mathbf{v}|$  que é, de um grosso modo, o tempo necessário para atravessar qualquer campo receptivo ao longo do caminho. O segundo termo em (2.12) é o termo de aprendizado dinâmico que depende da direção do movimento do animal e gera uma componente do vetor população que aponta na direção ao longo da trajetória. Nota-se que a realização do mesmo caminho com a mesma velocidade em sentido contrário muda o sinal do termo do aprendizado dinâmico, mas não muda a contribuição do termo estático. A magnitude do termo dinâmico depende do tamanho da assimetria temporal dado pela função janela (2.4) através do momento  $g_1$ , enquanto que o termo estático não depende desta assimetria mas apenas da magnitude da função janela  $g_0$ .

Para analisar o aprendizado envolvendo múltiplas trajetórias Abbott e Blum consideraram as trajetórias retas com velocidade unitária  $|\mathbf{v}| = 1$ . Usando a direção  $\theta$  em vez do vetor velocidade  $\mathbf{v}$  como parâmetro, onde  $\theta$  é o ângulo entre a direção do vetor  $\mathbf{v}$  e a linha

que conecta a posição de referência  $\mathbf{x}$  e a posição do alvo  $\mathbf{u}$  e  $D$  como sendo a distância entre o ponto de referência  $\mathbf{x}$  e a trajetória, eles mostraram que após  $K$  trajetórias o efeito no aprendizado é

$$\Delta \mathbf{p}(\mathbf{x}; \mathbf{u}) \propto \sum_{k=1}^K \Delta \mathbf{p}_1(\mathbf{x}; D_k, \theta_k). \quad (2.13)$$

Supondo que o tempo de aprendizado é longo o suficiente tal que, em média, o aprendizado ocorre em um grande número de trajetórias  $K \gg 1$  a alteração média no vetor população pode ser aproximado para

$$\begin{aligned} \Delta \mathbf{p}(\mathbf{x}; \mathbf{u}) &= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \Delta \mathbf{p}_1(\mathbf{x}; D_k, \theta_k) \\ &= \int dD P(D) \int d\theta P_D(\theta) \Delta \mathbf{p}_1(\mathbf{x}; D, \theta) \end{aligned} \quad (2.14)$$

onde  $P(D)$  é a probabilidade de encontrar uma trajetória á uma distância  $D$  do ponto de referência  $\mathbf{x}$  e  $P_D(\theta)$  é a probabilidade condicional que uma trajetória á uma distância  $D$  tenha a direção  $\theta$ . Por definição a direção  $\theta$  depende da localização do alvo  $\mathbf{u}$ . Para uma trajetória qualquer, não necessariamente reta

$$\Delta \mathbf{p}(\mathbf{x}; \mathbf{u}) = \int d\{X\} \Delta \mathbf{p}_1(\mathbf{x}; \{X\}) \quad (2.15)$$

onde  $\mathbf{p}_1$  é a mudança devido a uma única trajetória e a integração é sobre todo o conjunto das possíveis trajetórias. A dependência da posição do alvo aparece devido ao seu efeito sobre o conjunto de trajetórias possíveis  $\{X\}$  durante o aprendizado.

Embora as trajetórias sejam geradas por uma estratégia de procura aleatória, aquelas que saem do alvo não são permitidas. De acordo com o procedimento de aprendizado adotado, quando uma trajetória atinge o alvo ela é reiniciada em uma nova posição aleatória no ambiente. Portanto, não existem trajetórias iniciando no alvo, com isso a distribuição  $P_D(\theta)$  é não-uniforme. No caso de trajetórias retas em um alvo circular de pequeno diâmetro, o conjunto de trajetórias possíveis pode ser construída geometricamente. Usando argumentos de que em um ambiente homogêneo e isotrópico com suficiente número de trajetórias, a mudança  $\Delta \mathbf{p}$  deve apontar sempre na direção do alvo. Com isso

$$\Delta \mathbf{p}(\mathbf{x}; \mathbf{u}) = \frac{\mathbf{u} - \mathbf{x}}{|\mathbf{u} - \mathbf{x}|} F(|\mathbf{u} - \mathbf{x}|) \quad (2.16)$$

onde a função  $F$  depende da distância entre a localização  $\mathbf{x}$  e a posição do alvo  $\mathbf{u}$ .

Supondo um ambiente aberto onde todas as trajetórias começam no infinito e terminam no alvo então a função  $F(d)$  cai com  $d^{-1}$  para distâncias maiores que o tamanho típico dos campos receptivos. Portanto, o aprendizado tem o efeito de longo alcance que insere uma informação local sobre a posição do alvo em um mapa de navegação global. Para ver como isto acontece analisamos o termo de aprendizado dinâmico da equação (2.12), pois o termo estático não é sensível à direção do movimento e por isso o seu efeito é nulo devido a simetria do conjunto de trajetórias. Na soma de todas as possíveis trajetórias, aquelas que não chegam até o alvo se anulam, pois existem duas direções opostas do movimento para cada uma destas trajetórias e o efeito dinâmico do aprendizado para estas duas direções cancelam-se mutuamente. Para derivar a dependência de  $\frac{1}{d}$  é necessário observar que em um ambiente bidimensional a densidade de trajetórias que chegam até o alvo cai com o inverso da distância em relação ao alvo, isto porque o número de trajetórias que chegam menos o número de trajetórias que saem através de uma região circular em torno do alvo é o mesmo independente do tamanho desta região, pois todas as trajetórias começam no infinito e terminam somente no alvo. Em um ambiente finito com os contornos fixos e alguns obstáculos, o campo vetorial do mapa de navegação gerado não será mais radial, mas ainda será de longo alcance. Além disso, todas as trajetórias geradas pelo campo vetorial terminarão no alvo. O problema que surge em ambientes finitos é que podem existir trajetórias fechadas (*loops*) durante o aprendizado mesmo que estas trajetórias sejam compostas de segmentos de retas. Entretanto, desde que existam *loops* no sentido horário e anti-horário com a mesma probabilidade, no limite de um grande número de trajetórias o campo vetorial resultante terá o rotacional nulo, com isso pode-se determinar uma superfície potencial com um mínimo localizado no alvo. Se, após o aprendizado, o animal mover-se ao longo da trajetória gerada pelo campo vetorial, este movimento corresponderá à descida desta superfície potencial e com isso o animal deverá alcançar o alvo.

### 2.2.2 Simulações

Usando as idéias propostas por Blum e Abbott, que foram descritas anteriormente, Gerstner e Abbott [37] realizaram simulações de um rato fictício movendo-se ao longo de linhas retas

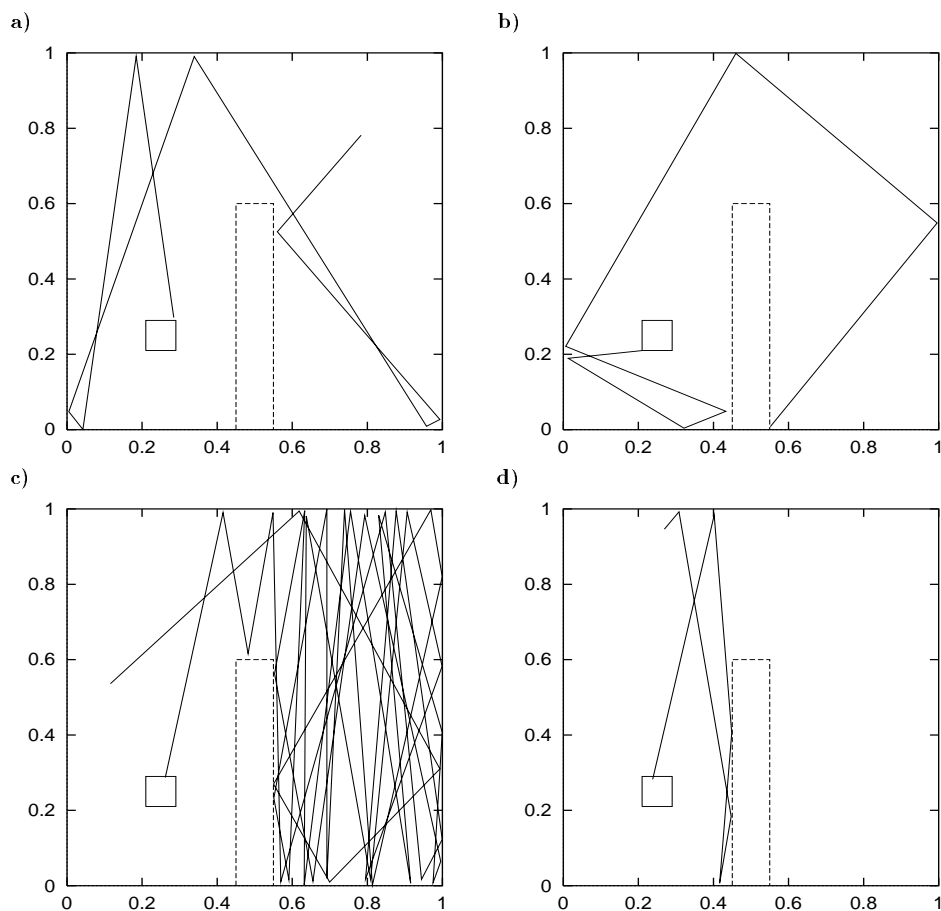
em um ambiente bi-dimensional. O ambiente consistia de uma caixa de tamanho  $L \times L$  onde o comprimento  $L$  era muito maior que a largura do campo receptivo ( $L \gg \sigma_f$ ). O movimento do rato iniciava em uma posição qualquer do ambiente e seguia em uma direção aleatória. Se uma das paredes fosse atingida, o rato mudava aleatoriamente a direção do movimento para o interior da caixa. O movimento parava se o alvo fosse atingido, porém o aprendizado continuava durante 100 intervalos de tempo enquanto o rato estivesse situado sobre o alvo. Finalmente a próxima tentativa começava em uma nova posição aleatória. Um conjunto de trajetórias gerados desta forma é mostrado na figura 2.2.

A idéia por trás deste procedimento é que o animal aprende apenas enquanto está procurando o alvo. Depois de ter encontrado o alvo, o animal permanecia nesta posição. Por exemplo, se o alvo fosse uma fonte de comida, o animal iria se alimentar após ter encontrado o alvo. Depois disso, ele andaria pelo ambiente sem a necessidade de aprender. O processo de aprendizagem recomeçava quando o animal sentisse fome e saísse a procura da fonte de comida novamente.

Nas simulações foi usado um ambiente quadrado de lado  $L = 1$ . Os campos receptivos possuem  $\sigma_f = 0.1$  e seus centros estão uniformemente distribuídos em uma grade de espaçamento 0.1. Nestas simulações para um único alvo usou-se 121 ( $11 \times 11$ ) neurônios.

O ‘animal’ movia-se através do ambiente ao longo de linha retas a uma velocidade constante  $|\mathbf{v}| = 0.05$  unidades de distância por unidade de tempo. O movimento era aleatoriamente espalhado nas bordas e suspenso cada vez que o alvo fosse atingido. O alvo era um quadrado de lado 0.1 e a função janela (2.4) tinha a constante  $\tau = 10$ . A contribuição de  $g_0$  e  $g_1$  podia ser ajustada pela variação da assimetria da função (2.4) através do parâmetro  $\zeta$ .

A componente estática do aprendizado gera uma mudança do vetor população na direção das posições que são visitadas mais frequentemente. Isto pode ser observado claramente em simulações na ausência de alvo. Na figura 2.3a foi atribuído  $\zeta = 0$  para realçar a componente estática. Após a exploração aleatória do ambiente na ausência do alvo, as características dominantes no mapa de navegação são as paredes do ambiente. As setas na figura 2.3a, representando a mudança na posição codificada devido a exploração e ao ajuste das sinapses, apontam para longe das paredes. Para o caso de  $\zeta = 1$  então  $g_0 = 0$ , com isso a componente estática é eliminada. O resultado pode ser visto na figura 2.3b.

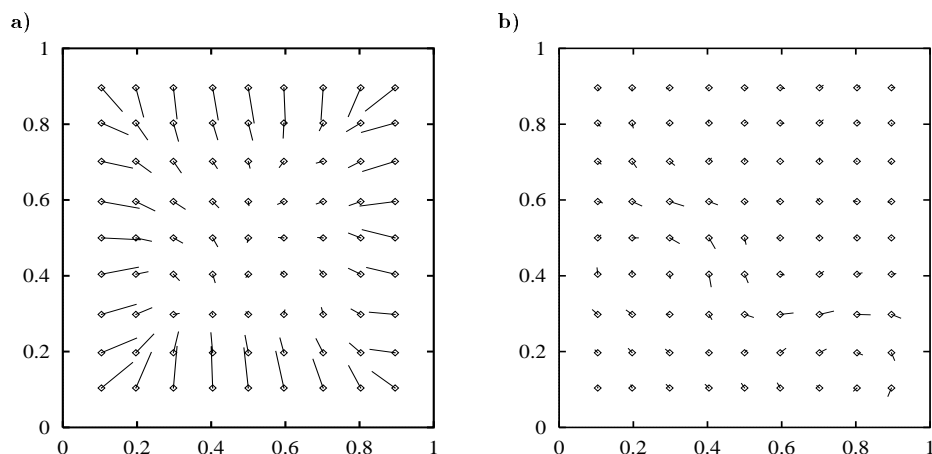


**Fig. 2.2:** Quatro exemplos de caminhos exploratórios. O ambiente consiste de uma caixa quadrada com uma barreira (pontilhado) no centro. Cada exploração começa em uma posição aleatória, atravessa o interior em linhas retas, é espalhado nas bordas e termina no alvo indicado pelo pequeno quadrado na posição  $(0.25, 0.25)$  (adaptado de [37]).

Agora analisando o mapa que surge quanto um alvo e um obstáculo são adicionados no ambiente. O ambiente e algumas trajetórias são mostradas na figura 2.2. A componente estática faz com que o animal evite os obstáculos e as paredes, além disso, esta componente é útil quando o animal encontra e pára no alvo. Por estas razões, foi mantida uma pequena contribuição da componente estática do aprendizado usando  $\zeta = 0.8$  em todas as simulações de agora em diante<sup>3</sup>.

<sup>3</sup> O valor de  $\zeta$  é arbitrário, qualquer valor entre 0.7 e 0.8 pode ser considerado um bom valor para  $\zeta$ .



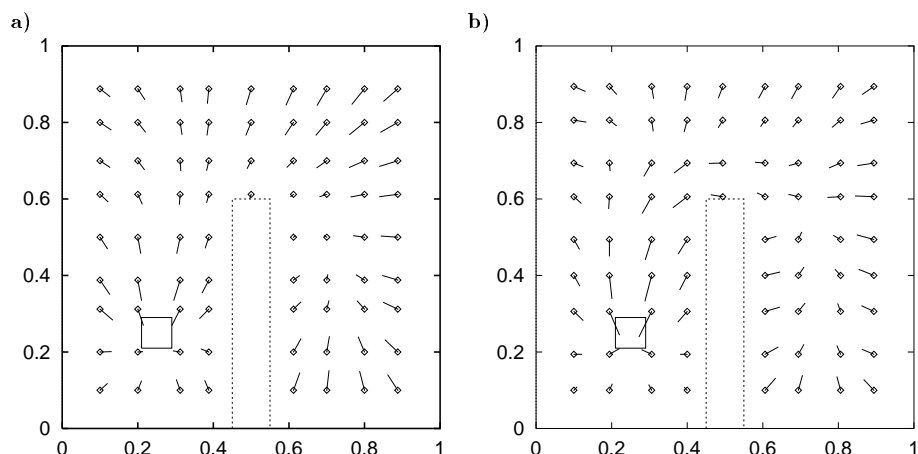


**Fig. 2.3:** Componentes a) estática e b) dinâmica do aprendizado (adaptado de [37]).

Conforme mencionado anteriormente, o aprendizado dinâmico contribui com uma componente de longo alcance para o mapa de navegação. Esta componente pode ser vista claramente na figura 2.3. Após um período de ajustes sinápticos e de exploração aleatória, o mapa de navegação resultante fornece um guia útil até o alvo a partir de qualquer ponto no ambiente. As setas do lado direito do ambiente não apontam diretamente para o alvo, mas para uma direção que contorna o obstáculo guiando o movimento em direção ao alvo. Na figura 2.4a, algumas setas à direita da barreira são pequenas. Para melhorar a estrutura do mapa a resolução foi aumentada para uma rede de  $21 \times 21$  células e os campos receptivos com largura  $\sigma_f = 0.05$ . O mapa obtido após 400 tentativas é mais regular, como pode ser observado na figura 2.4b.

## 2.3 Conclusões

Neste capítulo vimos que a formação de um mapa pelas células de posição do hipocampo está baseada em alguns princípios básicos: a codificação da posição espacial por um conjunto de células, a assimetria temporal das atualizações sinápticas, a exploração aleatória do ambiente. Através deste princípios básicos vimos que um rede de células de posição pode desenvolver mapas que indicam a direção que deve ser seguida para se alcançar o objetivo (alvo) de forma que os obstáculos existentes no ambiente sejam contornados e as paredes



**Fig. 2.4:** *Aprendizado da posição de um alvo. As setas mostram o deslocamento do vetor população após um período de aprendizado de a) 100 e b) 400 tentativas exploratórias. O mapa de navegação formado pode guiar o animal em direção ao alvo (pequeno quadrado) a partir de qualquer posição inicial. As figuras a) e b) diferem também pelas densidades dos campos receptivos. a) Rede de  $11 \times 11$  células e os campos receptivos com largura  $\sigma_f = 0.1$ . b) Rede de  $21 \times 21$  células e os campos receptivos com largura  $\sigma_f = 0.05$  (adaptado de [37]).*

sejam evitadas. Estes mapas são extremamente fáceis de serem seguidos, pois o animal deve simplesmente seguir em direção à posição codificada pelas atividades das células de posição. Conforme o animal se move, as posições codificadas conduzirão o animal ao longo dos caminhos armazenados no mapa.

No capítulo seguinte veremos que um mapa de navegação pode ser dado pelo gradiente de uma função potencial e que uma rede de células de posição é capaz de capturar este gradiente que surge quando da exploração aleatória de um ambiente.

## Capítulo 3

# Aprendizado não supervisionado de trajetórias

Neste capítulo vamos discutir a relação entre a exploração de um ambiente e o surgimento de um mapa de navegação em uma rede de células de posição. Em particular, veremos que para uma exploração aleatória a rede captura o gradiente de uma função potencial, que é um conhecido método no planejamento de trajetórias em robótica. Para uma exploração direcionada, onde as trajetórias exploradas e aprendidas anteriormente afetam o movimento, o mapa não representa mais o gradiente de uma função potencial <sup>1</sup>.

Assim como no modelo de Blum e Abbott, descrito no capítulo anterior, não vamos discutir como que as células nervosas tornam-se células de posição. Nosso objetivo é entender o problema de aprendizado em uma rede de células de posição, semelhante a de Blum e Abbott, já existentes.

O aprendizado em células de posição é um problema de aprendizado não supervisionado. Portanto, inicialmente analisamos os aspectos estatísticos do movimento bidimensional em um ambiente fechado. A idéia é que o movimento associado à percepção do ambiente produza os dados de entrada para a rede. Se estes dados tiverem uma distribuição de probabilidade não trivial, um simples esquema de aprendizado que extrai correlações, como o aprendizado de Hebb, é capaz de capturar a essência destes dados. Portanto, uma correta estatística é crucial para que haja aprendizado.

---

<sup>1</sup> Em IDIART, M. A. P.; TREVISAN, M. Unsupervised path learning in place-cells networks. *Submetido a Network: Computation in Neural Systems*

### 3.1 Mapa de navegação

Definimos um mapa de navegação como uma função vetorial  $\mathbf{v} = \mathbf{v}(\mathbf{r})$  que indica a direção preferencial a ser seguida quando se está na posição  $\mathbf{r}$ . O mapa guarda a informação dinâmica necessária para a navegação em termos de uma posição absoluta obtida de um mapa estático já existente, onde somente a posição é armazenada.

O mapa de navegação pode surgir de um movimento não intencional e aleatório onde “não intencional” significa que não há instrução explícita usada para localizar uma determinada característica do ambiente. Portanto, a navegação é capacidade que surge naturalmente e não uma capacidade pré-programada.

Suponha um “animal” seguindo uma dinâmica estocástica em um ambiente. Chamaremos isto de exploração aleatória e a única exigência é que este processo seja Markoviano. Cada passo depende somente do estado anterior dado pela posição  $\mathbf{r}(t)$  e orientação  $\phi(t)$  em relação as coordenadas absolutas do ambiente, ou seja:

$$\begin{aligned}\mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \Delta \mathbf{r}(t) \\ \Delta \mathbf{r}(t) &= (\cos \phi(t), \sin \phi(t)) \\ \phi(t) &= \phi(t - \Delta t) + \Delta \phi\end{aligned}\tag{3.1}$$

onde  $\Delta \phi$  é uma variável aleatória com a probabilidade definida por

$$P(\Delta \phi) = \frac{1}{\theta} \Theta\left(\Delta \phi + \frac{\theta}{2}\right) \Theta\left(\Delta \phi - \frac{\theta}{2}\right)\tag{3.2}$$

Nas discussões e resultados apresentadas aqui o animal se move com passos de tamanho constante  $\alpha \ll L$ , onde  $L$  é a dimensão do ambiente. Isto representa uma dinâmica difusiva e  $\frac{1}{\theta}$  fornece uma medida para o livre caminho médio ou uma medida de quão reto é o movimento.

Para entender a estatística precisamos definir quantidades que descrevam como o movimento acontece no ambiente. A primeira destas quantidades é a direção preferencial seguida pelo animal quando este atravessa o ambiente. Para fazer isto, dividimos o ambiente numa rede quadrada. O campo vetorial que fornece a direção preferencial de movimento em uma

célula  $C_{ab}$  é definido como

$$\mathbf{v}_{ab}(t) = \frac{1}{t} \sum_{t'=1}^t \Delta \mathbf{r}(t') \delta_{\mathbf{r}(t') \in C_{ab}} . \quad (3.3)$$

Do mesmo modo, a distribuição de probabilidade de encontrar o animal em uma célula  $C_{ab}$  é dado por

$$\rho_{ab}(t) = \frac{1}{t} \sum_{t'=1}^t \delta_{\mathbf{r}(t') \in C_{ab}} . \quad (3.4)$$

As funções  $\mathbf{v}_{ab}(t)$  e  $\rho_{ab}(t)$  estão relacionadas com as componentes estática e dinâmica do aprendizado de trajetórias discutidas no capítulo anterior. Se tomarmos os limites  $|C_{ab}| \rightarrow 0$  e  $\Delta t \rightarrow 0$  é possível mostrar que ambas as funções são contínuas no plano e  $\mathbf{v}(\mathbf{r}, t)$  aponta na direção do gradiente descendente de  $\rho(\mathbf{r}, t)$

$$\mathbf{v}(\mathbf{r}, t) \propto -\nabla \rho(\mathbf{r}, t) . \quad (3.5)$$

Imagine o animal movendo-se aleatoriamente em um ambiente fechado. Para tempos longos, o movimento aleatório resulta em  $\rho(\mathbf{r}, \infty) = \rho(\mathbf{r})$  que satisfaz a equação de Laplace

$$\nabla^2 \rho(\mathbf{r}) = 0 \quad (3.6)$$

com condições de contorno reflexivas, do tipo de Neumann

$$\mathbf{n} \cdot \nabla \rho(\mathbf{r}) = 0 \quad \text{para } \mathbf{r} \in \partial \Gamma_{\text{obstáculo}} \quad (3.7)$$

onde  $\partial \Gamma_{\text{obstáculo}}$  é um segmento contínuo junto a borda do ambiente  $\Gamma$ . O problema de contorno acima tem solução trivial  $\rho(\mathbf{r}) = \text{constante}$  e conseqüentemente não existe direção preferencial no ambiente, pois o gradiente  $\nabla \rho(\mathbf{r})$  se anula em todos os pontos.

Para escapar desta situação uma característica não trivial deve ser adicionada ao ambiente. Esta característica que torna o aprendizado possível chamamos de “alvo”.

### 3.1.1 O alvo

O alvo é uma operação pertencente ao ambiente e que age sobre o animal. Esta operação age da seguinte maneira: quando o animal alcança o alvo, ele é reposicionado aleatoriamente

no ambiente. Definimos o alvo desta maneira porque é exatamente isto que ocorre com o animal quando submetido ao experimento de Morris. A presença do alvo quebra a simetria da direção do campo vetorial, pois a região em torno do alvo torna-se um absorvedouro de trajetórias. O problema de contorno dado pela equação (3.6) torna-se agora uma equação de Poisson

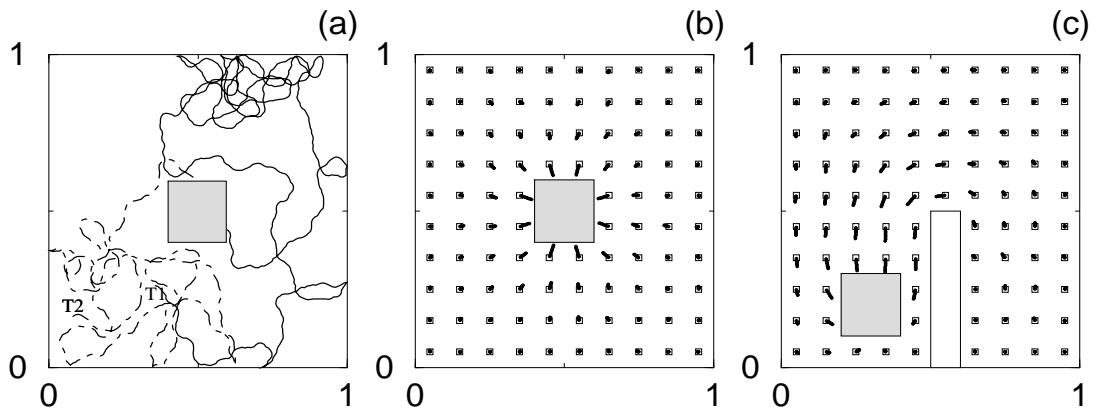
$$\nabla^2 \rho(\mathbf{r}) + \rho_0(\mathbf{r}) = 0 \quad (3.8)$$

com condições de contorno

$$\mathbf{n} \cdot \nabla \rho(\mathbf{r}) = 0 \quad \text{para } \mathbf{r} \in \partial\Gamma_{\text{obstáculo}} \quad (\text{reflexão})$$

$$\rho(\mathbf{r}) = 0 \quad \text{para } \mathbf{r} \in \partial\Gamma_{\text{alvo}} \quad (\text{absorção})$$

e  $\rho_0(\mathbf{r})d\mathbf{r}$  é a probabilidade que o animal apareça na posição  $\mathbf{r}$  após alcançar o alvo. Isto representa uma fonte de partículas em um problema de difusão. A figura 3.1 mostra o efeito do alvo sobre o campo vetorial  $\mathbf{v}(\mathbf{r})$  em um ambiente simples. Observe que esta figura é muito semelhante aquelas produzidas nas simulações do modelo de Blum e Abbott. No entanto, aqui não existe um modelo de rede neural subjacente. Isto mostra que os mapas surgem simplesmente da estatística de passagens.



**Fig. 3.1:** (a) Exemplos de trajetórias exploratórias com passos de tamanho  $\alpha = 0.01$  e  $\theta = 0.2$  em um ambiente com um alvo (quadrado cinza). (b) Direção preferencial do campo vetorial após 8000 tentativas de aprendizado. (c) Campo vetorial após 8000 tentativas em um ambiente com obstáculo, para  $\alpha = 0.01$  e  $\theta = 0.2$ .

Também é um resultado direto do problema de Poisson o fato que o alvo afeta a direção preferencial do campo vetorial em todo o ambiente, com o valor absoluto do campo  $|\mathbf{v}(\mathbf{r})| =$

$v(\mathbf{r})$  caindo com  $\frac{1}{d}$ , onde  $d$  é a distância entre o alvo e a posição  $\mathbf{r}$ . Assim, esta dependência de longo alcance não é devido ao processo de aprendizado especificamente, mas é uma propriedade da geometria e da estatística do ambiente na presença de um alvo.

Vários alvos em diferentes posições ou mudanças na geometria do ambiente alteram consideravelmente o que é mostrado na figura 3.1b. Longos corredores com um alvo nas extremidades terão um campo vetorial apontando para o alvo ao longo de toda a sua extensão. O efeito de um obstáculo pode ser visto na figura 3.1c.

O que torna o gradiente descendente de  $\rho(\mathbf{r})$  um bom mapa de navegação é o fato que os alvos, regiões onde a condição de contorno aplicada é  $\rho(\mathbf{r}) = 0$ , são os únicos mínimos existentes. Desta forma, a descida do gradiente de  $\rho(\mathbf{r})$  sempre levará ao alvo sem o risco de ficar preso em mínimos locais. Esta abordagem está ligada ao uso de funções harmônicas no planejamento de trajetórias em robótica<sup>2</sup>[41, 42, 43]. As funções harmônicas são soluções da equação de Laplace (3.6). Nestes algoritmos o mapa de navegação surge da solução da equação de Laplace com condições de contorno muito específicas. As condições de contorno são: nas paredes e nas obstáculos o potencial tem valor  $\rho(\mathbf{r} \in \partial\Gamma_{obstáculo}) = 1$  e nos alvos o potencial tem valor  $\rho(\mathbf{r} \in \partial\Gamma_{alvo}) = 0$ . Sendo o potencial  $\rho(\mathbf{r})$  uma função harmônica então não teremos mínimos locais além do alvo, pois o princípio do máximo garante isso (veja apêndice B). Existem várias maneiras de calcular este potencial incluindo uma técnica de Monte Carlo através de passeios aleatórios [44, 45].

Embora as duas abordagens parecem ser equivalentes no que diz respeito as propriedades do mapa gerado elas possuem algumas diferenças importantes. No caso dos algoritmos que utilizam funções harmônicas para navegação em robótica o potencial é uma entidade abstrata que é calculado após o robô ter marcado o ambiente com zeros nas posições dos alvos e uns nas posições dos obstáculos. Isto pode ser calculado usando-se passeios aleatórios, mas existem métodos melhores e mais rápidos para fazê-lo, como por exemplo técnicas numéricas do tipo Jacobi, Gauss-Siedel ou SOR [41]. No nosso caso os caminhos são trajetórias reais do animal no ambiente. O potencial  $\rho(\mathbf{r})$  representa a ocupação do ambiente durante a exploração, enquanto que  $\mathbf{v}(\mathbf{r}) = -\nabla\rho(\mathbf{r})$  é a direção preferencial a ser seguida quando se passa próximo da posição  $\mathbf{r}$ . Estas duas quantidades são mensuráveis e

<sup>2</sup> Mais adiante analisaremos este método em maiores detalhes.

podem ser armazenadas nas conexões sinápticas de uma rede neural, como é demonstrado no modelo de Blum e Abbott. Outra diferença essencial é que nos algoritmos que utilizam funções harmônicas pressupõem-se que os alvos são conhecidos com antecedência. Enquanto que neste caso nenhum conhecimento prévio sobre os alvos é necessário.

## 3.2 Contexto e múltiplos mapas

A trajetória do animal é sempre contínua. Então qual seria o significado das trajetórias interrompidas pela ação do alvo? Por exemplo, se tomarmos duas trajetórias contínuas  $p_1(\mathbf{r}_1 \rightarrow \mathbf{R}_1^t)$  e  $p_2(\mathbf{r}_2 \rightarrow \mathbf{R}_2^t)$  onde  $\mathbf{R}^t$  são pontos dentro do alvo e próximos a sua borda. Assumindo que elas são trajetórias seqüenciais para o animal, é válido perguntar o acontece com o segmento que conecta  $\mathbf{R}_1^t$  ao  $\mathbf{r}_2$ ? É importante notar que tal descontinuidade nas trajetórias é essencial para gerar o campo que aponta na direção do alvo. Uma possível explicação seria que o animal é removido do ambiente toda vez que ele encontra o alvo, como no experimento do labirinto de Morris [16]. Mas isto não funcionaria em condições mais realísticas como um animal solto na natureza ou um robô em um ambiente industrial.

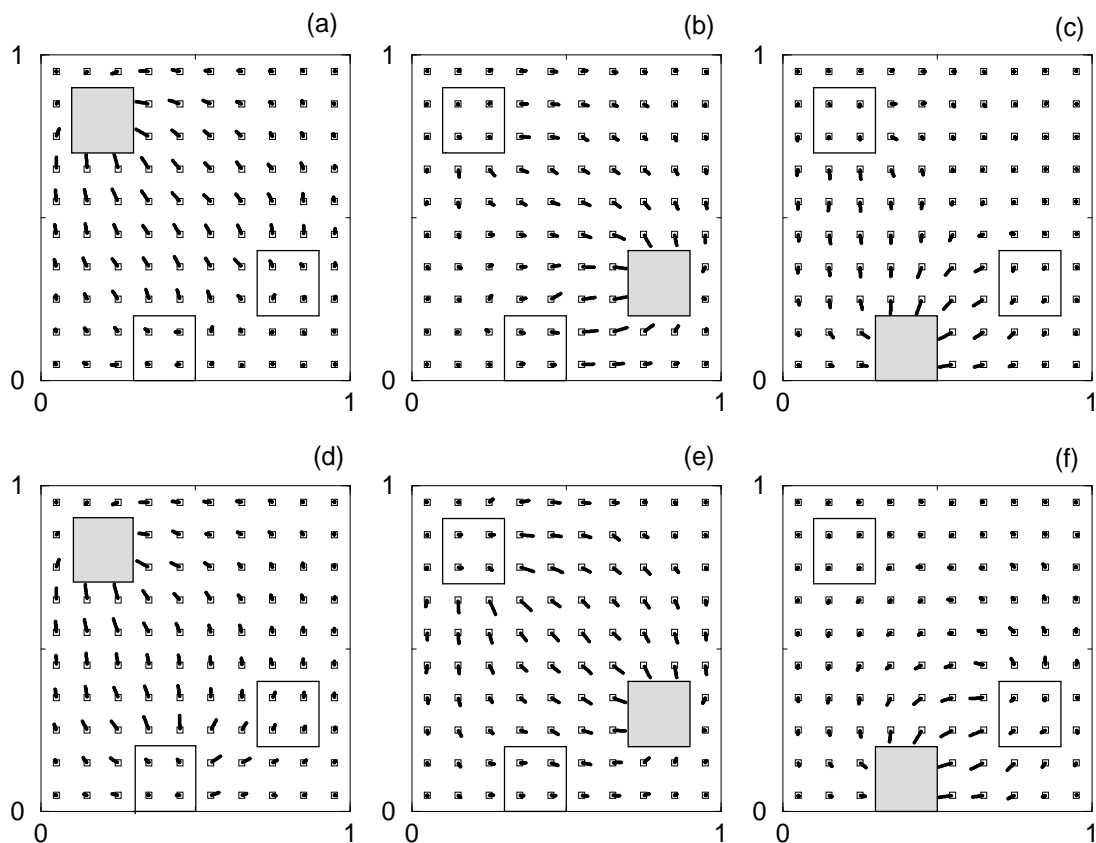
A solução para isto é uma reinterpretação da ação do alvo. Ao invés de considerar que o alvo age sobre o estado físico externo do agente (reposicionando-o), consideramos que o alvo age sobre um estado interno que serve como um acesso para os vários mapas do sistema [16, 24, 46] onde o campo vetorial  $\mathbf{v}(\mathbf{r})$  ou  $\rho(\mathbf{r})$  são armazenados. A troca entre diferentes mapas causa descontinuidades nas trajetórias, quando olhadas do ponto de vista de um dos mapas.

Consideramos que os mapas são funções de um estado interno  $\mathbf{S}$  do agente que chamaremos de “contexto” [47]. O contexto determina o padrão de comportamento ou a necessidade do agente. Por exemplo, no caso do agente ser um animal a contexto pode ser a vontade de comer, no caso de ser uma máquina o contexto pode ser uma tarefa relacionada a um determinado objeto no ambiente. Portanto, o estado  $\mathbf{S}$  determina o mapa a ser usado.

A reposição aleatória do agente após ter encontrado o alvo, para um dado mapa, simula um grande número de outros alvos uniformemente distribuídos. Porém isto nem sempre acontece, pois alguns ambientes podem apresentar apenas alguns alvos de interesse. Nes-



tes casos observamos o desenvolvimento de importantes rotas conectando os alvos entre si. A existência destas rotas podem também depender da associação entre os alvos. Tais associações podem existir ao nível da dinâmica dos estados  $\mathbf{S}$  que identificam os contextos. Elas podem representar tarefas onde os alvos devam ser encontrados em uma determinada seqüência, por exemplo. A figura 3.2 mostra isto. Todas as situações novas resultam em um problema de contorno específico da equação de Poisson.



**Fig. 3.2:** (a), (b) e (c) mostram o campo vetorial após 12000 tentativas para uma associação aleatória entre 3 alvos. (d), (e), (f) mostram o campo vetorial para 3 alvos que devem ser encontrados em uma seqüência pré-determinada. Em ambos os casos  $\alpha = 0.01$  and  $\theta = 0.2$ .

Um fenômeno interessante apresentado em [37] e visto na seção 2.2.2 é uma tendência repulsiva causada pelas paredes. Isto se deve a componente estática do aprendizado que surge como consequência do campo receptivo finito das células de posição. Aqui este efeito não está presente. A explicação é simples: as células próximas a parede têm em média

uma quantidade igual de componentes entrando e saindo da parede. Portanto elas possuem um campo vetorial nulo. Em outras palavras, o problema de contorno tem as condições de contorno de Neumann do tipo reflexiva  $\mathbf{n} \cdot \nabla \rho(\mathbf{r}) = 0$  para  $\mathbf{r} \in \partial\Gamma_{obstáculo}$ .

### 3.3 Usando os mapas durante o aprendizado

Vimos na última seção que o movimento aleatório no ambiente com alvos pode gerar um campo vetorial não trivial e muito expressivo sem a necessidade de supervisão. A simples alternância entre os mapas resultou na organização campos vetoriais que apontavam para o alvo responsável por esta alternância.

Até este ponto seria válido questionar a respeito do nome “alvo” usado aqui, pois isto não passa de uma operação que alterna entre os mapas internos e não um objetivo pré-definido para o animal.

Apesar disso, o campo vetorial revela que em média este movimento aleatório é na direção do alvo. A fim de transformar os alvos em objetivos para o animal, a informação armazenada nos mapas deve ser usada com um acessório para a navegação. Para isto mudamos a dinâmica de nossas simulações para incorporar o mapa da seguinte maneira

$$\begin{aligned} \mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \alpha \frac{\Delta \mathbf{r}(t) + \beta \mathbf{v}_{ab}}{\sqrt{1 + 2\beta \mathbf{v}_{ab} \cdot \Delta \mathbf{r}(t) + \beta^2 |\mathbf{v}_{ab}|^2}} \\ \Delta \mathbf{r}(t) &= (\cos \phi(t), \sin \phi(t)) \\ \phi(t) &= \phi(t - \Delta t) + \Delta \phi \end{aligned} \quad (3.9)$$

onde  $\mathbf{r}(t) \in C_{ab}$  e  $\Delta \phi$  é uma variável aleatória com a distribuição já mencionada. Com o qual garantimos que o movimento ainda é com passos de tamanho  $\alpha$  e que o campo vetorial domina a direção para grandes  $|\mathbf{v}|$ .

Em síntese a idéia deste modelo é que existe uma fase de aprendizado onde o mapa é construído e um uma fase de conduta onde o mapa é seguido. Nas simulações calculamos  $\mathbf{v}_{ab}$  sem a normalização, dessa forma a norma cresce com o número de vezes que o animal passa em  $C_{ab}$ . Portanto, a influência em (3.9) aumenta com o aprendizado e a dinâmica muda continuamente de um passeio aleatório para um passeio direcionado pelo mapa. A rapidez com que esta transição ocorre depende da frequência com que o ambiente é explorado e do

parâmetro  $\beta$  que está relacionado com o ruído no movimento. Isto é equivalente a fazer um *simulated annealing* através do aumento do parâmetro relacionado a temperatura  $\beta$  para um campo vetorial com norma fixa.

Aprendizado e conduta não necessariamente precisam ser excludentes. Entretanto, para melhor investigar a formação do mapa comparamos duas formas de aprendizado:

- Aprendizado exploratório

O agente tem seu estado interno ajustado para um estado exploratório  $\mathbf{E}$ . Ele executa um movimento aleatório no ambiente. A trajetória é armazenada em um mapa temporário  $\Delta\mathbf{v}_a$  e  $\Delta\rho_a$  até o momento em que um alvo relevante seja reconhecido. Isto provoca o surgimento do estado interno  $\mathbf{S}$  e o campo vetorial armazenado no mapa temporário é passado para o mapa  $\mathbf{S}$  atualizando  $\mathbf{v}_{ab}(\mathbf{S}) = \mathbf{v}_{ab}(\mathbf{S}) + \Delta\mathbf{v}_{ab}$  e  $\rho_{ab}(\mathbf{S}) = \rho_{ab}(\mathbf{S}) + \Delta\rho_{ab}$ . Após o estado muda novamente  $\mathbf{S} \rightarrow \mathbf{E}$ .

- Aprendizado direcionado

O agente tem seu estado interno previamente ajustado para um estado  $\mathbf{S}$ . Este estado seleciona um mapa específico  $\mathbf{v}_{ab}(\mathbf{S})$  que é usado para a navegação. Um mapa temporário também é criado onde instâncias parciais do campo de deslocamentos e o tempo de permanência são armazenados. A atualizações  $\mathbf{v}_{ab}(\mathbf{S}) = \mathbf{v}_{ab}(\mathbf{S}) + \Delta\mathbf{v}_{ab}$  e  $\rho_{ab}(\mathbf{S}) = \rho_{ab}(\mathbf{S}) + \Delta\rho_{ab}$  são completadas quando o alvo consistente com o estado  $\mathbf{S}$  é encontrado. Neste momento o estado interno muda  $\mathbf{S} \rightarrow \mathbf{S}'$ , um novo mapa é acionado e a procura recomeça.

O aprendizado exploratório é o que chamamos de exploração aleatória nas seções anteriores. O aprendizado direcionado acopla a fase de aprendizado e a fase de conduta, podendo também ser chamado de exploração direcionada.

Quando o mapa é acionado durante o aprendizado, a frequência de passagens por uma posição com a mesma direção do movimento causa um aumento na probabilidade de que a mesma direção seja escolhida na próxima vez. Este reforço lembra alguns modelos teóricos usados para explicar a formação de rios onde o fluxo de água esculpe o leito do rio e ao mesmo tempo é determinado pelo gradiente da superfície do leito do rio [48, 49]. O resultado

é o desenvolvimento de uma complexa rede de trajetórias que leva ao alvo. A estrutura formada é bastante dependente da história especialmente das primeiras trajetórias usadas para encontrar o alvo. Entretanto, diferentemente dos rios, aqui não existe um gradiente inicialmente apontando para o alvo (cuja a localização é inicialmente desconhecida). Além disso, devido ao ruído o agente pode andar o bastante para gerar trajetórias que se cruzam. O reforço desses laços (*loops*) pode gerar mapas com armadilhas (rotacionais do campo vetorial diferentes de zero) que impedem o agente de alcançar o alvo. A presença dos rotacionais implica que os mapas não são mais o gradiente da solução do problema de Poisson.

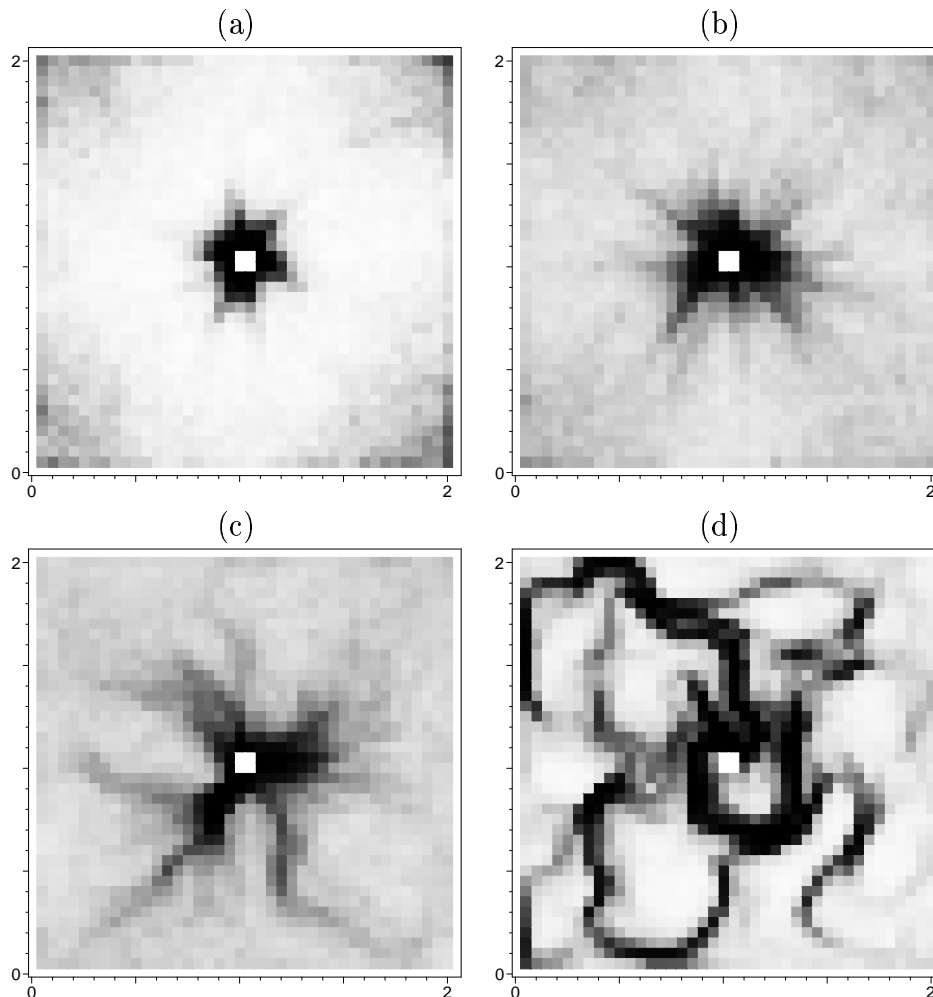
Os loops são facilmente produzidos, especialmente em ambientes finitos onde existem espalhamento nas paredes. Para evitar este efeito prejudicial no aprendizado Blum e Abbott [10] propuseram que apenas a história recente de abordagem do alvo deve ser armazenada no mapa. Os loops em geral aparecem longe do alvo, sendo que para as regiões próximas do alvo as trajetórias tendem ser mais retilíneas. No modelo citado anteriormente, Blum e Abbott consideraram que as atualizações dos pesos sinápticos eram ponderadas por um fator de tempo exponencial. No nosso caso é equivalente considerar que, ao invés de usar a equação (3.3), a direção preferencial do campo vetorial é calculada de acordo com

$$\mathbf{v}_{ab}(t) = \frac{1}{t} \sum_{p=1}^k \sum_{t'=1}^{t_p} e^{-\frac{(t_p-t')}{\tau}} (\Delta \mathbf{r}(t') \delta_{\mathbf{r}(t') \in C_{ab}}) \quad (3.10)$$

onde  $t_p$  é o gasto para realizar a trajetória  $p$ .

O efeito do fator exponencial é criar uma região de memória ao redor do alvo. Em nossas simulações esta região é grosseiramente  $\alpha \times \tau$ . A figura 3.3 mostra a estrutura de rios do mapa resultante para diferentes valores de  $\tau$ . Em analogia ao que é feito na literatura sobre rios, mostramos a taxa de visitação para um grande número de tentativas onde o agente é posicionado aleatoriamente no ambiente podendo seguir o mapa usando a dinâmica da equação (3.9), sem o aprendizado. As regiões mais escuras são as mais frequentemente usadas pelo agente quando ele se move para o alvo.

Vemos que grandes valores de  $\tau$  tendem a produzir trajetórias mais retorcidas, com isso reduzindo a eficiência do mapa. A figura 3.3d mostra o caso extremo  $\tau = \infty$ , onde é possível ver a formação de loops. Neste caso, se o aprendizado prosseguir um ponto crítico



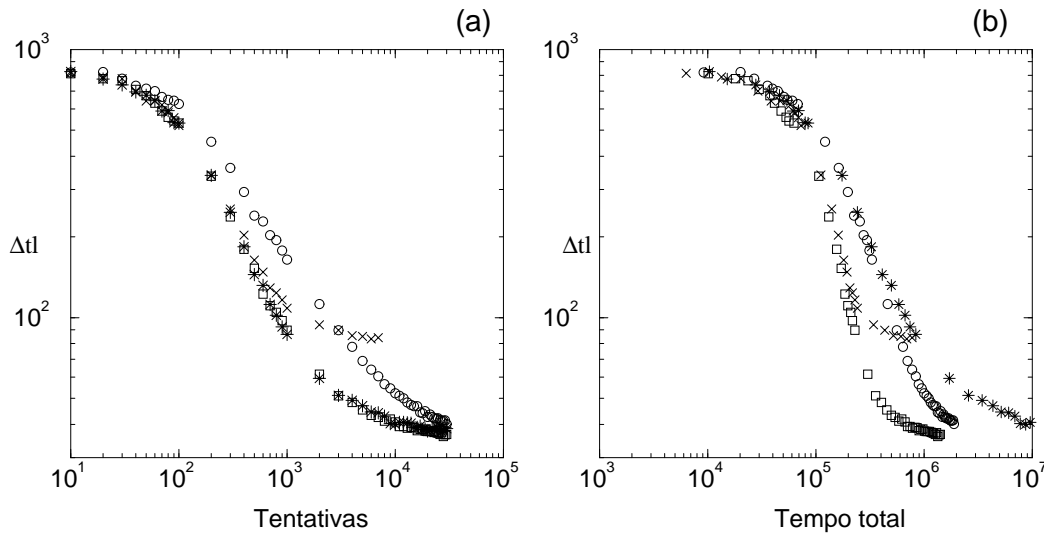
**Fig. 3.3:** *Estrutura de rios de um mapa para diferentes valores de  $\tau$ , durante o aprendizado direto em um ambiente  $2 \times 2$  com  $\alpha = 0.01$  e  $\theta = 0.2$ . (a)  $\tau/L = 25$ , (b)  $\tau/L = 50$ , (c)  $\tau/L = 500$  e (d)  $\tau/L = \infty$ .*

é alcançado impossibilitando que o agente alcance o alvo, com isso inutilizando o mapa.

Por outro lado, mapas gerados com pequenos valores de  $\tau$  são muito similares aos mapas gerados pelo aprendizado exploratório em termos da orientação do campo vetorial. Podemos ver nas figuras 3.3a e 3.3b que as trajetórias preferenciais são linhas retas como as indicadas na figura 3.1b. Mas elas não têm uma dependência de longo alcance, o que é uma característica do aprendizado exploratório. A intensidade do campo vetorial  $|\mathbf{v}_{ab}|$  decresce exponencialmente com a distância do alvo. Portanto, longe do alvo, o mapa é fraco e o movimento do agente quando segue o mapa é aproximadamente aleatório. Isto é visível na

figura 3.3. A situação ótima é onde  $\tau \times \alpha \sim L$ .

Vemos que o aprendizado direcionado não é tão bom quanto o aprendizado exploratório na geração do mapa de navegação. Apesar disso, o aprendizado direcionado tem a vantagem de rapidamente propiciar um mapa útil. Isto pode ser observado calculando-se tempo médio que o agente leva para alcançar o alvo (tempo de latência,  $\Delta t_l$ ) a partir de qualquer posição aleatória no ambiente como função do aprendizado. Medimos o aprendizado pelo número de tentativas  $m$  de aprender como chegar no alvo, figura 3.4a, e o tempo total gasto durante o aprendizado  $\sum_{p=1}^m t_p$ , figura 3.4b.



**Fig. 3.4:** Tempo de latência médio em função do aprendizado  $\tau/L = 25$  ( $\circ$ ),  $\tau/L = 250$  ( $\square$ ) e  $\tau/L = \infty$  ( $\times$ ) com aprendizado direcionado,  $\tau/L = \infty$  ( $*$ ) com aprendizado exploratório, para  $\alpha = 0.01$  e  $\theta = 0.2$  em um ambiente  $1 \times 1$  30.000 tentativas de aprendizado, exceto ( $\times$ ) que tem 6.000 tentativas. (a) Tempo médio de latência com função do número de tentativas de aprendizagem. (b) Tempo médio de latência como função do tempo total gasto durante o aprendizado.

A tentativa de aprendizado está compreendida entre o momento em que determinado mapa é acionado e o ponto em que o correspondente alvo é encontrado. Na figura 3.4a observamos que a redução muito grande do  $\tau$  afeta negativamente o desempenho do processo de aprendizado. Para pequenos valores de  $\tau$  o agente tem que estar muito próximo do alvo para “sentir” a presença do mesmo, sendo que longe o alvo o agente realiza um passeio

aleatório no ambiente. Para  $\tau \times \alpha$  da ordem do tamanho do ambiente (curva ( $\square$ )) obtemos a situação ótima, onde o mapa gerado equivale em desempenho ao mapa produzido pelo aprendizado exploratório. O aumento de  $\tau$  causa uma diminuição no desempenho culminando em loops (na curva ( $\times$ ) onde  $\tau = \infty$  o programa foi interrompido antes das 30.000 tentativas de aprendizado).

A figura 3.4b mostra como o desempenho muda se o parâmetro importante é o tempo total de aprendizado. Vemos que o aprendizado exploratório é a pior estratégia de aprendizado.

### 3.4 Modelos de rede

Como vimos no capítulo anterior, é possível capturar a informação armazenada no campo vetorial usando redes neurais. No modelo que Blum e Abbott propuseram a posição codificada pelas células de posição é deslocada na direção do objetivo devido a ação do aprendizado de Hebb [30] sobre estas células. O aprendizado de Hebb pode ser escrito como

$$w_{ij} = \sum_p \int dt \int dt' g(t') s_i[\mathbf{r}_p(t)] s_j[\mathbf{r}_p(t-t')] \quad (3.11)$$

onde a soma sobre  $p$  representa a série de trajetórias e  $s_i(\mathbf{r})$  é a atividade do neurônio  $i$ , que tem o campo receptivo centrado em  $\mathbf{R}_i$ , quando o rato está na posição  $\mathbf{r}$ . Usualmente  $s_i(\mathbf{r}) = s(|\mathbf{r} - \mathbf{R}_i|)$ , com  $s(x)$  sendo uma função Gaussiana. A função  $g(t')$  é a função janela da LTP que indica como o aumento da eficiência sináptica varia com o intervalo de tempo  $t'$  entre as atividades pré e pós-sináptica.  $g(t') = 0$  para  $t' < 0$ , pois os experimentos de LTP mostram que a potenciação sináptica ocorre apenas se a atividade pré-sináptica  $j$  preceder a atividade pós-sináptica  $i$ .

A atividade dos neurônios nesta rede é dado por

$$\sigma_i(\mathbf{r}) = f \left[ \sum_j w_{ij} \sigma_j(\mathbf{r}) + s_i(\mathbf{r}) \right] \quad (3.12)$$

onde  $f[h]$  é a função transferência.

As atividades neurais  $\{s_i(\mathbf{r})\}_{i=1,\dots,N}$  codificam a posição  $\mathbf{r}$ , por outro lado, as atividades  $\{\sigma_i(\mathbf{r})\}_{i=1,\dots,N}$  quando decodificadas [38, 18, 40] indicam a posição deslocada  $\mathbf{r}^*(\mathbf{r})$ . A

diferença entre a posição real e a posição deslocada, devido a integração das trajetórias na equação (3.10), aponta para a direção preferencial seguida pelo animal. Portanto, podemos considerar o mapa formado pela posição deslocada  $\Delta(\mathbf{r}) = \mathbf{r}^*(\mathbf{r}) - \mathbf{r}$  com sendo qualitativamente equivalente ao campo vetorial definido anteriormente na equação (3.3). Realmente, podemos afirmar que no limite quando o tamanho da subdivisão do ambiente  $C_{ab}$  tende a zero,  $\mathbf{v}_{ab}$  torna-se um campo vetorial contínuo e

$$\mathbf{v}(\mathbf{r}) = \Delta(\mathbf{r}) F[\mathbf{r}^*(\mathbf{r}) - \mathbf{r}] \quad (3.13)$$

onde  $F$  é uma função escalar.

Em resumo a idéia básica deste modelo é que a estatística de passagens em uma certa posição no ambiente reforça as conexões sinápticas, da rede que codifica as posições, somente em uma direção. O resultado é que quando a rede é estimulada após o aprendizado ocorre a propagação da ativação causada pela potenciação das sinápses em direção dos neurônios que codificam a próxima posição [50].

Esta idéia pode ser estendida, se as sinápses retardadas são usadas [51, 52, 53] toda a trajetória até o alvo pode ser recuperada. Considerando que a função janela pode ser escrita como

$$g(t') = \sum g_n(t'; \delta_n) \quad (3.14)$$

onde cada  $g_n$  é responsável pelas correlações com um intervalo de tempo específico  $\delta_n$ . Em particular,  $n = 0$  pode ser uma função impulso ( $g_0(t', \delta_0 = 0) = \delta(t')$ ) que, para campos receptivos sobrepostos, corresponde a atividade sincronizada armazenada na parte estática do mapa ( $\rho_{ab}$ ).

Além da propagação da ativação nas células de posição existem outras alternativas possíveis de armazenamento de trajetórias. Por exemplo, existem neurônios no hipocampo que respondem seletivamente não apenas às posições mas também à direção do movimento. Durante a recuperação de uma trajetória eles podem fornecer diretamente a indicação da direção a ser seguida. Uma possibilidade é que estes neurônios estejam uma camada acima dos neurônios descritos anteriormente, onde eles detectam o deslocamento na posição codificada pelas atividades  $\{s_i(\mathbf{r})\}_{i=1,\dots,N}$  e as atividades  $\{\sigma_i(\mathbf{r})\}_{i=1,\dots,N}$ . Células de posição que



selecionam a direção podem também surgir a partir de células que conectam-se localmente a um determinado número de células de posição, mas elas próprias não possuem campos receptivos de posições. A seletividade de posições vem da seletividade de posições de outras células e intensa seletividade de direção vem da assimetria das conexões.

## 3.5 Conclusões

O ponto principal do trabalho apresentado neste capítulo foi mostrar que uma operação local pode desbalancear a estatística do movimento aleatório em um ambiente para permitir a formação de um mapa de direções preferenciais. Este mapa pode ser aprendido por uma rede de uma forma não supervisionada. O operador tem uma localização fixa no ambiente e o mapa resultante acaba indicando esta localização. O operador naturalmente torna-se um alvo.

A operação efetuada pelo alvo é uma mudança do estado interno do agente. Uma maneira diferente de ver isto é considerar que o alvo representa alguma coisa importante o suficiente para causar uma alteração interna no agente. O estado interno que foi alterado deve agir como uma sugestão para um sistema de memória que armazena de alguma maneira o mapa das direções preferenciais. Isto pode ser apenas uma indicação ou alguma coisa mais complexa como um padrão de atividade neural que traz à tona um dentre os vários mapas em uma rede neural.

Discutimos também o efeito de se usar mapas para guiar o agente durante o aprendizado. Na fase inicial do aprendizado os alvos são encontrados acidentalmente. As primeiras trajetórias que chegam até o alvo assemelham-se com passeios aleatórios. No aprendizado direcionado, elas são adicionadas ao mapa e a forma seguinte é aprendida.

Se o grau de ruído ( $1/\beta$ ) na equação (3.9) é grande ( $\beta \ll 1$ ) o aprendizado é menos sensível às trajetórias iniciais. No caso extremo ( $\beta = 0$ ) temos o aprendizado exploratório e um mapa ótimo é produzido. Obter o mapa ótimo por meio de um aprendizado exploratório não necessariamente é a melhor solução, especialmente se o tempo é um fator importante no que diz respeito a desempenho. Como vimos na figura 3.4, os agentes que seguem um movimento exploratório (bastante ruidoso) demoram para encontrar o alvo. Aquilo que

parece ótimo em termos das tentativas de aprendizado torna-se muito custoso em termos de tempo. Por outro lado, um baixo ruído produz trajetórias sub-ótimas muito rapidamente, mas o ganho obtido em tempo de formação do mapa é perdido no momento da utilização deste mapa. Realmente, o aprendizado torna-se instável podendo produzir mapas que são completamente inúteis devido a existência de loops.

Esta situação pode ser evitada se modularmos o tamanho da contribuição da trajetória que será incorporado ao mapa dependendo do tempo gasto para encontrar o alvo. Ajustando  $\tau$ , que é o parâmetro que controla quanto tempo no passado uma trajetória deve ser adicionada na equação (3.10), podemos reduzir alguns dos efeitos do movimento aleatório sobre as primeiras versões do mapa e evitar os loops.

Acreditamos que uma escolha adequada de  $\beta$  e  $\tau$  pode produzir um ótimo aprendizado para vários tipos de ambientes e disposições de alvos. Para isto é necessário fazer uma cuidadosa pesquisa no espaço de parâmetros do sistema.

## Capítulo 4

# Utilizando redes neurais para armazenar diferentes mapas

No capítulo anterior introduzimos o conceito de contexto como sendo formas de recuperação de mapas de navegação. Neste capítulo apresentamos uma maneira para armazenar e recuperar mapas de navegação, para diferentes ambientes, em uma rede neural associando cada ambiente a um estado interno específico da rede neural que chamamos de contexto. O modelo é construído sobre a idéia de que o mapa de navegação pode ser escrito como uma solução da equação de Laplace com condições de contornos específicas aplicadas aos obstáculos e aos objetivos presentes no ambiente. A compressão de informação obtida permite-nos obter um bom desempenho no armazenamento dos mapas com um reduzido número de conexões sinápticas [54] .

Também no capítulo anterior, mencionamos uma técnica que faz uso de funções harmônicas no planejamento de trajetórias de robôs. Esta técnica é muito interessante, pois as funções harmônicas são bastante conhecidas e utilizadas em várias áreas, como por exemplo na matemática, física e engenharia. Para nós, o importante é que o cálculo numérico das funções harmônicas envolve uma dinâmica onde o valor do potencial nos sítios vizinhos é substituído pelo valor médio dos sítios vizinhos. Isto é semelhante a dinâmica de uma rede neural atratora.

Neste capítulo vamos propor uma ligação entre os algoritmos artificiais e os modelos biológicos para o sistema de memória espacial e de navegação. Em nossa analogia, entretanto, as

atividades neurais que resultam do mapa harmônico representam a probabilidade de colisão<sup>1</sup> [43] em vez da posição do agente em um sistema de referência absoluto. Extendemos esta idéia combinando a rede de células que representam a probabilidade de colisão com uma rede de contexto para armazenar múltiplos ambiente.

A seguir descrevemos o método das funções harmônicas e o modelo completo da rede neural que implementa este método para vários ambientes. Os resultados são discutidos no final do capítulo.

## 4.1 O método das funções harmônicas

Uma função harmônica no domínio  $\Gamma \subset R^d$  é uma função que satisfaz a equação de Laplace

$$\nabla^2 p = 0 \quad (4.1)$$

No caso do planejamento de trajetórias em robótica o contorno de  $\Gamma$  consiste dos contornos de todos os obstáculos ( $\partial\Gamma_o$ ) e alvos ( $\partial\Gamma_a$ ).

No método das funções harmônicas o ambiente é explorado e as posições dos obstáculos e dos objetivos são armazenados em uma rede. A equação de Laplace (4.1) é resolvida numericamente sob as condições de que os sítios ocupados pelos obstáculos tenham o valor fixado em  $p_o$  e os sítios ocupados pelos alvos tenham o valor fixado em  $p_a$ , com  $p_a < p_o$ . Para fins práticos consideramos  $p_o = +1$  e  $p_a = -1$ .

Para um ambiente bidimensional o cálculo numérico pode ser implementado pela seguinte dinâmica, também conhecida como método de Jacobi:

$$p_{\mathbf{r}}(t+1) = \begin{cases} +1 & \text{se obstáculo} \\ -1 & \text{se alvo} \\ h_{\mathbf{r}}(t) & \text{caso contrário} \end{cases} \quad (4.2)$$

$p_{\mathbf{r}}(t)$  é a atividade da célula que representa a probabilidade de colisão ( potencial ) no ponto  $\mathbf{r}$  da grade de ocupação. A função  $h_{\mathbf{r}}(t)$  depende do algoritmo que usado para resolver o

<sup>1</sup> A probabilidade de colisão é definida da seguinte forma: Considere um processo Markoviano em uma rede e considere  $\mathcal{A}$  como sendo o conjunto dos sítios absorvedores desta rede. Dividindo  $\mathcal{A}$  em dois subconjuntos disjuntos  $\mathcal{O}$  e  $\mathcal{G}$ , então a probabilidade de colisão  $p(x)$  em um determinado estado  $x$  é a probabilidade que um processo, iniciado em  $x$ , seja absorvido em  $\mathcal{O}$  antes de atingir algum estado em  $\mathcal{G}$ .

problema de Laplace. Para o caso de primeiros vizinhos em uma rede quadrada as posições são escritas como  $\mathbf{r} = (i, j)$  e com isso temos

$$h_{ij}(t) = \frac{1}{4} \left( p_{i+1,j}(t) + p_{i-1,j}(t) + p_{i,j+1}(t) + p_{i,j-1}(t) \right) \quad (4.3)$$

Se considerarmos também os segundos vizinhos teremos

$$\begin{aligned} h_{ij}(t) &= \frac{1}{5} \left( p_{i+1,j}(t) + p_{i-1,j}(t) + p_{i,j+1}(t) + p_{i,j-1}(t) \right) \\ &+ \frac{1}{20} \left( p_{i+1,j+1}(t) + p_{i-1,j+1}(t) + p_{i+1,j-1}(t) + p_{i-1,j-1}(t) \right) \end{aligned} \quad (4.4)$$

Portanto genericamente podemos escrever

$$h_{\mathbf{r}}(t) = \sum_{\langle \mathbf{r}' \rangle} W_{\mathbf{r}\mathbf{r}'} p_{\mathbf{r}'}(t) \quad (4.5)$$

onde  $\langle \mathbf{r}' \rangle$  são as somas sobre os vizinhos, os pesos  $W_{\mathbf{r}\mathbf{r}'}$  são conectados de acordo com a topologia da grade de maneira que o estado estável da rede seja a solução da equação de Laplace.

A dinâmica (4.2) fixa o valor das atividades das células que representam obstáculos e alvos e permite que as células de espaço livre (células cuja posição não está ocupada nem por obstáculo nem por alvo) relaxem. O potencial harmônico gravado nas atividades células de espaço livre é a interpolação entre as atividades das células que representam obstáculos e alvos. Portanto o mapa de navegação é obtido diretamente do gradiente descendente das atividades.

## 4.2 O modelo da rede

Existem dois elementos neurais distintos que precisamos considerar. As células relacionadas a probabilidade de colisão, discutidas anteriormente, representam regiões no ambiente. Elas são neurônios contínuos cujas atividades, limitadas entre  $p_0$  e  $p_a$ , dependem da localização dos alvos e obstáculos no ambiente. Introduzimos um segundo conjunto de neurônios chamado de neurônios de contexto. Estes neurônios são binários cujos padrões de atividades

estão associados, via aprendizado sináptico, a uma configuração específica de ambiente que são representadas por um padrão de três estados

$$\rho_{\mathbf{r}} = \begin{cases} +1 & \text{se há um obstáculo no sítio } \mathbf{r} \\ -1 & \text{se há um alvo no sítio } \mathbf{r} \\ 0 & \text{para um sítio livre} \end{cases} \quad (4.6)$$

Consideramos uma rede composta de três sub-redes com as seguintes propriedades:

- *Rede 1:* É uma rede feed-forward que associa o padrão de ambiente  $\{\rho_{\mathbf{r}}^{\mu}\}_{\mathbf{r} \in \Gamma_1}$  ao padrão de contexto  $\{\xi_i^{\mu}\}_{i=1, \dots, M}$ .
- *Rede 2:* É uma rede feed-forward que associa o padrão de contexto  $\{\xi_i^{\mu}\}_{i=1, \dots, M}$  ao padrão de ambiente  $\{\rho_{\mathbf{r}}^{\mu}\}_{\mathbf{r} \in \Gamma_1}$ .
- *Rede 3:* É uma rede atratora com conexões locais que implementam a dinâmica na equação (4.2).

### 4.2.1 Rede 1 - Recuperação do contexto

Um determinado ambiente pode ser associado ao um estado interno específico (contexto) através da intensidade sináptica. Esta associação pode ser obtida através de um algoritmo de aprendizado supervisionado. Consideramos um conjunto de treinamento composto de  $P$  padrões de ambiente com correspondente padrão de contexto

$$(\{\rho_{\mathbf{r}}^{\mu}\}_{\mathbf{r} \in \Gamma}, \{\xi_a^{\mu}\}_{a=1, \dots, M})_{\mu=1, \dots, P}$$

Os padrões  $\rho$  são  $N \times N$  vetores com componentes  $\pm 1$  ou zero. Considerando que a maior parte do ambiente é espaço livre, estes vetores serão bastante esparsos. Os padrões  $\xi$  são vetores  $M$ -dimensional com componentes escolhidas aleatoriamente a partir de uma distribuição

$$P(\xi) = \frac{1}{2} \delta(\xi - 1) + \frac{1}{2} \delta(\xi + 1) \quad (4.7)$$

A atividade de um neurônio de contexto é dada por

$$s_a = \text{sgn}(h_a) \quad (4.8)$$

com o campo local dado por

$$h_a = \sum_{\mathbf{r}} J_{a\mathbf{r}} p_{\mathbf{r}} \quad (4.9)$$

Os pesos sinápticos são calculados de acordo com uma regra de Hebb iterativa (por exemplo o algoritmo de Rosenblatt) com conexões dadas por

$$\Delta J_{a\mathbf{r}} = \Theta(-h_a^\mu \xi_a^\mu) \eta \xi_a^\mu \rho_{\mathbf{r}}^\mu \quad (4.10)$$

onde  $\Theta(x)$  é a função Heaviside. A rede 1 permite a recuperação de um contexto após uma exploração parcial de um ambiente conhecido.

### 4.2.2 Rede 2 - Recuperação do ambiente

Os contextos por sua vez podem ser usados como uma chave para recuperar a configuração completa das posições dos obstáculos e dos alvos no ambiente. Para isso deve ser feita e armazenada uma associação entre o contexto e o ambiente. Esta é a regra do segundo grupo de conexões que chamamos de rede 2. Estas sinapses mandam informações das células do contexto para as células do ambiente. A atividade de uma célula do ambiente é dado por

$$m_{\mathbf{r}} = f(H_{\mathbf{r}}, \lambda) \quad (4.11)$$

com uma função de transferência de três estados

$$f(x, \lambda) = \Theta(x - \lambda) - \Theta(-x - \lambda)$$

O parâmetro  $\lambda$  determina quando um dado padrão de entrada deve ser classificado como +1, 0 ou -1. Ou, em outras palavras, se um determinado contexto representa um obstáculo, espaço livre ou um alvo em uma dada posição. Para manter  $\lambda$  independente do aprendizado as entradas locais devem ser normalizadas, portantoo

$$H_{\mathbf{r}} = \frac{1}{|\mathbf{J}_{\mathbf{r}}|} \sum_a J_{\mathbf{r}a} s_a \quad (4.12)$$

com

$$|\mathbf{J}_{\mathbf{r}}| = \sum_a (J_{\mathbf{r}a})^2$$

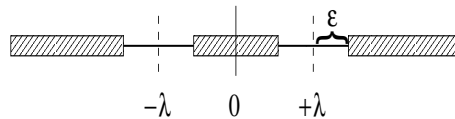
O procedimento adequado para ensinar uma rede de três estados é diferente do procedimento empregado anteriormente. Usamos uma adaptação do algoritmo de Rosenblatt baseado em uma função custo fragmentada (veja figura 4.1) [55]. Para o caso de três estados isto representa

$$\Delta J_{ra} = \begin{cases} +\Theta(\tau_l^+ - H_r^\mu) \eta \xi_a^\mu & \text{se } \rho_r^\mu = +1 \\ -\Theta(H_r^\mu + \tau_l^-) \eta \xi_a^\mu & \text{se } \rho_r^\mu = -1 \\ \{ \Theta(-H_r^\mu - \tau_s^-) - \Theta(H_r^\mu - \tau_s^+) \} \eta \xi_a^\mu & \text{se } \rho_r^\mu = 0 \end{cases} \quad (4.13)$$

onde os parâmetros de aprendizado  $\tau$  definem o intervalo onde o campo de entrada  $H$  permanece após o aprendizado. Considerando

$$\tau_l^\pm = \pm(\lambda + \epsilon)$$

$$\tau_s^\pm = \pm(\lambda - \epsilon)$$



**Fig. 4.1:** Função custo fragmentada.

A presença do parâmetro  $\epsilon$  produz um aprendizado mais severo do que o imposto pelo conjunto de treinamento associado a equação (4.11). De fato, para aprender o conjunto de treinamento  $\epsilon = 0$  é suficiente. Entretanto, um  $\epsilon$  finito permite uma recuperação mais robusta na presença de ruídos adicionais. Ambos  $\lambda$  e  $\epsilon$  podem ser otimizados para produzir um armazenamento máximo para um dado conjunto de treinamento e nível de ruído.

### 4.2.3 Rede 3 - Geração do mapa de navegação

A rede de células que representam a probabilidade de colisão integra a recuperação do ambiente com a geração do mapa usado para a navegação. Uma maneira natural para fazer



isto é considerar a seguinte dinâmica

$$p_{\mathbf{r}} = g(h_{\mathbf{r}}|H_{\mathbf{r}}, \lambda) \quad (4.14)$$

onde  $h_{\mathbf{r}}$  é a entrada local definida na equação (4.5) e  $H_{\mathbf{r}}$  é o contexto de entrada definida na equação (4.12). A função de transferência é

$$g(x|y, \lambda) = \begin{cases} +1 & \text{se } y > \lambda \\ x & \text{se } y \in [-\lambda, \lambda] \\ -1 & \text{se } y < -\lambda \end{cases} \quad (4.15)$$

Conseqüentemente os padrões de ambiente armazenados  $\rho^{\mu}$ , servem como uma máscara para a dinâmica de Laplace.

A equação (4.14) em conjunto com (4.8) e (4.11) compõem um atrator do sistema que armazena mapas de navegação para múltiplos ambientes. Neste caso estamos interessados apenas na demonstração da plausibilidade. Para isso implementamos o armazenamento de ambientes simples.

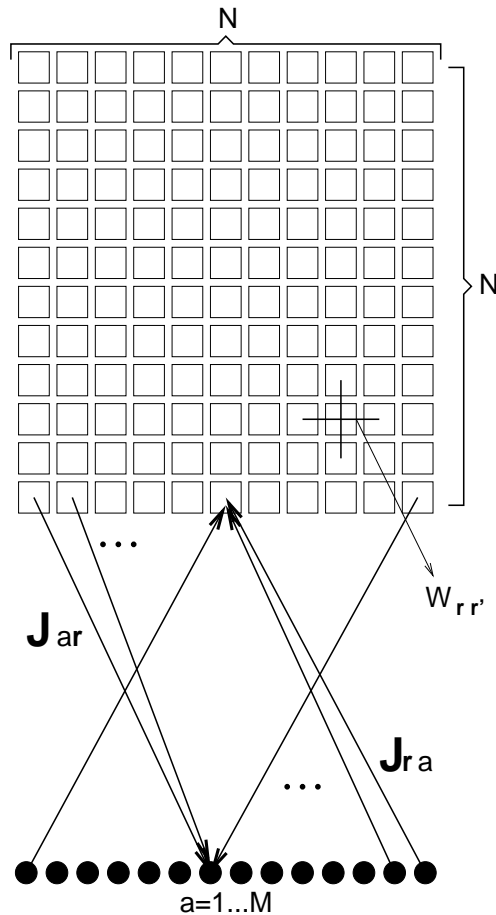
A figura 4.2 mostra a arquitetura geral da rede para um pequeno ambiente.

### 4.3 Resultados

As figuras 4.3 e 4.4 mostram o ambiente real e o mapa de navegação calculado pela rede neural. As figuras exemplificam a recuperação do mapa sugerida pelo contexto.

Durante o aprendizado os padrões de ambiente foram apresentados de maneira cíclica e a equação (4.13) foi calculada em um período de 30 iterações. No caso específico das figuras 4.3 e 4.4 três diferentes ambientes foram armazenados. Os parâmetros de aprendizado  $\lambda = 0.4$  e  $\epsilon = 0.2$  produziram os melhores resultados pra o armazenamento dos dados.

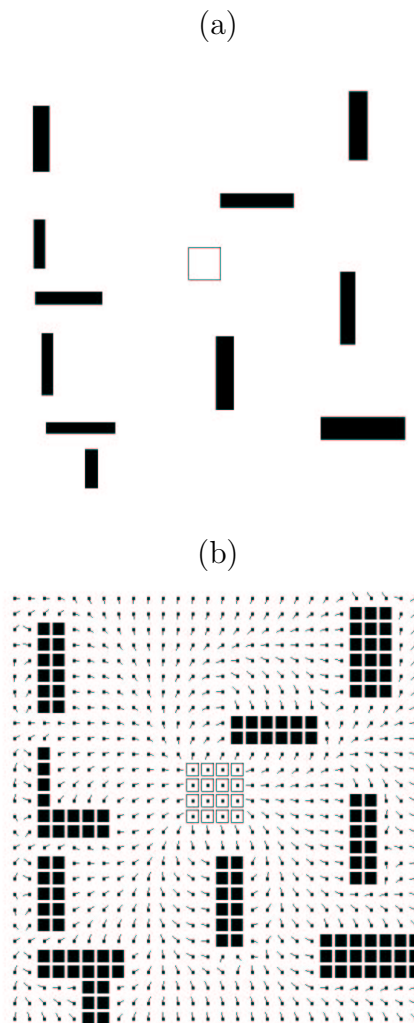
Na 4.3b e 4.4b os quadrados preto e branco representam as porções do ambiente recuperadas como obstáculos e alvos respectivamente pela rede de contexto. As setas representam o gradiente descendente do potencial do potencial harmônico calculado pela equação (4.14) usando a expressão (4.3) após 300 iterações.



**Fig. 4.2:** *Arquitetura da rede neural para um pequeno ambiente coberto por  $N \times N$  células que representam a probabilidade de colisão e uma rede de recuperação de contexto com  $M$  componentes de entrada.*

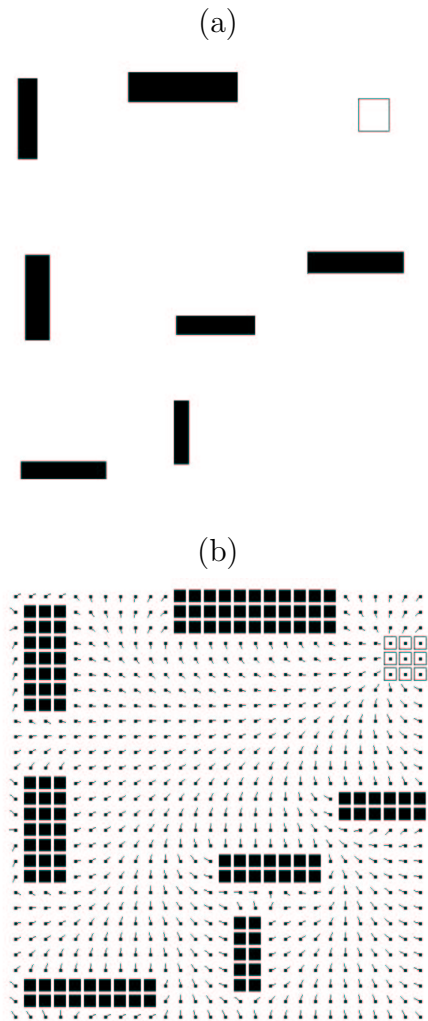
## 4.4 Conclusões

A rede proposta é composta de  $L = N^2 + M$  neurônios. Mas o número de pesos sinápticas escala com  $\sim N^2(M + 2)$ . O que é mais importante é que o número de sinapses que são modificadas é igual a  $\sim N^2M$ . Isto não é comum em redes neurais atratoras, como a rede de Hopfield, cujo o número de sinapses é sempre da ordem de  $\sim L^2$ . De fato, para o armazenamento de padrões aleatórios é sempre melhor usar o número máximo de parâmetros sinápticos possível. Apesar disso, o uso de um baixo número de sinapses não reduz o desempenho da rede neste caso. Isto é possível porque os padrões que interessam ser armazenados não são aleatórios. Eles compartilham de uma importante propriedade



**Fig. 4.3:** Mapa de navegação para um ambiente com alguns obstáculos (retângulos pretos) e um alvo (retângulos brancos). (a) O ambiente real. (b) Mapa de navegação calculado por uma rede neural com  $N \times N = 30 \times 30$  e  $M = 64$ .

que é o fato deles serem soluções da equação de Laplace para as condições de contorno de Dirichlet. Realmente, as condições de contorno que descrevem o ambiente determinam completamente todo o padrão de atividade das células que representam a probabilidade de colisão, portanto é necessário armazenar apenas as condições de contorno. Isto é um bom exemplo de compressão de informação obtida através de um sistema físico.



**Fig. 4.4:** Outra configuração de obstáculos e alvo. (a) O ambiente real. (b) Mapa de navegação calculado pela mesma rede neural especificada anteriormente.

## Capítulo 5

# Potencial ótimo para navegação

No capítulo anterior apresentamos uma rede neural artificial capaz de armazenar e recuperar mapas de navegação para diferentes ambientes. Esta rede basicamente guarda as localizações dos obstáculos e dos objetivos que estão inseridos no ambiente e depois resolve a equação de Laplace para este ambiente. As condições de contorno são dadas de maneira que as posições ocupadas pelos obstáculos tenham o valor fixado em  $p_o$  e as posições ocupadas pelos alvos tenham o valor fixado em  $p_a$ , com  $p_a < p_o$ . O resultado desse cálculo é uma função harmônica que não possui mínimos além daqueles pontos onde estão situados os objetivos. Com isso, qualquer dinâmica que siga o gradiente descendente dessa função harmônica levará a um dos objetivos. Por exemplo, considere uma dinâmica da seguinte forma

$$\frac{d\mathbf{r}}{dt} = -v_0 \frac{\nabla \rho(\mathbf{r}) + \boldsymbol{\eta}(t)}{|\nabla \rho(\mathbf{r}) + \boldsymbol{\eta}(t)|} \quad (5.1)$$

onde o gradiente descendente é seguido com uma velocidade constante  $v_0$  e  $\boldsymbol{\eta}(t)$  é um termo de ruído aleatório que representa o erro na medida do gradiente.

Uma possível falha deste método é que para pontos longe do alvo a solução da equação de Laplace tende a ser plana, especialmente na direção perpendicular ao gradiente. Nestas circunstâncias o termo aleatório na equação (5.1) é dominante sobre o termo gradiente e a dinâmica torna-se ineficiente, ou seja, torna-se um passeio aleatório em duas dimensões. Isto indica que a forma do potencial harmônico pode não ser uma escolha adequada em um ambiente ruidoso. A questão que nos interessa é saber qual a forma que o potencial deve ter para que o objetivo seja alcançado no menor tempo possível e na presença de ruído <sup>1</sup>?

---

<sup>1</sup> Em IDIART, M. A. P.; TREVISAN, M. Optimal potential shapes for navigating a random walker. *Submetido ao Journal of Physics A: Mathematical and General*

Salvo a sua aplicação prática este problema pode ser transformado em um problema físico interessante. De uma forma muito simples podemos imaginar um animal como sendo uma partícula que, devido aos erros inerentes, difunde em uma superfície potencial. O nível de erro pode ser modelado como uma temperatura. O potencial tem um único mínimo, portanto, conduz a partícula na direção deste mínimo. Dentro desta estrutura podemos fazer uma pergunta relacionada ao problema da difusão e que é análoga ao famoso problema da braquistócrona da mecânica: qual é a forma do potencial que minimiza o tempo médio da primeira passagem da partícula para alcançar a posição do alvo para uma dada temperatura?

## 5.1 Teoria geral

O problema do passeio aleatório pode ser estabelecido da seguinte maneira: considere um processo Markoviano em uma rede  $d$ -dimensional onde uma partícula pode saltar apenas para os sítios que são primeiros vizinhos. Supomos que existe uma tendência em saltar em uma determinada direção e que a probabilidade de ir de um sítio para outro depende da intensidade de uma ligação  $\omega(\mathbf{r}, \mathbf{r}')$  existente entre eles. A probabilidade de transição de ir de  $\mathbf{r}$  para  $\mathbf{r}'$  é escrita em termos das ligações como

$$W(\mathbf{r}, \mathbf{r}') = \frac{1}{Z} e^{\beta \omega(\mathbf{r}, \mathbf{r}')} \quad (5.2)$$

onde

$$Z = \sum_{\langle \mathbf{r}'' \rangle} e^{\beta \omega(\mathbf{r}, \mathbf{r}'')} \quad (5.3)$$

e  $\langle \mathbf{r}'' \rangle$  é a soma sobre todos os vizinhos.

Uma forma adequada para o tratamento numérico do problema é através da equação mestra para a probabilidade  $P(n, t)$ , para  $n = 1, \dots, N$ , de encontrar a partícula na posição  $n$  no tempo  $t$  tendo ela começado no ponto  $n_0$ .

$$\frac{\partial P(n, t)}{\partial t} = \sum_m P(m, t) M_{mn} \quad (5.4)$$

onde  $M_{mn} = W_{mn} - \delta_{mn}$ . Considerar um sítio absorvedor  $a$  é equivalente a atribuir  $P(a, t) = 0$ , reduzindo o problema para um conjunto de  $(N - 1)$  equações, (veja por exemplo [56]),

que podem ser escritas como

$$\frac{\partial \mathbf{P}_a(t)}{\partial t} = \mathbf{P}_a(t) \cdot \mathbf{M}_a \quad (5.5)$$

onde  $\mathbf{P}_a$  é um vetor de  $(N - 1)$  dimensões obtidas pela retirada do elemento  $a$  do vetor original  $\mathbf{P}$  e  $\mathbf{M}_a$  é uma matriz  $N - 1 \times N - 1$  obtida pela retirada da coluna  $a$  e da linha  $a$  da matriz  $\mathbf{M}$ .

O tempo médio que uma partícula leva para difundir de um ponto inicial  $n_0$  até o sítio absorvedor é

$$\begin{aligned} \langle t \rangle_{n_0} &= \int_0^{\infty} dt \sum_n P_a(n, t) \\ &= \sum_n (\mathbf{M}_a^{-1})_{n_0 n} \end{aligned} \quad (5.6)$$

onde a condição inicial para  $\mathbf{P}$  é  $P(n, 0) = \delta_{n_0 n}$ .

Um bom mapa de navegação tem que ser capaz de desempenhar uma boa navegação a partir de qualquer ponto do ambiente, então para nós é mais interessante minimizar a média de (5.6) sobre todos os sítios, ou seja

$$\langle t \rangle = \frac{1}{N - 1} \sum_{n_0} \langle t \rangle_{n_0}. \quad (5.7)$$

## 5.2 Problema unidimensional

Com algumas manipulações algébricas é possível mostrar que, quando  $n_0 = 1$  e  $a = N + 1$  são os extremos de uma cadeia unidimensional, o tempo médio da primeira passagem de uma extremidade a outra é

$$\langle t \rangle_1 = N + 2 \sum_{n=1}^{N-1} \sum_{m=n+1}^N e^{-\beta(\omega_m - \omega_n)} \quad (5.8)$$

e se fizermos a média sobre todos os possíveis pontos iniciais temos

$$\langle t \rangle = \frac{1}{2}(N + 1) + \frac{2}{N} \sum_{n=1}^{N-1} \sum_{m=n+1}^N m e^{-\beta(\omega_m - \omega_n)} \quad (5.9)$$

onde  $\omega_n$  é a ligação entre os sítios  $n$  e  $n + 1$ . O problema de otimização que propomos é encontrar o conjunto de ligações  $\omega_n^*$ ,  $n = 1 \dots N - 1$ , que minimiza (5.8) ou (5.9) sujeito

a uma determinada condição. É absolutamente necessário impor uma condição sobre  $\omega_n$ , caso contrário o mínimo nos dois casos ocorre na situação trivial onde todos  $\omega_n = \infty$  com  $\omega_n - \omega_{n-1} = \infty$ . Portanto assumimos que  $\omega_n$  está limitado entre 0 e 1.

O mínimo de (5.9) pode ser obtido através da técnica de Monte Carlo com temperatura zero. Iniciamos com todas as ligações iguais com valor 1/2 e selecionamos aleatoriamente com igual probabilidade as ligações que podem ter o seu valor aumentado ou diminuído de uma certa quantidade. A mudança somente é aceita se  $\langle t \rangle$  diminuir. Não esperamos mínimos locais para o caso unidimensional, portanto, uma minimização determinística é suficiente para encontrar a solução que procuramos.

Para valores grandes de  $N$  a soma em (5.9) pode ser aproximada para uma integral e a expressão para o tempo médio da primeira passagem torna-se

$$\langle t \rangle = \frac{N}{2} + 2 N^2 \Gamma(\beta) \quad (5.10)$$

onde a função  $\Gamma(\beta)$  é definida como

$$\Gamma(\beta) = \int_0^1 dx \int_x^1 dy y e^{-\beta(\omega(y)-\omega(x))}. \quad (5.11)$$

A figura 5.1 mostra as formas dos potenciais que minimizam  $\Gamma(\beta)$  para  $\beta$  variando de 0.1 a 50. O aumento da temperatura causa um afastamento da solução harmônica  $\omega(x) = x$  em direção a uma solução mais abrupta para a forma do potencial.

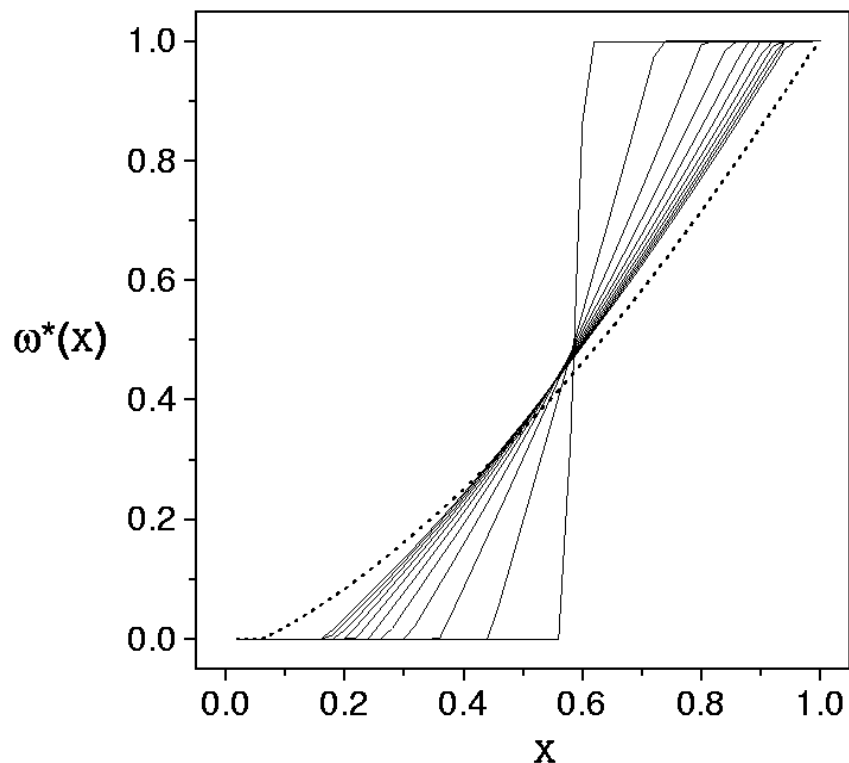
No limite de grandes temperaturas é possível obter um expressão analítica para  $\omega^*(x)$ . A equação (5.10) torna-se

$$\langle t \rangle \cong N^2 \left[ \frac{2}{3} - \beta \int_0^1 dx (3x^2 - 1)\omega(x) \right] \quad (5.12)$$

e a condição do tempo mínimo implica que  $\omega^*(x)$  seja uma função degrau com o degrau situado em  $x = 1/\sqrt{3}$ , ou  $\omega^*(x) = \Theta(x-1/\sqrt{3})$ . O tempo ótimo é  $\langle t \rangle^* = N^2(2/3 - \beta 2/(3\sqrt{3}))$  contra um  $\langle t \rangle_h = N^2(2/3 - \beta 1/4)$  para  $\omega(x) = x$ .

A figura 5.2 mostra o comportamento de  $\Gamma(\beta)$  para quatro escolhas de potencial. A solução degrau mencionada anteriormente (linha pontilhada) tem seu melhor desempenho para pequenos  $\beta$  mas é rapidamente suplantada pela solução harmônica (linha tracejada).



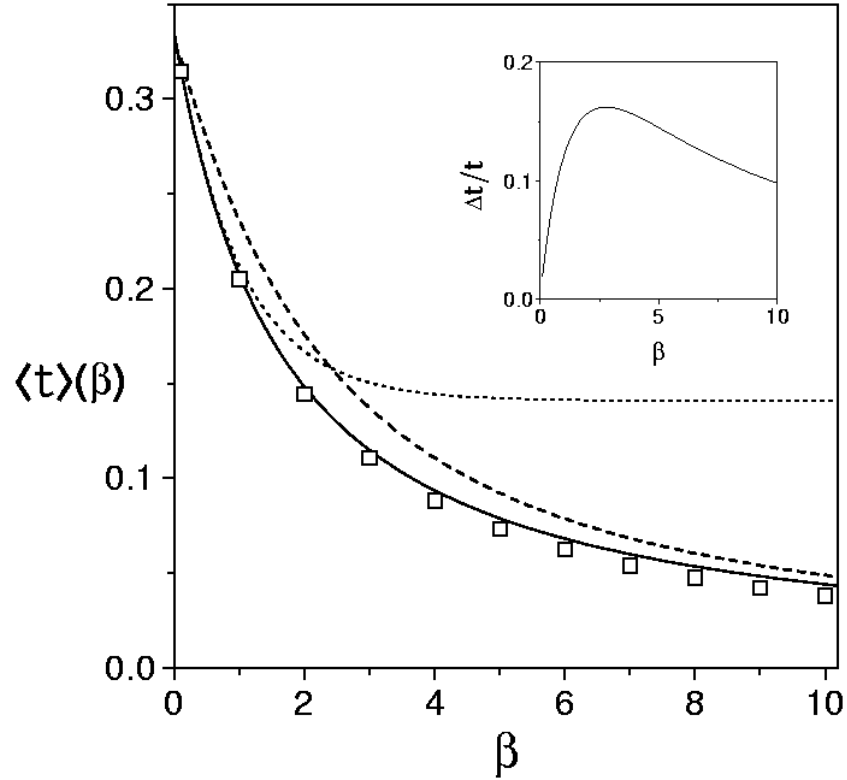


**Fig. 5.1:** Potencial ótimo obtido por simulações de Monte Carlo para  $N=50$ . Iniciando a partir da curva mais acentuada temos  $\beta = 0.1$  e a seqüência  $\beta = 1, 2, 3, \dots, 10$ . A linha pontilhada corresponde a  $\beta = 50$ .

O potencial ótimo obtido numericamente (símbolos), veja figura 5.1, pode ser aproximado por uma função linear que é a interpolação entre a função degrau e a solução harmônica produzindo um resultado que pode ser comparado (linha cheia).

Como era de se esperar para grandes  $\beta$  a forma do potencial não é importante. Para baixas temperaturas todos os  $w(x)$  dão o mesmo resultado para o caso unidimensional.

A otimização do potencial mostra uma economia máxima no tempo na ordem de 15% como podemos ver no detalhe da figura 5.2 em comparação entre solução harmônica e a solução aproximada por uma função linear. Apesar do pequeno ganho, isto nos permite entender a natureza da solução ótima que é propiciar barreiras de transição localizadas para a partícula. Isto é uma boa estratégia porque, para altas temperaturas, a partícula não é tão sensível a pequenos valores do gradiente do potencial. Portanto, o melhor é concentrar altos gradientes em determinadas posições.

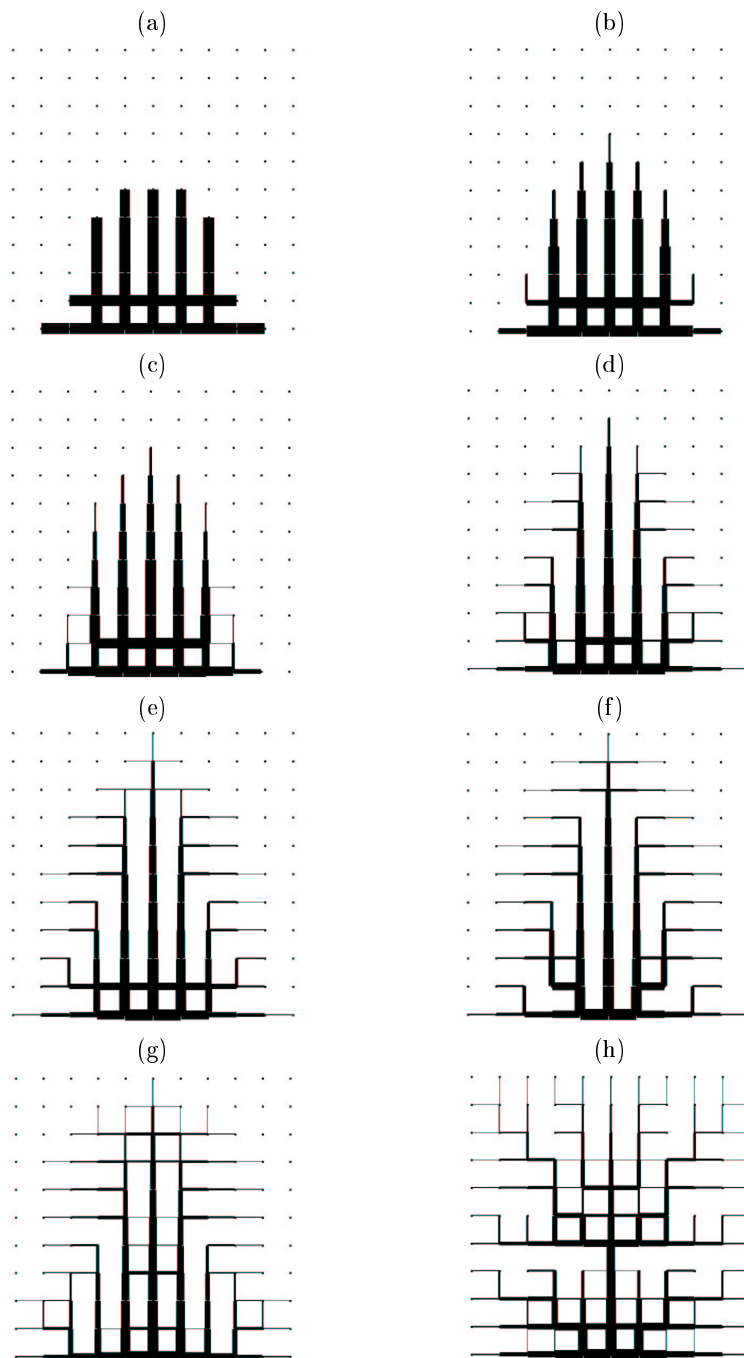


**Fig. 5.2:** O fator de tempo  $\Gamma$  como função do parâmetro de ruído  $\beta$ . A linha tracejada corresponde ao potencial linear  $\omega(x) = x$ . A linha pontilhada é para  $\omega(x) = \Theta(x - 1/\sqrt{3})$ . A linha contínua corresponde a uma interpolação entre estes dois potenciais que é dada por uma função linear fragmentada da forma  $\omega(x) = 0$  se  $x < \frac{1}{\sqrt{3}} \frac{2}{2+\beta}$ ,  $\omega(x) = \frac{(2+\beta)}{2} (x - \frac{1}{\sqrt{3}}) + \frac{1}{\sqrt{3}}$  se  $\frac{1}{\sqrt{3}} \frac{2}{2+\beta} < x < \frac{1}{\sqrt{3}} \frac{2+\sqrt{\beta}}{2+\beta}$ , e  $\omega(x) = 1$  se  $x > \frac{1}{\sqrt{3}} \frac{2+\sqrt{\beta}}{2+\beta}$ . Os ( $\square$ ) representam a solução ótima obtida pelo método de Monte Carlo para  $N=50$ . Os potenciais linear e fragmentado fazem  $\Gamma$  anular-se com  $1/(2\beta)$  para grandes  $\beta$ , as simulações numéricas parecem indicar uma dependência similar.

### 5.3 Problema bidimensional

A figura 5.3 mostra oito configurações otimizadas de ligações para diferentes temperaturas de um passeio bidimensional em uma rede quadrada. Os valores das ligações, variando entre 0 e 1, são representados pelas espessuras das linhas. Como é indicado na equação (5.2), a partícula tende a deixar um determinado sítio usando a ligação mais espessa dentre as

quatro possíveis ligações entre os sítios vizinhos.



**Fig. 5.3:** *Potencial ótimo obtido através de simulações de Monte Carlo para o caso bidimensional para uma rede quadrada com  $N=11 \times 11$  sítios. O sítio absorvedouro (alvo) está centralizado na borda inferior. (a)  $\beta = 0.1$  (b)  $\beta = 1.0$  (c)  $\beta = 2.0$ , (d)  $\beta = 4.0$  (e)  $\beta = 6.0$  (f)  $\beta = 8.0$ , (g)  $\beta = 10.0$  e (h)  $\beta = 1000.0$ .*

A partir de uma análise cuidadosa da figura concluímos que a estrutura ótima para o potencial em duas dimensões é governado por três fatores ou princípios.

O primeiro está relacionado com a dependência da dimensão do passeio aleatório. Uma propriedade conhecida da difusão de partículas em um espaço Euclidiano é que a raiz quadrática média do deslocamento em um tempo  $t$  é independente da dimensão, mais especificamente  $\langle r^2 \rangle \sim t$ . O que pode, à primeira vista, parecer surpreendente, pois inocentemente poderíamos pensar que o aumento dos graus de liberdade do problema dificultaria o movimento da partícula a partir de sua posição inicial. Apesar disso, a intuição funciona bem quando questionamos sobre o tempo necessário para alcançar um determinado ponto. Neste caso vemos que o número de dimensões é importante [57].

No nosso problema queremos guiar a partícula em direção ao alvo o mais rápido possível. Podemos aumentar esta velocidade se reduzirmos a dimensionalidade do passeio. Isto pode ser realizado através do aumento da rugosidade do potencial nas direções perpendiculares ao movimento desejado. Isto é observado em todas as configurações mostradas na figura 5.3. O alvo está localizado no centro da borda inferior do ambiente, por causa disso é que a partícula deve alcançar o centro do ambiente e descer em direção ao alvo. Como resultado do processo de otimização vemos um sistema de “canais” com linhas paralelas que vão em direção ao centro ou para baixo.

O segundo fator é o efeito condutor da inclinação interna do canal. Isto já foi discutido no caso unidimensional. A eficiência do gradiente interno do canal que guia a partícula na direção do alvo varia com o ruído. De fato, para obter o melhor desempenho o potencial deve adaptar-se ao ruído. Quando aumenta-se o ruído, a melhor estratégia é concentrar altos gradientes em posições chaves. Na figura 5.3 isto pode ser observado através da retração e alargamento dos canais.

O terceiro fator é a minimização da trajetória. Para ser eficiente o sistema de canais deve ser projetado para fornecer o mínimo trajeto até o alvo. Em nossas simulações a escolha de uma rede quadrada com primeiros vizinhos é um pouco inadequada porque isto gera uma grande degenerescência da trajetória mínima. Esta é a razão pela qual as estruturas na figura 5.3 são diferentes das estruturas usuais de rios que são feitas em redes hexagonais ou com conexões de segundos vizinhos, veja por exemplo [58].

A figura 5.4 mostra como o tempo médio de alcançar o alvo varia com diferentes configurações de potencial. Como anteriormente, comparamos diferentes estruturas em diferentes níveis de ruído. As estruturas apresentadas são: a configuração de ligações para  $\beta = 0.1$  (linha pontilhada) que representa a estrutura ótima para altas temperaturas, a configuração de ligações para  $\beta = 4.0$  (linha cheia) que representa um estágio intermediário, a configuração de ligações para  $\beta = 100.0$  (linha ponto-tracejada) que representa as baixas temperaturas e uma configuração de ligações que representam uma solução do tipo harmônica (linha tracejada). Neste caso vemos que para  $\beta = 4.0$  existe um ganho considerável da ordem de 40% quando escolhemos a solução ótima para esta temperatura em vez do potencial do tipo harmônico.

## 5.4 Conclusões

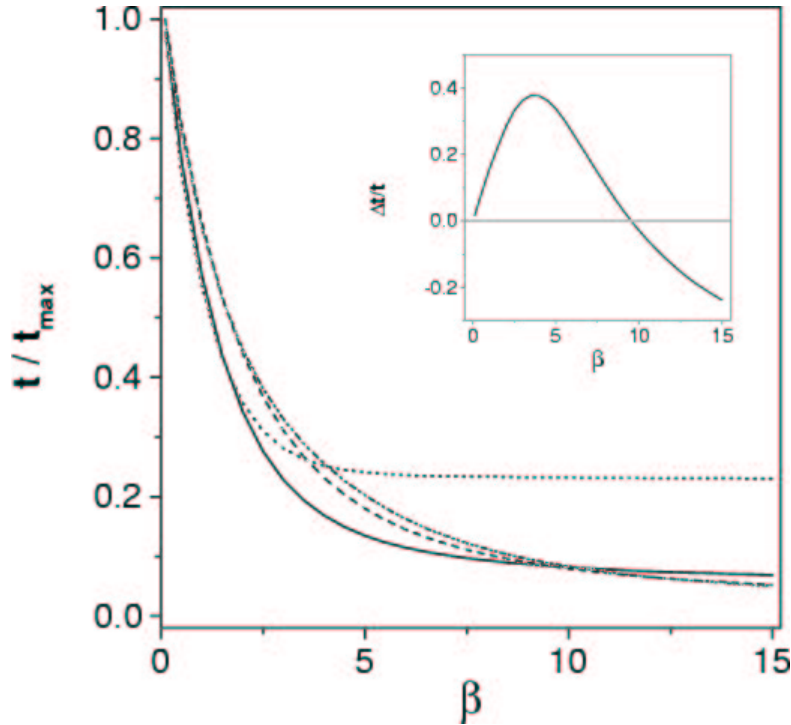
Neste capítulo apresentamos as simulações realizadas em ambientes discretos que mostram que a superfície potencial ótima para uma partícula que é guiada na direção de um objetivo é uma estrutura ramificada que lembra a estrutura de um sistema de rios.

O resultado pode ser sistematizado em termos de três princípios. O primeiro é a redução da dimensionalidade do passeio. Este princípio é responsável pela surgimento de estruturas semelhantes a rios que são armadilhas unidimensionais que reduzem a dimensionalidade do passeio de duas para uma dimensão e portanto reduzem o tempo que a partícula leva para ser absorvida.

O outro princípio é o princípio condutor que diz qual é a melhor escolha para o gradiente dentro do canal que minimiza o tempo passagem para um determinado nível de ruído  $\beta$ . Se restringirmos o potencial entre 0 e 1, diminuindo a região de grandes gradientes, teremos um achatamento do potencial nas regiões longe dos alvos e dos obstáculos. Mas a diminuição do gradiente nestas regiões é compensada quando o agente passar de um sítio para o outro dentro desta região.

O último princípio está relacionado com a minimização do comprimento do rio e tende a orientar os rios o máximo possível na direção dos alvos.

É válido mencionar que os princípios listados acima não se assemelham ao princípios



**Fig. 5.4:** Valor relativo do tempo médio da primeira passagem em uma rede quadrada com uma rede bidimensional  $N=11 \times 11$  em função  $\beta = 1/T$ . A linha pontilhada corresponde a configuração ótima de ligações para  $\beta = 0.1$ . A linha contínua é para  $\beta = 4$  e a linha ponto-tracejada é para  $\beta = 100$ . A linha tracejada corresponde a configuração harmônica para as ligações. Nesta configuração as ligações horizontais crescem linearmente a partir do zero, nas bordas da esquerda e da direita, para um valor máximo no centro; e as ligações verticais crescem linearmente a partir do zero na borda superior até um valor máximo na borda inferior, de tal forma que as ligações que conectam-se diretamente ao sítio absorvedouro (alvo) tenham valor um.  $t_{max}$  é o tempo médio que a partícula leva para alcançar o alvo em um passeio não direcionado.

responsáveis pela formação de rios no contexto da dissipação de energia em sistemas de rios reais [59, 60]. A partir destes estudos, é possível derivar uma função energia que dependa de dois fatores: o comprimento do canal e o fluxo no canal. A redução do comprimento do canal através de conexões entre os sítios vizinhos próximos resulta em altos fluxos em alguns canais, enquanto que baixos fluxos implicam em longos canais conectando sítios distantes. A

estrutura ramificada surge naturalmente quando um compromisso entre as duas tendências é obtido [58].

O que impede uma solução onde há uma série de trajetórias dirigindo-se radialmente ao alvo é a discretização do ambiente. No espaço contínuo não há limite para a rugosidade de uma superfície. A discretização é um artifício para os modelos de rios mas não para um mapa de navegação. Um mapa de navegação deve ser possível de se calcular em tempo real o que torna a discretização ingrediente necessário.

Para o problema de navegação a conclusão é que a estrutura ramificada, como um fractal, que surge dentro do contexto dos modelos de crescimento [61] bem como na organização dos canais dos sistemas de rios pode ter aplicação prática, pois é capaz de produzir mapas que podem contornar o problema de imprecisões que existem na navegação.

O resultado também pode ter implicações no entendimento do processo biológico de navegação. Conforme vimos anteriormente, o campo vetorial armazenado nas conexões sinápticas das células do hipocampo pode ser atualizado enquanto o animal anda no ambiente. No capítulo 3 mostramos que uma otimização deste processo produz um campo vetorial que tem uma estrutura ramificada. Estas estruturas surgem de uma realimentação feita na probabilidade da escolha de uma trajetória. A frequência de passagens por uma posição com a mesma direção de movimento causa um aumento na probabilidade da mesma direção ser escolhida na próxima vez. Este reforço lembra algumas classes de modelos de sistemas rios, onde o fluxo de água esculpe a bacia do rio e ao mesmo tempo é determinado pelos gradientes no solo da bacia do rio [48, 62].

# Conclusões e Perspectivas Futuras

Neste trabalho apresentamos um estudo, organizado em diferentes aspectos, de como que o cérebro atua no processo de aprendizado de trajetórias. Mais especificamente, estudamos como que as informações necessárias para navegar em um ambiente, inicialmente desconhecido, estão estruturadas no hipocampo dos roedores. Vimos que este processo pode ser simplificado e equacionado em um modelo matemático que pode ser utilizado no projeto de sistemas de navegação autônomos, como no modelo proposto por Blum e Abbott.

No capítulo 2 analisamos o modelo proposto por Blum e Abbott onde eles mostram que o aprendizado de Hebb ou o modelo fisiológico equivalente (modelo de LTP) [29] agindo sobre as células do hipocampo é capaz de gerar um deslocamento na posição estimada que é fornecido por um mapa codificado em uma rede de células de posição existente no hipocampo [13, 24, 18]. Este deslocamento quando comparado a posição real indica a direção a ser seguida. O campo vetorial que surge da posição deslocada é armazenado nas conexões sinápticas das células do hipocampo que podem ser atualizadas enquanto o animal explora o ambiente.

No capítulo 3 mostramos que a existência de objetivos no ambiente permite a formação de um mapa de direções preferenciais, semelhante ao mapa de navegação codificado em uma rede de células de posição do hipocampo, devido a quebra de simetria que o objetivo causa na estatística do movimento aleatório no ambiente. Ainda no capítulo 3, mostramos que uma otimização no processo de abordagem do alvo através da incorporação das informações obtidas em explorações anteriores no mapa atual. Esta otimização diminui o tempo total necessário para a formação do mapa, mas o ganho obtido em tempo de formação do mapa pode ser perdido no momento da utilização deste mapa, pois esta forma de aprendizado pode produzir mapas que são completamente inúteis devido a existência de loops. Este



problema pode ser contornado se modularmos o tamanho da contribuição da trajetória que será incorporado ao mapa dependendo do tempo gasto para encontrar o alvo. Ajustando este parâmetro que controla quanto tempo no passado uma trajetória deve ser incorporada ao mapa, podemos reduzir alguns dos efeitos do movimento aleatório sobre as primeiras versões do mapa e evitar os loops.

No capítulo 4 apresentamos um modelo de rede neural capaz de armazenar e recuperar mapas de navegação para diferentes ambientes. O projeto desta rede neural foi baseado no fato que um mapa de navegação pode ser descrito como o gradiente de uma função potencial. A grande vantagem desta abordagem é o baixo número de sinapses da rede. É importante ressaltar que este baixo número de sinapses não reduz o desempenho da rede, porque os padrões que interessam ser armazenados não são aleatórios. Estes padrões são soluções da equação de Laplace para as condições de contorno de Dirichlet. As condições de contorno que descrevem o ambiente determinam completamente todo o padrão de atividade das células de posição, com isso é necessário armazenar apenas as condições de contorno.

No capítulo 5 propomos que o problema de navegação de um robô como sendo uma partícula difundindo em uma superfície potencial com um único ponto de mínimo. O nível de erro deste sistema foi modelado como uma temperatura. Dentro dessas condições, nosso objetivo principal era saber qual é a forma do potencial que minimizava o tempo médio da primeira passagem da partícula para alcançar a posição do objetivo (ponto de mínimo) para uma dada temperatura. Os resultados mostraram que superfície potencial tem uma estrutura ramificada. A frequência de passagens por uma posição com a mesma direção de movimento causa um aumento na probabilidade da mesma direção ser escolhida na próxima vez. Este reforço assemelhasse muito com alguns modelos de sistemas rios encontrados na literatura [48, 62].

Futuramente, pretendemos estudar o surgimento de células de posição e fazer uma conexão entre o campo potencial e as células de posição. Também, pretendemos obter uma equação matemática que gere a estrutura ramificada do campo potencial que surge quando se faz a otimização no processo de abordagem do objetivo. Além disso, pretendemos calcular a capacidade máxima de armazenagem de ambientes em uma rede neural semelhante a que foi proposta no capítulo 4

## Apêndice A

# Alteração do vetor população induzida pelo treinamento

Assumimos que o ajuste da sinápses aumenta a taxa de disparo do neurônio pós-sináptico  $i$  de uma quantidade que é dada pelo produto do peso sináptico  $w_{ij}$  pela taxa de disparo do correspondente neurônio pré-sináptico  $j$ , onde  $w_{ij}$  é dado pela equação (2.5). O efeito total para um dado neurônio pós-sináptico  $i$  é obtido pela soma de todos os correspondentes neurônios pré-sinápticos  $j$ . Após o treinamento, a atividade neurônio  $i$  é dada pela equação (2.6). Para calcular como as alterações das atividades dos neurônios induzidas pelo treinamento afetam o valor do vetor população definido em (2.7) inserimos a atividade após o treinamento dada pela equação (2.6) na equação (2.7)

$$\mathbf{p} = \frac{\sum_i \left( A_i G(|\mathbf{x} - \mathbf{x}_i|) + \sum_j w_{ij} A_j G(|\mathbf{x} - \mathbf{x}_j|) \right) \mathbf{x}_i}{\sum_i \left( A_i G(|\mathbf{x} - \mathbf{x}_i|) + \sum_j w_{ij} A_j G(|\mathbf{x} - \mathbf{x}_j|) \right)} \quad (\text{A.1})$$

Assumimos que antes do aprendizado ( quando  $w_{ij} = 0$ ), o vetor população representa a localização do animal

$$\mathbf{p} = \frac{\sum_i \mathbf{x}_i A_i G(|\mathbf{x} - \mathbf{x}_i|)}{\sum_i A_i G(|\mathbf{x} - \mathbf{x}_i|)} = \mathbf{x}. \quad (\text{A.2})$$

Expandindo o denominador da equação (A.1) até a ordem linear em  $w$  obtemos

$$\mathbf{p} = \mathbf{x} + \frac{\sum_{ij} (\mathbf{x}_i - \mathbf{x}) w_{ij} s_j(\mathbf{x})}{\sum_i s_i(\mathbf{x})}. \quad (\text{A.3})$$

De acordo com a equação (2.9) temos

$$\Delta \mathbf{p}(\mathbf{x}) = \frac{\sum_{ij} (\mathbf{x}_i - \mathbf{x}) w_{ij} s_j(\mathbf{x})}{\sum_i s_i(\mathbf{x})}. \quad (\text{A.4})$$

## Apêndice B

# Princípio do máximo das funções harmônicas

O princípio do máximo das funções harmônicas é consequência de uma propriedade das funções harmônicas que diz que o valor de uma função harmônica em um ponto  $P$  é igual a média aritmética dos valores desta função em uma esfera qualquer centrada em  $P$ , desde que esta função seja regular e contínua dentro da região limitada pela esfera.

**Princípio do Máximo:** *Considere  $u$  uma função regular e harmônica em uma região conectada  $G$  e contínua até e sobre o contorno  $\Gamma$ . Então o máximo e o mínimo desta função são assumidos sempre sobre o contorno; o máximo e o mínimo são assumidos no interior se e somente se  $u$  é uma constante.*

Para provar este teorema consideramos um subconjunto  $F$  do domínio fechado  $G + \Gamma$  constituído dos pontos nos quais  $u$  assume seu valor máximo  $M$  em  $G + \Gamma$ . Sendo  $u$  contínuo no domínio  $G + \Gamma$ ,  $F$  é um conjunto fechado. Se  $F$  contém um ponto interno  $P_0$  de  $G$ , então existe uma família de esferas centradas em  $P_0$  todas contidas em  $G$ , e o valor médio da função  $u$  em cada um destas esferas é  $u(P_0) = M$ . Mas já que  $u \leq M$ , isto é possível somente se  $u = M$  em toda parte das esferas centradas em  $P_0$  e contidas inteiramente em  $G$ . Portanto,  $F$  contém pelo menos todos os pontos dentro da maior esfera possível centrada em  $P_0$  contida inteiramente em  $G + \Gamma$ . O mesmo argumento pode ser repetido para qualquer outro ponto interno de  $F$ , por isso  $F$  deve coincidir com  $G + \Gamma$ . Mas isto significa que  $u$  é constante (igual a  $M$ ). Portanto, para qualquer  $u$  que não seja constante,  $F$  pode ser composto apenas dos pontos do contorno. Um argumento correspondente mostra que o

mínimo  $m$  é assumido sobre o contorno e apenas para  $u$  constante é assumido no interior de  $G$ . Uma consequência direta deste teorema é que se existe uma função regular e harmônica em  $G$  e contínua em  $G + \Gamma$  que é constante sobre o contorno, então esta função é contínua em toda a região <sup>1</sup>.

---

<sup>1</sup> Para maiores detalhes sobre funções harmônicas consulte [63]

# Bibliografia

# Referências Bibliográficas

- [1] BROOKS, R. A. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, New York, v. 2, n. 1, p. 14–23, Mar. 1986.
- [2] BEER, R. D. *Intelligence as adaptive behavior: an experiment in computational neuroethology*. San Diego: Academic Press, 1990.
- [3] McFARLAND, D. J.; BOESSER, T. *Intelligent behavior in animals and robots*. Cambridge: MIT Press, 1993.
- [4] NEHMZOW, U. Animal and robot navigation. *Robotics and Autonomous Systems*, Amsterdam, v. 15, n. 1/2, p. 71–81, July 1995.
- [5] HARTMANN, G.; WEHNER, R. The ants path integration system: a neural architecture. *Biological Cybernetics*, New York, v. 73, n. 6, p. 483–497, Nov. 1995.
- [6] PFEIFER, R. Cognition-perspectives from autonomous agents. *Robotics and Autonomous Systems*, New York, v. 15, n. 1/2, p. 47–70, July 1995.
- [7] O'KEEFE, J.; NADEL, L. *The hippocampus as a cognitive map*. London: Clarendon Press, 1978.
- [8] BATTAGLIA, F. P.; TREVES, A. Attractor neural networks storing multiple space representations: a model for hippocampal place fields. *Physical Review E*, Woodbury, v. 58, n. 6, p. 7738–7753, Dec. 1998.
- [9] BURGESS, N.; RECCE, M.; O'KEEFE, J. A model of hippocampal function. *Neural Networks*, New York, v. 7, n. 6/7, p. 1065–1081, 1994.

- [10] BLUM, K. L.; ABBOTT, L. F. A model of spatial map formation in the hippocampus of the rat. *Neural Computation*, Cambridge, v. 8, n. 1, p. 85–93, Jan. 1996.
- [11] ARLEO, A.; GERSTNER, W. Spatial cognition and neuro-mimetic navigation: A model of hippocampal place cell activity. *Biological Cybernetics*, New York, v. 83, n. 3, p. 287–299, Sept. 2000.
- [12] McCULLOCH, W. S.; PITTS, W. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, Oxford, v. 5, p. 115–133, 1943.
- [13] O'KEEFE, J.; DOSTROVSKY, J. The hippocampus as spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, Amsterdam, v. 34, n. 1, p. 171–175, 1971.
- [14] O'KEEFE, J.; NADEL, L.; KEIGHTLEY, S.; KILL, D. Fornix lesions selectively abolish place learning in the rat. *Experimental Neurology*, San Diego, v. 48, n. 1, p. 152–166, 1975.
- [15] OLTON, D. S.; WALKER, J. A.; GAGE, H. Hippocampal connections and spatial discrimination. *Brain Research*, Amsterdam, v. 139, n. 2, p. 295–308, 1978.
- [16] MORRIS, R. G. M.; GARRUD, P.; RAWLINS, J. N. P.; O'KEEFE, J. Place navigation impaired in rats with hippocampal lesions. *Nature*, London, v. 297, n. 5868, p. 681–683, June 1982.
- [17] TAUBE, J. S.; MULLER, R. U.; RANCK, J. B. Head-direction cells recorded from the postsubiculum in freely moving rats: description and quantitative analysis. *Journal of Neuroscience*, Washington, v. 10, n. 2, p. 420–435, Feb. 1990.
- [18] WILSON, M. A.; McNAUGHTON, B. L. Dynamics of hippocampal ensemble code for space. *Science*, Washington, v. 261, n. 5124, p. 1055–1058, Aug. 1993.
- [19] O'KEEFE, J.; BURGESS, N. Geometric determinants of the place fields of hippocampal neurons. *Nature*, London, v. 381, n. 6581, p. 425–428, May 1996.



- [20] PARKINSON, J. K.; MURRAY, E. A.; MISHKIN, M. A selective mnemonic role for the hippocampus in monkey: memory for the location of objects. *Journal of Neuroscience*, Washington, v. 8, n. 11, p. 4159–4167, Nov. 1988.
- [21] MURRAY, E. A.; DAVIDSON, M.; GAFFAN, D.; OLTON, D. S.; SUOMI, S. Effects of fornix transection and cingulate cortical ablation on spatial memory in rhesus monkeys. *Experimental Brain Research*, New York, v. 74, n. 1, p. 173–186, 1989.
- [22] ONO, T.; NAKAMURA, K.; FUKUDA, M.; TAMURA, R. Place recognition responses of neurons in monkey hippocampus. *Neuroscience Letters*, Clare, v. 121, n. 1/2, p. 194–198, Jan. 1991.
- [23] McNAUGHTON, B. L.; CHEN, L. L.; MARKUS, E. J. Dead reckoning, landmark learning, and the sense of direction: a neurophysiological and computational hypothesis. *Journal of Cognitive Neuroscience*, Cambridge, v. 3, n. 2, p. 190–196, Sept. 1991.
- [24] MULLER, R. U.; KUBIE, J. L.; SAYPOFF, R. The hippocampus as a cognitive graph. *Hippocampus*, New York, v. 1, n. 3, p. 243–246, July 1991.
- [25] TRAUB, R. D.; MILES, R.; MULLER, R. U.; GULYAS, A. I. Functional organization of the hippocampal ca3 regions: implications for epilepsy, brain waves and spatial behavior. *Network: Computation in Neural Systems*, Bristol, v. 3, n. 4, p. 465–488, Nov. 1992.
- [26] WORDEN, R. Navigation by fragment fitting: a theory of hippocampal function. *Hippocampus*, New York, v. 2, n. 2, p. 165–188, Apr. 1992.
- [27] WAN, H. S.; TOURETZKY, D. S.; REDISH, A. D. Towards a computational theory of rat navigation. In: CONNECTIONIST MODELS SUMMER SCHOOL, 1993. *Proceedings of the ...* Hillsdale: Lawrence Erlbaum Associates, 1994. p. 11–19
- [28] BEAR, M. F.; CONNORS, B. W.; PARADISO, M. A. *Neuroscience: exploring the brain*. Philadelphia: Williams and Wilkins, 1996.

- [29] ROLLS, E. T.; TREVES, A. *Neural networks and brain function*. Oxford: Oxford University Press, 1998.
- [30] HEBB, D. *The organization of behavior*. New York: John Wiley, 1949.
- [31] ABBOTT, L. F.; BLUM, K. L. Functional significance of long-term potentiation for sequence learning and prediction. *Cerebral Cortex*, Cary, v. 6, n. 3, p. 406–416, May/June 1996.
- [32] LEVY, W. B.; STEWARD, D. Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience*, Oxford, v. 8, n. 4, p. 791–797, 1983.
- [33] GUSTAFSSON, B.; WIGSTROM, H.; ABRAHAM, W. C.; HUANG, Y. Y. Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. *Journal of Neuroscience*, Washington, v. 7, n. 3, p. 774–780, Mar. 1987.
- [34] LEVY, W. B. A computational approach to hippocampal function. In HAWKINS, R. D.; BOWER, G. H. (Eds.) *Computational models of learning in simple neural systems*. San Diego: Academic Press, 1989. p. 243–305.
- [35] DEBANNE, D.; GAHWILER, B. H.; THOMPSON, S. M. Asynchronous pre- and postsynaptic activity induces associative long-term depression in area ca1 of the rat hippocampus in vitro. *Proceedings of the National Academy of Sciences USA*, Washington, v. 91, n. 3, p. 1148–1152, Feb. 1994.
- [36] MARKRAM, H.; SAKMANN, B.; HELM, P. J. Dendritic calcium transients evoked by single back-propagating action-potentials in rat neocortical pyramidal neurons. *Journal of Physiology*, Cambridge, v. 485, n. 1, p. 1–20, May 1995.
- [37] GERSTNER, W.; ABBOTT, L. F. Learning navigational maps through potentiation and modulation of hippocampal place cells. *Journal of Computational Neuroscience*, Dordrecht, v. 4, n. 1, p. 79–94, Jan. 1997.

- [38] GEORGOPOULOS, A. P.; SCHWARTZ, A.; KETTNER, R. E. Neuronal population coding of movement direction. *Science*, Washington, v. 233, n. 4771, p. 1416–1419, Sept. 1986.
- [39] MULLER, R. U.; KUBIE, J. L. The firing of hippocampal place cells predicts the future position of freely moving rats. *Journal of Neuroscience*, Washington, v. 9, n. 12, p. 4101–4110, Dec. 1989.
- [40] SALINAS, E.; ABBOTT, L. F. Vector reconstruction from firing rates. *Journal of Computational Neuroscience*, Dordrecht, v. 1, n. 1, p. 89–107, Jan. 1994.
- [41] CONNOLLY, C. I.; GRUPEN, R. A. On the applications of harmonic functions to robotics. *Journal of Robotic Systems*, New York, v. 10, n. 7, p. 931–946, Oct. 1993.
- [42] CONNOLLY, C. I. *Harmonic functions as a basis for motor control an planning*. 1994. 111 f. Thesis. (PhD Computer Science) - University of Massachusetts, Amherst, May 1994.
- [43] CONNOLLY, C. I. Harmonic functions and collision probabilities. *International Journal of Robotic Research*, Thousand Oaks, v. 16, n. 4, p. 497–507, Aug. 1997.
- [44] KNAPP, A. W. Connection between brownian motion and potential theory. *Journal of Mathematical Analysis and Applications*, San Diego, v. 12, n. 2, p. 328–349, Oct. 1965.
- [45] HERSH, R.; GRIEGO, R. J. Brownian motion and potential theory. *Scientific American*, New York, v. 220, n. 3, p. 67–74, Mar. 1969.
- [46] McNAUGHTON, B. L. et al. Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *Journal of Experimental Biology*, Cambridge, v. 199, n. 1, p. 173–185, Jan. 1996.
- [47] DAMASIO A. *Descarte's error*. New York: Avon Books, 1994.
- [48] LEHENY, R. L.; NAGEL, S. R. Model for the evolution of river networks. *Physical Review Letters*, Woodbury, v. 71, n. 9, p. 1470–1473, Aug. 1993.

- [49] LEHENY R. L. Simple model for river network evolution. *Physical Review E*, Woodbury, v. 52, n. 5, p. 5610–5620, Nov. 1995.
- [50] MEHTA, M. R.; BARNES, C. A.; McNAUGHTON, B. L. Experience-dependent, asymmetric expansion of hippocampal place fields. *Proceedings of The National Academy of Sciences USA*, Washington, v. 94, n. 16, p. 8918–8921, Aug. 1997.
- [51] KLEINFELD, D. Sequential state generation by model neural networks. *Proceedings of The National Academy of Sciences USA*, Washington, v. 83, n. 24, p. 9469–9473, Dec. 1986.
- [52] SOMPOLINSKY, H.; KANTER, I. Temporal association in asymmetric neural networks. *Physical Review Letters*, Woodbury, v. 57, n. 22, p. 2861–2864, Dec. 1986.
- [53] GUTFREUND, H.; MEZARD, M. Processing of temporal sequences in neural networks. *Physical Review Letters*, Woodbury, v. 61, n. 2, p. 235–238, July 1988.
- [54] TREVISAN, M.; IDIART, M. A. P.; SILVA Jr., E. P. e. Storage of multiple navigation maps using neural networks. In: CONGRESSO BRASILEIRO DE REDES NEURAIS, 5., 2001, Rio de Janeiro. *Anais do ...* Rio de Janeiro: Conselho Nacional de Redes Neurais, 2001. p. 625–629. 1 Shape CD.
- [55] MERTENS, S.; KOHLER, H. M.; BOS, S. Learning grey-toned patterns in neural networks. *Journal of Physics A: Mathematical and General*, Bristol, v. 24, n. 20, p. 4941–4952, Jan. 1991.
- [56] REICHL, L. E. *A modern course in statistical physics*. 2. ed. New York: John Wiley, 1998.
- [57] BALAKRISHNAN, V. Random walks on fractals. *Materials Science Engineering B*, Lausanne, v. 32, n. 3, p. 201–210, July 1995.
- [58] RINALDO, A. *et al.* Minimum energy and fractal structures of drainage networks. *Water Resource Research*, Washington, v. 28, n. 9, p. 2183–2195, Sept. 1992.

- 
- [59] SUN, T.; MEAKIN, P.; JOSSANG, T. The topography of optimal drainage basins. *Water Resource Research*, Washington, v. 30, n. 9, p. 2599–2610, Sept. 1994.
- [60] SUN, T.; MEAKIN, P.; JOSSANG, T. Minimum energy dissipation river networks with fractal boundaries. *Physical Review E*, Woodbury, v. 51, n. 6, p. 5353–5359, June 1995.
- [61] MEAKIN P. An eden model for randomly branched structures. *Physica Scripta*, Stockholm, v. 45, n. 2, p. 69–74, Feb. 1992.
- [62] GIACOMETTI, A.; MARITAN, A.; BANAVAR, J. R. Continuum model for river networks. *Physical Review Letters*, Woodbury, v. 75, n. 3, p. 577–580, July 1995.
- [63] COURANT, R.; HILBERT, D. *Methods of mathematical physics*. New York: John Wiley, 1962. v. 2.